

Developing guidelines for business intelligence modules in information technology programmes at universities using critical systems heuristics

Theo Mahalepa
12295825

Thesis submitted in fulfillment of the requirements for the degree
Magister Scientiae in Computer Science at the Vaal Triangle
Campus of the North-West University

Supervisor: Prof R. Goede

May 2016

It all starts here TM

Acknowledgements

I thank God almighty for being with me through this educational journey. I could not have completed this study without the help, guidance, and motivation from the following people:

Firstly, I am ineffably indebted to Prof. Roelien Goede for her conscientious guidance and encouragement to accomplish this study. I will also like to thank Carin Venter for her contribution to ensuring that this work of high standard.

Secondly, I am my grateful to the participants of this study who volunteered their time.

Thirdly, I will like to thank Ntate for inspiring me to reach greater heights.

Lastly, and most importantly, I would like to thank my wife, Sibulele Mahalepa, for her endless love, understanding, patience, support, and encouragement during the writing of this dissertation.

Declaration

I, Theo Mahalepa, declare that

Developing guidelines for business intelligence modules in information technology programmes at universities using critical systems heuristics

is my own work and that all the sources I have used or quoted have been indicated and acknowledged by means of complete references.

Signature: _____

Date: _____

Abstract

Data warehousing (DW)/ Business intelligence (BI) is continually being implemented across many industries; however, universities in South Africa face a variety of challenges in training the next generation of DW/BI workers. The complexity of the situation arises from the different goals of the involved and affected stakeholders. The problem addressed by this study is the difficulty traditional universities experience to provide the DW/BI industry with graduates that are able to contribute to the organisation without substantial further training.

A systems approach is recommended and applied to determine how, in this changing environment of DW/BI education practice and with different stakeholders' perspectives, we can improve our understanding of it. Two systems methodologies are chosen as instruments to be used in the research problem. A research paradigm is chosen; the process of how to develop guidelines is also presented to bring clarity on what is a guideline is, and how it is used.

Research aspects are extracted from the DW/BI literature review, and from current DW/BI programmes at universities. These research aspects are used to formulate an interpretive questionnaire leading to an empirical study. A purposeful selection of participants is conducted for this interpretive study. The required information is the understanding of the former students, lecturers, and industry experts. Interpretive coding and cross case analysis are selected as applicable methods for this study. The results of data analysis are used to create guidelines for DW/BI module.

Four sets of guidelines concerning the content and development process for a DW/BI module are presented as findings of this study. The significance of the findings reflects on the current DW/BI problem situation, as well as the ideal.

Further study, recommendations, as well as assumptions and limitations of the study are also presented.

Keywords: Data warehousing (DW), Business intelligence (BI), Guideline, Module, Education

TABLE OF CONTENTS

Acknowledgements	i
Declaration	ii
Abstract	iii
LIST OF FIGURES	xiv
LIST OF TABLES	xvii
Chapter 1 Introduction.....	1
1.1 Introduction	1
1.2 Concepts central to the study	3
1.2.1 Business intelligence	3
1.2.2 Systems thinking	4
1.3 Background of the problem area.....	5
1.4 Study objectives.....	7
1.4.1 Theoretical objectives	7
1.4.2 Empirical objectives:	7
1.5 Overall research methodology	8
1.6 Chapter layout	8
1.6.1 Chapter 2: Research methodology	8
1.6.2 Chapter 3: Business intelligence and data warehousing	8
1.6.3 Chapter 4: Guidelines from DW/BI literature review	9
1.6.4 Chapter 5: Guidelines from institutions of higher education	9
1.6.5 Chapter 6: Systems thinking methodologies	9
1.6.6 Chapter 7: Stakeholders' perspectives on DW/BI education	9
1.6.7 Chapter 8: Stakeholders' perspectives on DW/BI education (continuation)	9
1.6.8 Chapter 9: Summary	10
1.7 Summary.....	10

Chapter 2 Research methodology	13
2.1 Introduction	13
2.2 Guidelines.....	14
2.3 Key principles for developing guidelines	14
2.3.1 Position of the study	16
2.4 Research paradigms	17
2.4.1 Positivism	19
2.4.2 Interpretive research.....	20
2.4.3 Critical social theory	24
2.4.4 Design science	27
2.4.5 Position of the study	29
2.5 Interpretive research methodology.....	30
2.5.1 Principles for interpretive research	30
2.5.1.1 Position of the study (Interpretive principles applied)	32
2.6 Interpretive research methods	33
2.6.1 Participants selection.....	33
2.6.1.1 Position of the study	34
2.6.2 Data collection methods	34
2.6.2.1 Grounded theory as data collection method.....	35
2.6.2.2 Interviews.....	35
2.6.2.3 Ethnography	36
2.6.2.4 Collection of relevant documents	36
2.6.2.5 Position of the study	36
2.6.3 Data analysis.....	37
2.6.3.1 Grounded theory as analysis method.....	37
2.6.3.2 Content analysis with interpretive coding	38
2.6.3.3 Cross case analysis.....	39
2.6.3.4 Position of the study	40
2.6.4 Research evaluation: trustworthiness of the study	40
2.6.5 Position of the study	42
2.6.6 Ethical considerations.....	42
2.7 Concise research plan for this study	43

2.8	Summary.....	46
	Chapter 3 Business intelligence and data warehousing	49
3.1	Introduction	49
3.2	Evolution of business intelligence.....	50
3.3	Business intelligence and analytics	51
3.4	Business intelligence components.....	54
3.4.1	Data.....	54
3.4.1.1	Data Hierarchy.....	55
3.4.1.2	Data storage and retrieval	56
3.4.2	Processes.....	57
3.4.3	Tools and technologies (hardware and software).....	58
3.4.4	Human talent and education.....	58
3.5	Data warehousing	59
3.6	The BI/DW life cycle	61
3.6.1	Project planning and management	62
3.6.2	Requirements collection	62
3.6.3	BI/DW architecture	63
3.6.4	Data warehouse architecture types	63
3.6.4.1	Centralized data warehouse	64
3.6.4.2	Independent data mart.....	64
3.6.4.3	Dependent data warehouse	65
3.6.4.4	Homogeneous data warehouse.....	66
3.6.4.5	Heterogeneous data warehouse	66
3.6.5	Illustrative example of data warehouse concepts.....	68
3.6.6	Source systems.....	69
3.6.7	Extract, transform and load system	73
3.6.7.1	Source to target mapping	74
3.6.7.2	Extraction, transform, and load architecture.....	76
3.6.7.3	Data extraction.....	77
3.6.7.4	Data transformation	78
3.6.7.5	Data loading.....	81
3.6.8	Presentation server	81
3.6.8.1	Illustrative example: Star schema.....	81

3.6.8.2	Dimensional modelling.....	83
3.6.8.3	Dimensions	85
3.6.8.4	Fact tables	91
3.6.8.5	Online analytical processing	92
3.6.8.6	Data presentation (Reporting)	95
3.6.8.7	Traditional data delivery methods.....	96
3.6.8.8	Standard reports	97
3.6.8.9	Performance monitoring (score cards and dashboards)	98
3.6.9	Deployment	98
3.7	Maintenance and growth	99
3.8	Future business intelligence	99
3.8.1	Real time business intelligence.....	100
3.8.2	Big data	100
3.8.3	Mobile business intelligence	100
3.9	Summary.....	101
Chapter 4 Guidelines from DW/BI literature review.....		103
4.1	Introduction	103
4.2	Methodology.....	104
4.2.1	Information technology body of knowledge	105
4.2.2	Identify skills / knowledge according to DW/BI literature review	105
4.3	Information technology body of knowledge.....	106
4.4	Guidelines from DW/BI literature.....	108
4.4.1	Project planning and management	108
4.4.2	Business requirements definition	109
4.4.3	Technical architecture	110
4.4.4	Source systems	111
4.4.5	Extract, transform, and load systems.....	112
4.4.6	Presentation server	118
4.4.7	BI/DW applications	119
4.4.8	Metadata	120
4.4.9	Deployment	121
4.4.10	Growth and maintenance.....	121

4.4.11	Skills from literature review	121
4.4.12	Categorisation of skills from literature review	127
4.5	Summary	137
 Chapter 5 Guidelines from institutions of higher education		139
5.1	Introduction	139
5.2	Methodology	140
5.2.1	Identify skills according programmes/modules of institutions on higher learning	140
5.2.2	Compare the lists of skills from DW/BI literature and skills from universities' programmes and modules.	141
5.2.3	Present a combined list of skills.....	141
5.3	Skills from universities' programme/module content.....	142
5.3.1	Data collection	142
5.3.2	General programme characteristics	144
5.3.3	Data warehouse module characteristics	148
5.3.3.1	Prerequisites	148
5.3.3.2	Module content	148
5.4	Findings: Skills from universities' programmes and modules	152
5.4.1	General IT skills.....	152
5.4.2	DW/BI specific skills from module contents.....	155
5.5	Comparison of university skills and skills from literature.....	156
5.5.1	Comparison of advance general IT skills from literature and skills from universities	156
5.5.2	Comparison of DW/BI specific skills from literature and skills from universities	158
5.6	Guidelines for DW/BI programmes from literature and universities' programmes	158
5.6.1	Prerequisite skills from literature and universities' programmes.....	159
5.6.2	BI specific skills from literature and universities' programmes.....	160
5.6.3	Advanced general IT skills from literature and universities' programmes	161

5.7	Summary.....	162
Chapter 6 Systems thinking methodologies.....		165
6.1	Introduction	165
6.2	Systems and systems thinking	166
6.3	Illustrative problem situation.....	168
6.4	Soft systems methodology.....	169
6.4.1	Soft systems methodologies main activities.....	171
6.4.1.1	Perceived real-world problem situation	172
6.4.1.2	Making relevant purposeful activities.....	173
6.4.1.3	Comparison and structured discussions.....	174
6.4.1.4	Define or take the action to improve the problem situation	174
6.4.2	The SSM learning cycle (making purposeful activity models)	174
6.4.2.1	The PQR formula.....	176
6.4.2.2	CATWOE	176
6.4.2.3	Root definition	178
6.4.2.4	Primary vs issue based model.....	178
6.4.2.5	Purposeful activity models	179
6.4.3	Position of the study	180
6.5	Critical systems thinking	181
6.6	Critical systems heuristics	182
6.6.1	Boundary judgements.....	183
6.6.2	Demonstration of critical systems heuristics	187
6.6.3	Position of the study	189
6.7	Summary.....	189
Chapter 7 Stakeholders' perspectives on DW/BI education.....		193
7.1	Introduction	193
7.2	Interpretive interview design	193
7.2.1	Participants.....	194
7.2.2	Interview questions.....	194

7.2.2.1	Section 1: Biographical information	194
7.2.2.2	Section 2: Stakeholder Identification	195
7.2.2.3	Section 3: Module content	196
7.3	Data collection process	198
7.3.1	Participants.....	198
7.3.2	Interview process.....	200
7.4	Data analysis	201
7.4.1	Preparing the data.....	201
7.4.2	Define units of analysis.....	201
7.4.3	Develop categories of scheme.....	201
7.4.4	Test your coding schemes on a sample of text	202
7.4.5	Code all the text.....	202
7.4.6	Assess the coding consistency	202
7.4.7	Draw Conclusions from the coded data	202
7.4.8	Report your findings	202
7.4.9	Step 1: Preparing the data.....	203
7.4.10	Step 2: Define the unit of analysis	204
7.4.11	Step 3: Develop categories and a coding scheme	204
7.4.12	Step 4: Testing the coding scheme on a sample of text.....	205
7.4.13	Step 5: Code all text	206
7.4.14	Step 6: Assess the coding consistency.....	210
7.5	Draw conclusions from the coded data.....	211
7.5.1.1	Theme 1: DW/BI roles	212
7.5.1.2	Theme 2: DW/BI education preference?	214
7.5.1.3	Theme 3: Beneficiaries	217
7.5.1.4	Theme 4: DW/BI module purpose	219
7.5.1.5	Theme 5: Decision makers	223
7.5.1.6	Theme 6: Conditions outside control of decision makers.....	225
7.5.1.7	Theme 7: Decision makers' powers	227
7.5.1.8	Theme 8: Affected stakeholders.....	230
7.5.1.9	Theme 9: Affected stakeholders' control	232
7.5.1.10	Theme 10: Involved stakeholders.....	234
7.5.1.11	Theme 11: Expertise involved	237
7.5.1.12	Theme 12: Reflected views	239
7.5.1.13	Theme 13: Success measures	241
7.5.1.14	Theme 14: Success assurer (Guarantor)	244

7.5.1.15	Theme 15: Prerequisites.....	247
7.5.1.16	Theme 16: Elective subjects.....	249
7.5.1.17	Theme 17: Assessment component	251
7.5.1.18	Theme 18: Student preparation	252
7.5.1.19	Theme 19: A general perception of what universities are teaching	254
7.5.1.20	Theme 20: DW concepts experience	256
7.5.1.21	Theme 21: What to teach DW/BI students	258
7.5.1.22	Theme 22: Other concerns	261
7.5.1.23	Theme 23: DW/BI processes.....	263
7.5.1.24	Conclusions drawn from DW/BI process	265
7.6	Conclusions from the data	265
7.7	Summary.....	269
 Chapter 8 Stakeholders' perspectives on DW/BI education (continuation)....		271
8.1	Introduction	271
8.2	Critical systems heuristics (boundary findings)	272
8.3	Conclusions drawn from boundary findings.....	274
8.4	Soft systems methodology application	276
8.5	SSM application to this study.....	277
8.6	Student perspective	281
8.7	BI professional perspective.....	287
8.8	Faculty member perspective	292
8.9	Conclusions drawn from SSM activities.....	297
8.10	Summary.....	304
 Chapter 9 Summary		307
9.1	Introduction	307
9.2	Research objectives revisited	307

9.2.1	Theoretical objectives	308
9.2.2	Empirical objectives:	309
9.3	Combined set of guidelines.....	310
9.3.1	First proposed guideline (Prerequisites)	311
9.3.2	Second proposed guideline (DW/BI specific skills)	313
9.3.3	Third proposed guideline (Advance general IT skills).....	314
9.3.4	<i>Fourth proposed guideline (Processes to improve the DW/BI module)</i>	<i>316</i>
9.4	Recommendations for further study	319
9.5	Assumptions and limitations of the study.....	320
9.6	Summary.....	321

References	322

LIST OF FIGURES

Figure 2-1: Properties of a guideline	15
Figure 2-2: Paradigm framework.....	18
Figure 2-3: Illustration: summary of positivism.....	20
Figure 2-4: Illustration: summary of interpretive research.....	23
Figure 2-5: Illustration: summary of critical social theory	26
Figure 2-6: Illustration: summary of design science.....	28
Figure 2-7: Illustration: inductive approach.....	39
Figure 2-8: Illustration: deductive approach.....	39
Figure 2-9: Research plan.....	44
Figure 3-1: Business intelligence and analytics usage (Mirji, 2011).....	52
Figure 3-2: Analytics categories	54
Figure 3-3: Data hierarchy (Clark, 2004b)	56
Figure 3-4: Kimball DW/BI life cycle diagram (Kimball <i>et al.</i> , 2008a:3)	62
Figure 3-5: Centralised DW architecture	64
Figure 3-6: Independent data mart architecture.....	65
Figure 3-7: Dependent DW architecture	65
Figure 3-8: Homogeneous DW architecture	66
Figure 3-9: Heterogeneous DW architecture	67
Figure 3-10: High level DW/BI system architecture model (Kimball <i>et al.</i> , 2008d:101).....	67
Figure 3-11: OLTP vs OLAP logical organisation of data	69
Figure 3-12: Illustrative Example: Entity relationship diagram	70
Figure 3-13: Representing a relationship on an entity-relationship diagram	71
Figure 3-14: Three ways of thinking of a relationship	72
Figure 3-15: Fragmented ETL data-flow.....	73
Figure 3-16: Illustrative example: ETL workflow	74
Figure 3-17: Illustrative example: visual representation of source to target mapping.....	76
Figure 3-18: Illustrative example: ETL architecture	77
Figure 3-19: Illustrative example: Connectivity	78
Figure 3-20: Illustrative example: Attribute case and space transformation	79
Figure 3-21: Illustrative example: Cell phone number standardisation.....	79
Figure 3-22: Illustrative example: SQL scrip to convert number to monetary value	79
Figure 3-23: Illustrative example: MS SQL statement to remove unwanted spaces	80
Figure 3-24: Illustrative example: Merge data records.....	80
Figure 3-25: Illustrative example: Transaction perspectives (dimensions)	82
Figure 3-26: Illustrative example: Telecommunication transactions star schema.....	83

Figure 3-27: Illustrative example: snow flake schema	85
Figure 3-28: Illustrative example: FactTransaction and Dim_PricePlan Relationship.....	86
Figure 3-29: Illustrative example: Dim_Time	86
Figure 3-30: Illustrative example: Time hierarchy	87
Figure 3-31: Illustrative example: SCD type one	88
Figure 3-32: Illustrative example: SCD type two.....	89
Figure 3-33: Illustrative example: SCD type three	89
Figure 3-34: Illustrative example: SCD type four	90
Figure 3-35: Illustrative example: Fact transaction	91
Figure 3-36: Illustrative example: Summary table: granularity: monthly call category	92
Figure 3-37: Illustrative example: Transaction cube	93
Figure 3-38: Illustrative example: Roll up and down	94
Figure 3-39: Illustrative example: Cube slice and dice.....	94
Figure 3-40: Illustrative example: Self-service BI (Microsoft Excel pivot)	97
Figure 3-41: Illustrative example: Report parameters	98
Figure 3-42: Illustrative example: Dashboard	98
Figure 4-1: Kimball lifecycle diagram (Kimball et al., 2008a:3)	108
Figure 4-2: High level DW/BI system architecture model (Kimball et al., 2008b:101)	111
Figure 6-1: Adaptive system (Adapted from Checkland and Poulter (2006:203))	167
Figure 6-2: The hard and soft systems stances (Checkland, 2000:19)	170
Figure 6-3: SSM learning cycle (Adopted from Checkland and Poulter (2006:207))	171
Figure 6-4: DW/BI Education problem transformation element.....	172
Figure 6-5: DW/BI education problem rich picture	173
Figure 6-6: Purposeful activity building guidelines (Adopted from Checkland and Poulter (2006:220))	175
Figure 6-7: Illustration: DW/BI education purpose activity model.....	180
Figure 6-8: Applying SSM to knowledge management in the DW/BI education problem ...	181
Figure 6-9: The 'eternal triangle' of boundary judgments, facts, and values (Adopted from (Ulrich, 2005b:1))	185
Figure 6-10: Table of boundary categories (Adopted from Ulrich (2005b:1))	186
Figure 7-1: Participants' IDs from Atlas.ti	204
Figure 7-2: Streamlined codes-to-theory model for qualitative inquiry (Saldana, 2008:12)	205
Figure 7-3: Code families	212
Figure 8-1: DW/BI process activity model.....	280
Figure 8-2: DW/BI education rich picture according to student's perspective (P7).....	282
Figure 8-3: Conceptual activity module according to student group.....	286

Figure 8-4: DW/BI education rich picture according to DW/BI professional (P3) perspective	288
Figure 8-5: Conceptual model according to DW/BI professional.....	292
Figure 8-6: DW/BI education rich picture according to faculty member (Participant P1)	294
Figure 8-7: A conceptual activity model according to university faculty members (lecturer)	297
Figure 8-8: Accommodative activity model.....	304
Figure 9-1: Activity model for improving a DW/BI module.....	319

LIST OF TABLES

Table 2-1: Guidelines principles: application in the study	16
Table 2-2: A summary of interpretive research characteristics - Adapted from Joubish <i>et al.</i> (2011:2084)	22
Table 2-3: Interpretive research principles	32
Table 3-1: Comparison of OLTP vs. OLAP (<i>Chen et al., 2002</i>)	68
Table 3-2: Terms and definitions.....	82
Table 3-3: SCD summary (Ross, 2013b).....	90
Table 4-1: IT body of knowledge with core topics underlined (Lunt et al., 2008:27).....	106
Table 4-2 Summary: 38 ETL subsystems and required skills as per Kimball (2004:1).....	113
Table 4-3: ETL skills set and knowledge	117
Table 4-4: Literature review skill sets/knowledge	123
Table 4-5: General IT skills / knowledge.....	128
Table 4-6: Knowledge area ranking.....	132
Table 4-7: Advance general skills.....	133
Table 4-8: Development of prerequisite and advanced general IT skills.....	134
Table 4-9: DW/BI specific skills/knowledge	136
Table 5-1: Top 12 universities in South Africa	142
Table 5-2: Top 12 universities in South Africa	143
Table 5-3: DW/BI module content	144
Table 5-4: Module content sources reference table.....	147
Table 5-5: DW/BI module content	149
Table 5-6: General skills from module contents.....	153
Table 5-7: DW/BI specific skills/knowledge from module contents	155
Table 5-8: Advanced general IT skills mapped to universities' skills.....	157
Table 5-9: DW/BI specific skills mapped to universities' skills	158
Table 5-10: Development of prerequisite and advanced general IT skills	159
Table 5-11: Second guideline DW/BI specific skills	160
Table 5-12: Advanced general skills.....	161
Table 6-1: CATWOE analysis of DW/BI education problem situation	177
Table 6-2: Evaluation criteria for DW/BI education problem	178
Table 6-3: Four perspectives for examining selectivity (Adopted Ulrich (2005b:1)).....	185
Table 6-4: Table of boundary categories (Adopted from Ulrich (1983:33))	187
Table 6-5: Boundary judgements: DW/BI education problem (MSc study perspective)	188
Table 7-1: Biographical information of participants.....	194
Table 7-2: Section 2: Stakeholders	195

Table 7-3: Module content	197
Table 7-4: Participants	198
Table 7-5: Participants' summary	200
Table 7-6: 8 step process to perform content analysis (Zhang & Wildemuth, 2009a:308) .	203
Table 7-7: Code scheme unit test on sample of participants (P10 and P11).....	206
Table 7-8: Alphabetical extract of undirected codes and counts for all participants' responses	207
Table 7-9: Undirected code scheme testing on participant entire response	209
Table 7-10: Count of codes per document for code family: <i>DW/BI Roles</i>	213
Table 7-11: Count of codes per document for code family: <i>DW/BI education preference</i> ..	215
Table 7-12: Count of codes per document for code family: <i>Why participants prefer DW/BI education</i>	216
Table 7-13: Count of codes per document for code family: <i>Current DW/BI beneficiaries</i> ...	217
Table 7-14: Count of codes per document for code family: <i>Ideal DW/BI education beneficiaries</i>	219
Table 7-15: Count of codes per document for code family: <i>Current DW/BI education purpose</i>	220
Table 7-16: Count of codes per document for code family: <i>Ideal DW/BI education purpose</i>	222
Table 7-17: Count of codes per document for code family: <i>Current decision makers</i>	224
Table 7-18: Count of codes per document for code family: <i>Ideal decision makers</i>	224
Table 7-19: Count of codes per document for code family: <i>Current conditions outside control</i>	226
Table 7-20: Count of codes per document for code family: <i>Current decision maker's powers</i>	227
Table 7-21: Count of codes per document for code family: <i>Ideal decision makers' powers</i>	228
Table 7-22: Count of codes per document for code family: <i>Current affected stakeholders</i>	230
Table 7-23: Count of codes per document for code family: <i>Ideal affected stakeholders</i>	231
Table 7-24: Count of codes per document for code family: <i>Current affected stakeholders' control</i>	232
Table 7-25: Count of codes per document for code family: <i>Ideal affected stakeholders' control</i>	233
Table 7-26: Count of codes per document for code family: <i>Current stakeholders involved</i>	235
Table 7-27: Count of codes per document for code family: <i>Ideal stakeholders involved</i>	236
Table 7-28: Count of codes per document for code family: <i>Current expertise involved</i>	238
Table 7-29: Count of codes per document for code family: <i>Ideal expertise involved</i>	238
Table 7-30: Count of codes per document for code family: <i>Current reflected views</i>	240

Table 7-31: Count of codes per document for code family: <i>Ideal reflected views</i>	241
Table 7-32: Count of codes per document for code family: Current success measures.....	242
Table 7-33: Count of codes per document for code family: <i>Ideal</i> success measures.....	243
Table 7-34: Count of codes per document for code family: <i>Current success assurer (guarantor)</i>	244
Table 7-35: Count of codes per document for code family: <i>Ideal success assurer (guarantor)</i>	246
Table 7-36: Count of codes per document for code family: <i>Prerequisites</i>	248
Table 7-37: Count of codes per document for code family: <i>Elective subject</i>	249
Table 7-38: Count of codes per document for code family: <i>Assessment component</i>	251
Table 7-39: Count of codes per document for code family: <i>Student preparation</i>	253
Table 7-40: Count of codes per document for code family: <i>Perception of what universities are teaching</i>	255
Table 7-41: Count of codes per document for code family: <i>DW concepts experience</i>	257
Table 7-42: Count of codes per document for code family: <i>What to teach DW/BI students</i>	259
Table 7-43: Count of codes per document for code family: <i>Other concerns</i>	261
Table 7-44: Count of codes per document for code family: <i>DW/BI process</i>	264
Table 7-45: A list of skills that support the guidelines of Chapter 5.....	265
Table 7-46: Items not previously covered in Chapter 5.....	268
Table 8-1: Current (“is”) and ideal (“ought to”) DW/BI education state.....	272
Table 8-2: Steps or process to redesign DW/BI module	279
Table 8-3: SSM activities	281
Table 8-4: SSM rich picture elements according to student.....	282
Table 8-5: DW/BI CATWOE according to students’ group.....	284
Table 8-6: SSM rich picture elements according to DW/BI professional (P3)	289
Table 8-7: DW/BI CATWOE according DW/BI professionals group.....	290
Table 8-8: SSM rich picture elements according to faculty member (P1 - lecturer).....	294
Table 8-9: DW/BI module CATWOE according faculty members	295
Table 8-10: SSM rich picture elements according to student, DW/BI professional and faculty member groups.....	299
Table 8-11: CATWOE of DW/BI education according students, DW/BI professionals and faculty members.....	301
Table 9-1 Advanced general IT skills prerequisite	311
Table 9-2 DW/BI specific skills	313
Table 9-3 Advanced general IT skills	315
Table 9-4 CSH analysis of problem situation.....	317

Chapter 1 Introduction

1.1 Introduction

The purpose of this chapter is to introduce the research study, present the motivation for the research study, and discuss how it will be conducted in terms of the research methodology. The objective of this study is to develop guidelines for data warehousing/business intelligence modules in information technology programmes at universities; the guidelines are based on an understanding of the needs of various stakeholders. Two methodologies, i.e. critical systems heuristics and soft systems methodology, are used to understand the various stakeholders' perspectives.

In the data warehouse/business intelligence industry data are collected, transformed according to business rules, loaded, and presented from source systems and files for decision making purposes (Ranjan, 2009a:60). Data from different source systems, which are integrated and presented to assist strategic decision making in organisations, are stored in data warehouses (Kimball *et al.*, 2008c:5). A form of intelligence is applied to get information from data. Intelligence has many definitions that have strong similarities between them, i.e. patterns, trends, abstract thought, understanding, self-awareness, reasoning, learning, planning, and problem solving (Legg & Hutter, 2007:391, Ranjan, 2009a:60). Any intelligence that can be gathered from historical data through processes, using tools and human talent, forms business intelligence (Kamoor & Sherif, 2012:1626)

The data warehouse/business intelligence industry has a need for better qualified graduates that possess skills and knowledge that can be applied to get intelligence from data. This need is directed towards universities to produce graduates that possess such knowledge and skills. Students should be prepared for different roles that are within the realm of business intelligence. The goal is to get students to make sense of the vast amounts of data collected about all dimensions of a business, and to help decision makers to take sound business decisions. Human talent is one of the components that are crucial to implement business intelligence. Talent supply from universities into the industry is of utmost importance to the managers.

Business intelligence content can be offered as: embedded within another course; a full course; a major subject; a degree programme; or a non-degree programme. While business intelligence has gained some ground in academia, its integration “into university business courses has not kept pace with market needs” (Olsen & Bryant, 2012:1).

Universities are expected to meet the needs of stakeholders, but there may be challenges that affect business intelligence teaching. Industry, students, and faculty members are intuitively identified as stakeholders that are involved in, or affected by, data warehousing and business intelligence programmes. These three groups of stakeholders may have different views about data warehousing and business intelligence programmes. Each stakeholder group may want their views to be reflected in the programmes. This becomes a complex problem area; in trying to satisfy one stakeholder, all other stakeholders are affected. The systems approach is used in this problem environment to incorporate the needs of different stakeholders in guidelines for data warehouse/business intelligence modules in information technology programmes.

A systems approach is needed to guide universities to identify the content that they should teach data warehousing and business intelligence students, and the processes to follow to improve these programmes. Systems thinking applies the systems approach, which aims to reduce the complexity of our rational thinking in trying to understand and unfolding realities (Ulrich & Reynolds, 2010a:244). There is a need for a set of guidelines that can help universities to better prepare students for industry.

This remainder of the chapter is organised into five sections. Section 1.2 presents the concepts central to the study. Section 1.3 provides a detailed problem statement for the research study, background and motivation for this study. Section 1.4 presents the objectives of the study. Section 1.5 contains an overall discussion of the research plan with the chapter classification presented in Section 1.6. The chapter concludes with a summary in Section 1.7.

1.2 Concepts central to the study

This section presents concepts that are central to the study. Business intelligence (BI) and systems thinking are discussed next.

1.2.1 Business intelligence

Business intelligence (BI) is an umbrella term to describe concepts and methods to improve business decisions that are based on facts and derived from data (Al-Qirem, 2012b:93, Mulcahy, 2007). This umbrella term covers data, technology, analytics, and human knowledge. BI is concerned with reporting that derives knowledge from data; it is done through analytical tools and in order to support decision making. This function differs from operational information technology (IT) that concentrates on daily transactions. BI is concerned with the overall coordinated function of technology, people, and operating procedures in the execution of on-going tasks.

Different stakeholders such as academic professionals, industry, and business experts attach different definitions and meanings to the term BI. For example, just to name a few, consider the following definitions:

Ranjan (2008:463) defines BI as the conscious, methodical transformation of data from any and all data sources into new forms to provide information that is business-driven and results-oriented.

Kamoor and Sherif (2012:1626) define BI a discipline that has components such as skills, technologies, applications, and business processes for better decision making, especially strategic ones.

Rud (2009:3) defines BI as “a term that encompasses all the capabilities required to turn data into intelligence.”

All these definitions point to one or more components that are covered under the BI umbrella term. The resulting knowledge and wisdom that is gained from the use of BI can be used to support business decisions in order to increase profitability, sales,

revenue, and customer base. This knowledge can influence decision making by understanding employee behaviour, customer behaviour, and to know what competitors are doing in the market (Ranjan, 2009a:60).

The core of BI is data, which are collected and stored in flat files, operational and legacy systems, as well as data warehouses (DWs), and transformed into business information. *Because of the close relationship between BI and DW in university programmes, the term warehouse/business intelligence (DW/BI) is used in this dissertation to refer to a BI system based on a data warehouse.* Other DW/BI related concepts, such as flat files, operational and legacy systems, and data warehousing architecture are presented in detail in Chapter 3.

1.2.2 Systems thinking

Formally methods such as linear thinking and analysis, which is the process of deconstructing a topic or problem to study the smaller pieces, were the methods mostly used to solve problems (Bartlett, 2001:1). Complex realities of life, where nothing stays the same, necessitated better approaches for addressing them. One such approach is systems thinking. Systems thinking provides a means to understand the problem situation holistically, because it is a framework for seeing interrelationships rather than things (Senge, 1990:31).

Holistic thinking is a feature of systems thinking; it allows seeing the bigger picture (whole system is examined) instead of examining each part (Bartlett, 2001:1). Holistic improvement of the system is the aim of systems thinking, rather than focusing on snapshots of isolated parts (Bartlett, 2001:1, Senge, 1990:31). Systems thinking allows one to understand the dynamics relationships (one affects the other) and complexities that influence the system (Ackoff, 1971:661). The relationships define the links between systems' emerging properties. Checkland (1997a:668) describes emerging properties as features of the entire system, which is not present in any of the parts, and resulting from the interaction between the individual parts of the system.

Different authors attach various definitions to the term “system”, of which hold similar meanings. The following are some of the definitions of highly regarded pioneers in systems thinking:

Bertalanfy (1976:298) defines a system as an entity maintaining its existence through the reciprocal interaction of its parts. Checkland (1997a:667) defines the system as a whole seen as having properties making it more than the sum of its parts. Ackoff (1971:662) defines a system as a set of interrelated elements.

Systems and systems thinking are used in this study to identify different stakeholders and to articulate the different perspectives of these stakeholders. The fundamentals of systems thinking is discussed in more detail in Chapter 6 of this study.

The background to the problem area is presented in the next section.

1.3 Background of the problem area

Information technology (IT) degree programmes are designed for students to learn general IT concepts that revolve around the usage of data, and the applications of hosts of tools that make data stratified and timely (Harper, 2013, Reeves, 2015a). Universities face the challenge of how to make students data warehouse/business intelligence (DW/BI) ready. Mrdalj (2007:134) argues that one of the basic challenges in teaching a DW/BI course is its high overlap with statistics, databases, and various business disciplines. Separate domains of knowledge become interconnected, and integration become essential. A distinction between DW/BI specific, and other or general information technology (IT), skill sets must be made. Universities should identify which elective subjects can be taught in parallel with the DW/BI modules.

It is still to be established if students have knowledge of what goes on beyond the higher education institution environment in terms of the required skills, knowledge, and attitudes expected of them in the workplace (Komla & Offei-Ansah, 2011b:54). Students should know what expertise is (or ought to be) possessed by DW/BI

professionals, or what counts (should count) as relevant knowledge in this field. The experience of current and former (past) students that enrolled for DW/BI modules needs to be incorporated in the development of the module. DW/BI requires different skills for effective execution. Companies need employees who are able to apply technologies, applications, and processes related to DW/BI.

The development and content of DW/BI modules at universities should be aligned to its purpose. There is a need to establish what is, or ought to be, the purpose of DW/BI modules, such that it serves the interests of those who must benefit from it. Whose interests are served by DW/BI, i.e. the involved and affected stakeholders must be identified and their views should be shared or incorporated in DW/BI modules' development. Different stakeholders may have different viewpoints of what expertise is, or ought to be, possessed by DW/BI professionals; or what counts, or should count, as relevant knowledge in this field. For this reason, it is very difficult for universities, being service providers, to meet and exceed expectations when expectations are not known. A systems approach, to make unknown expectations of stakeholders known, is needed.

Universities face a challenge to meet stakeholders' expectations whilst operating within certain academic boundaries, i.e. the National Qualifications Framework (NQF) and other sub committees such as faculty boards that, amongst others, govern the scientific content and required hours of academic programmes. (Trauth *et al.*, 2003:294) highlight that accreditation standards also limit curriculums; furthermore, there is a delay between the design and implementation of curriculum changes. There may be other challenges that universities face and they need to be unearthed.

The problem addressed by this study is that there is a lack of guidance to universities to assist them to develop of DW/BI modules that truly prepare students for the needs of industry, while satisfying the needs of an academic programme. There is a greater need for universities to know what content to teach, and what processes steps to follow to improve DW/BI modules.

1.4 Study objectives

The main objective of this study is to develop guidelines for DW/BI modules in IT programmes at universities; the guidelines are based on an understanding of the needs of various stakeholders. Two methodologies, i.e. critical systems heuristics (CSH) and soft systems methodology (SSM), are used to understand the various stakeholders' perspectives.

1.4.1 Theoretical objectives

The following theoretical objectives are set for this study:

1. Demonstrate an understanding of what constitutes a guideline as well as an understanding of general research methodology in order to justify the selection of the interpretive research paradigm.
2. Demonstrate an understanding of DW/BI components within the IT discipline.
3. Demonstrate an understanding of SSM and CSH in the context of systems thinking.

1.4.2 Empirical objectives:

The following empirical objectives are presented for this study:

1. Investigate what DW/BI modules and programmes universities currently teach, by collecting and analysing document data from universities' DW/BI programmes.
2. Investigate the perspectives of stakeholders, using CSH and SSM, as well as understand the needs of the stakeholders in terms of DW/BI module content and the process to design a DW/BI module. Although the aim of this objective is to identify stakeholders, it is intuitively assumed that the stakeholders will at least include industry, universities, and former students.

3. Develop guidelines representative of the needs of all identified stakeholders, grounded in the literature review, universities' programmes, and from interview data analysis.

1.5 Overall research methodology

The chosen research methodology for this study is given and substantiated in Chapter 2. The study is performed in the interpretive research paradigm. Guidelines are compiled firstly from DW/BI literature. The guidelines are enhanced after DW/BI modules' documents from universities in South Africa are analysed. The findings from content analysis of data collected from interpretive interviews with stakeholders are then used to refine the guidelines. The interview questions are guided by the literature reviews on DW/BI, the document analysis of module and programme guides of universities, and theory of CSH and SSM. Finally, guidelines for the content of a DW/BI module in an IT programme, as well as guidelines for the process of developing such a module, are presented.

1.6 Chapter layout

This study is organised into chapters and are presented as follows:

1.6.1 Chapter 2: Research methodology

A discussion on research methodologies is presented. The researcher chooses, and motivates the choice of the interpretive research methodology to follow in this study. The application of the methodology and a research plan for this study is also presented.

1.6.2 Chapter 3: Business intelligence and data warehousing

A discussion on DW/BI literature is presented, with a purpose of creating a shared understanding of the underlying concepts.

1.6.3 Chapter 4: Guidelines from DW/BI literature review

Specific DW/BI skills required by DW/BI professionals are identified in this chapter from the literature review presented in previous chapter. A division made between general IT and specific DW/BI skills.

1.6.4 Chapter 5: Guidelines from institutions of higher education

The secondary empirical element of the study from document data is presented in this chapter. Document data gathered from other universities' DW/BI programmes are used to understand what they currently teach. This chapter provides a combined list skills from DW/BI literature and university programme/module documentation.

1.6.5 Chapter 6: Systems thinking methodologies

The systems thinking methodologies, soft systems methodology (SSM) and critical systems heuristics (CSH), are discussed in detail in this chapter

1.6.6 Chapter 7: Stakeholders' perspectives on DW/BI education

This chapter reports on the primary empirical element of the study. The interpretive interview process of data collection, data analysis, and findings are reported on in Chapter 7. The interview content is guided by concepts of systems thinking methodologies, as discussed in Chapter 6. Data analysis is performed using content analysis. The findings of this chapter focus on the content of a DW/BI module.

1.6.7 Chapter 8: Stakeholders' perspectives on DW/BI education (continuation)

Findings on the boundary questions of CSH, identified in the data analysis of Chapter 7, are presented in this chapter. The findings of this chapter focus on the process to follow when designing a DW/BI module. SSM components are used to present this process.

1.6.8 Chapter 9: Summary

This chapter marks the conclusion of the current study. The research process is summarised and the consolidated guidelines are presented. Limitations of the study and further work to be done are also outlined.

1.7 Summary

The problem area introduced in this chapter is the difficulty traditional universities experience to provide the DW/BI industry with graduates that are able to contribute to the organisation without substantial further training.

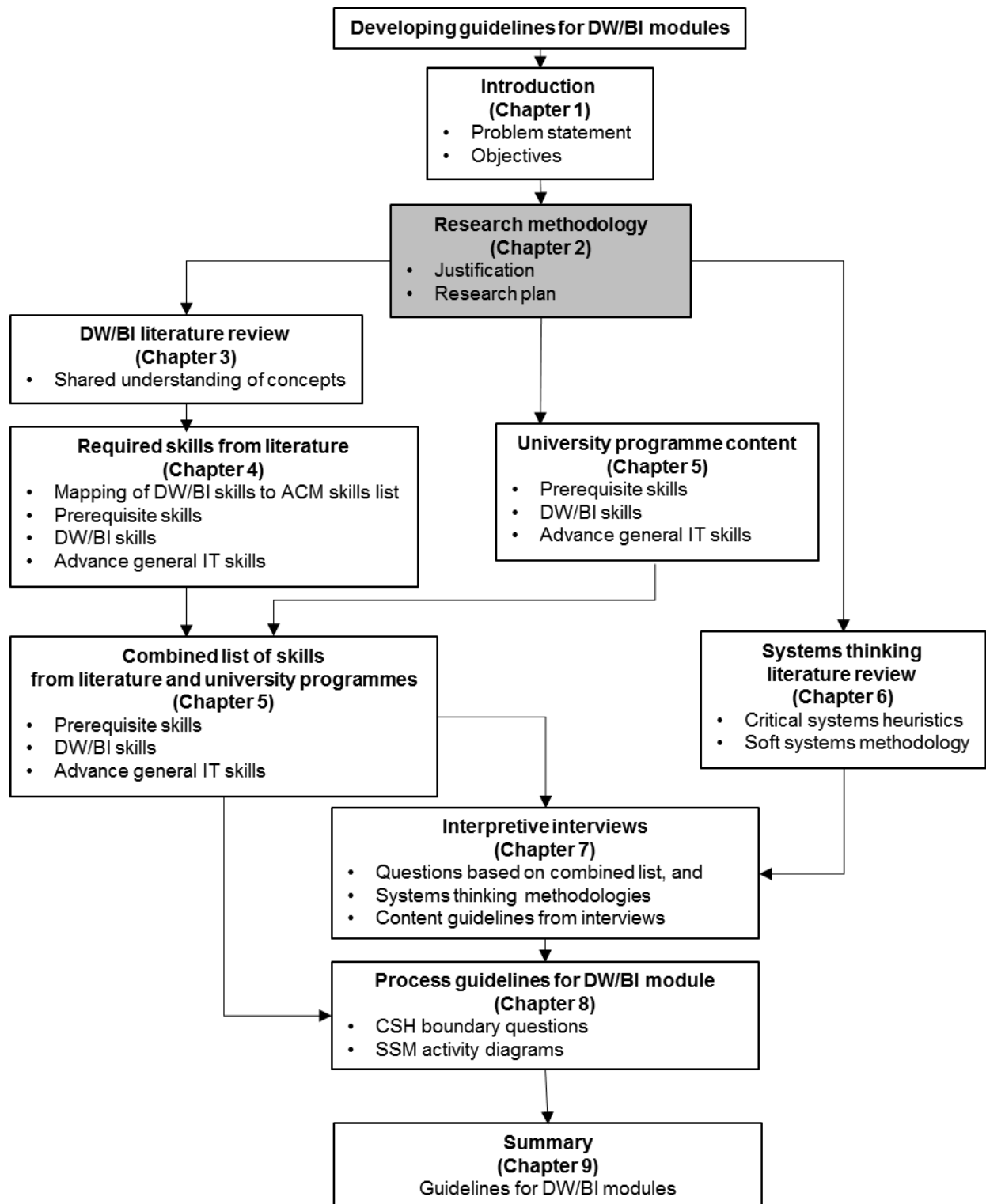
Business intelligence refers to strategic decision making in organisations based on the processing of data/information already available in the organisation. Data from sources systems in large organisations are integrated in DWs. DWs are key technological systems for BI. Different stakeholders are identified; they have different needs and limitations. Universities have to operate within specific limitations, such as teaching hours and qualification programme requirements. The DW/BI industry functions in an ever changing information and technological environment. Factors such as these motivate the use of a holistic approach to address this problem. Systems thinking provide tools for understanding and action in a complex environment where holistic thinking is required. CSH and SSM are selected as suitable systems thinking methodologies to use in this study.

The overall objective of this study is thus to develop guidelines for DW/BI modules in IT programmes at universities; it is based on an understanding of the needs of various stakeholders through CSH and SSM.

This objective is achieved by means of interpretive research methods. Initial guidelines are formulated based on DW/BI literature and programme/module documents of universities. These guidelines, along with systems thinking methodologies (CSH and SSM) are used to develop an interpretive interview. The participants of the study are representatives of identified stakeholder groups.

Guidelines are presented in terms of both the content and the development process of a DW/BI module.

The next chapter provides detail on research methodologies and a justification for the use of interpretive methods in this study.



Chapter 2 Research methodology

2.1 Introduction

The industry needs graduates that are better prepared to contribute to an organisation. Higher education is critical in preparing students for a knowledge-based work society. Higher education institutions provide formal education to students, but real learning begins when they start working; Komla and Offei-Ansah (2011a:53) argue that not enough studies have been conducted “to investigate the constraints and challenges of industrial and institutional linkages”. The objective of this study is to develop guidelines for data warehousing/business intelligence modules in information technology programmes at universities; the guidelines are based on an understanding of the needs of various stakeholders. Two methodologies, i.e. critical systems heuristics and soft systems methodology, are used to understand the various stakeholders’ perspectives.

Conducting and evaluating research require knowledge of philosophical assumptions about what constitutes valid research, and the relevant methods to be applied. Relevant methodologies that are used to develop knowledge in a research study are informed by these philosophical assumptions (Rowlands, 2005a:81). The purpose of this chapter is to present the methodology used to achieve the main objective of this study.

Following this introduction, this chapter is organised as follows: Section 2.2 explains the concept of guidelines and characteristics in general since the aim of the study includes the development of guidelines. Section 2.4 explains research paradigms (positivism, interpretive research, critical social theory, and design science) in order to motivate the selection of the interpretive paradigm for this study. Section 2.5 provides a more detailed discussion on the interpretive research methodology. Interpretive data collection and analysis methods are presented in Section 2.8. Section 2.7 provides a concise research plan for this study and therefore an explanation of the adjacent research diagram. A chapter summary is given in Section 2.8.

2.2 Guidelines

The relationship between business intelligence (BI) and information communication technology (ICT) pedagogy in institutions of higher education is not straightforward (Yeasmin & Rahman, 2012b:154). A set of guidelines may ease the integration of business intelligence/data warehouse (DW/BI) in academic programmes. The aim of this section is to describe the nature of guidelines. Clearing the misconceptions, misunderstandings, and misinterpretations around guidelines assist the development of useful guidelines in support of the objectives of the study.

Guidelines are recommendations aimed at identifying interventions that will ensure the best possible outcome (NHMRC, 2011). They are grounded by the history of the problem situation and are traditionally developed on consensus by a panel of experts in the field (Chena *et al.*, 2011:288, Jackson *et al.*, 2009:303).

Guidelines differ from policies and laws. Policies and laws are enforced and become mandatory once the government turns them into legislature; other organisations and institutions incorporate them into their code of conduct. A guideline can be seen as recommendation, advice, proposal, direction, suggestion, counsel, or specification. A guideline is not a policy or rule that must be enforced; rather, it can be viewed as a recommended practice (NHMRC, 2011). It allows room for application of judgment to determine what is appropriate for the specific situation (West, 2011).

2.3 Key principles for developing guidelines

Boon and Tan (2006:197) and The National Health and Medical Research Council of Australia (NHMRC, 1999:13) recommend nine key principles for developing guidelines. These principles are more general and can be adopted where applicable; they can be made more specific to the problem area. In this study, these principles are used as a reference point to guide the process of formulating the guidelines for DW/BI modules.

Principle one suggests that guidelines must focus on the outcomes. Guidelines are used with the expectation of improved outcomes, to lessen variability, and must be

valid and usable (Zimlichman & Meilek-Weiss, 2004:626). The validity and usability of guidelines are properties that must be assessed. The expected outcomes, after having used guidelines, must be explicit and clear to all. Figure 2-1 below illustrates the three properties of a guideline, i.e. that it must be valid, usable, and it must be assessed.

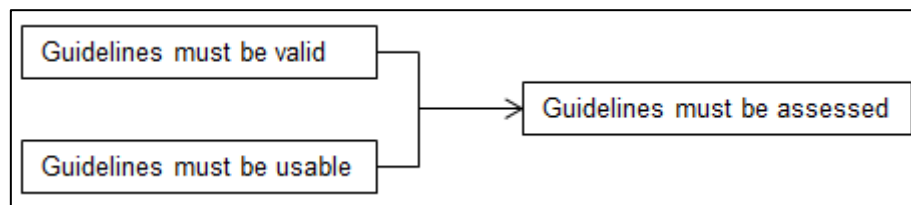


Figure 2-1: Properties of a guideline

Principle two dictates that guidelines must be based on evidence. Guidelines must be informed by evidence (data and history) that can support their propositions (NHMRC, 1999). Evidence in the form of data that are synthesised can be the base of good guidelines. *Principle three* involves the method used to synthesise the available evidence; it should be applicable and relevant. This process should be a systematic approach that encourages decision makers to carefully consider the entire process (NHMRC, 1999:17, Johnson, 2011). All stakeholders should be included (*principle four*). In order to be usable, the proposed guidelines must not be above, or outside the scope of, the stakeholders' expectations and needs (NHMRC, 1999:17).

Guidelines should allow for flexibility and adaptability in their implementation (*principle five*). Depending on the intended use, they should be flexible for conditions such as geography, environment, resources, and costs constraints; they should also accommodate different consumers' values and preferences (NHMRC, 1999:17). The guideline formulation process must consider all constraints and challenges (*principle six*). The recommended guidelines should be viable, sound, and probable within the limits of the involved and affected stakeholders (NHMRC, 1999:17).

Guidelines must be clear and understandable, and usable (*principle seven*). So, a good guideline is clear, specific, and understandable (NHMRC, 1999:17). This

principle suggests that, if a guideline is not clear, it is the fault of the panel/team that formulated the guideline. However, if a guideline is clear and understandable, but not adhered to, then it is the fault of those who did not adhere to them. Sun Tzu said: “If the words of the commander are not clear and distinct, if orders are not thoroughly understood, then the general is to blame. But if the orders are clear but the soldiers disobey, then it is the fault of their officers” (Giles, 1910:5).

Guidelines are formulated with a purpose and goal in mind. Their implementation outcomes should be evaluated to gauge the impact of their use (*principle eight*). The implementation outcomes must be gauged against their initial objectives to see if they serve their intended purposes (NHMRC, 1999:18). The last principle (*principle nine*) advocates that guidelines should be revised regularly. Changing conditions and principle two (evidence based) suggest that a guideline should be revised using the latest information; new research outcomes should be taken into account (NHMRC, 1999:19). Guidelines are time and conditions oriented, because these variables change all the time.

2.3.1 Position of the study

Guidelines (in this instance for DW/BI programmes) are developed with the intention to provide support to institutions of higher education’s policies and procedures, not to contravene them, otherwise they will be invalid (Ballarat, 2010a). Table 2-1 below presents the application of these steps in this study.

Table 2-1: Guidelines principles: application in the study

Principles	Application to this study
1. Guidelines must focus on the outcomes.	The outcome results of guidelines for DW/BI modules are better programmes that produce better students. This step is applied throughout this study (this is the focus of the study).
2. Guidelines be based on evidence	DW/BI guidelines are systematically formulated from data gathered from other universities DW/BI programmes and findings from the empirical data. This principle is applied in Chapter 3, Chapter 4, Chapter 5, Chapter 7, Chapter 8 and Chapter 9.
3. Method used to	Multiple cross referencing and comparison is applied until the

synthesise the available evidence and should be applicable and relevant.	final set of guidelines is established. This process necessitates coding scheme to help with traceability of skills and their sources (i.e. where and how they were identified). This principle is applied in Chapter 4, Chapter 5, Chapter 7, Chapter 8 and Chapter 9. The work of Zhang and Wildemuth (2009b:309): “set of systematic and transparent procedures” involving an 8 step process guiding content analysis is applied in Chapter 7. These steps involve coding data in order to make the findings traceable to the data.
4. All stakeholders should be included	Guidelines for DW/BI modules are developed with the objective of representing the views of the involved and affected stakeholders. The targeted participants are experience DW/BI professionals in the industry, former students and faculty members. This principle is applied in Chapter 7, Chapter 8 and Chapter 9.
5. Guidelines should allow flexibility and adaptability in their implementation	Guidelines for DW/BI education should be flexible enough to allow their use in different institutions of higher education that have different constrains such as access to resources, policies, etc. They are not case specific but generic to be used by other universities.
6. Guideline formulation process must consider all constraints and challenges.	Constraints and challenges that affect DW/BI modules are identified from data collected both from data gathered from other universities DW/BI programmes and findings from the empirical data. This principle is enhanced by the use of systems thinking methodologies in this study.
7. Guidelines must be clear and understandable for all to use.	Guidelines are aimed at clarifying “what to teach in the DW/BI module content?” and “what process should be used to improve DW/BI education?” Care was taken to use common terminology in the formulation of the guidelines.
8. Guidelines are formulated with a purpose and goal in mind.	The purpose of guidelines is to help universities develop DW/BI programmes that will better prepare students for industry. This principle is applied throughout this study.
9. Guidelines should be revised regularly.	Guidelines are based on current data and knowledge and must be reviewed periodically. This principle is applied as a recommendation in Chapter 9.

Research paradigms are discussed in the next section.

2.4 Research paradigms

The term “paradigm” was given philosophical meaning by Kuhn (1970:5). According to Kuhn (1970:1) paradigms are based on past achievements that will be learned by newcomers, so that they can be part of the paradigm body. Kuhn (1996:10) suggests

that a paradigm defines "universally recognized scientific achievements that, for a time, provide model problems and solutions for a community of researchers. These are the sum total of ways of conducting research with a set of beliefs and assumptions that embody and support our hypotheses, methodology, and analysis" (Kuhn, 1970:5). We cannot practice science without some set of perceived beliefs and perspectives (Kuhn, 1970:1). For example, the Figure 2-2 is an illustration of a paradigm properties.

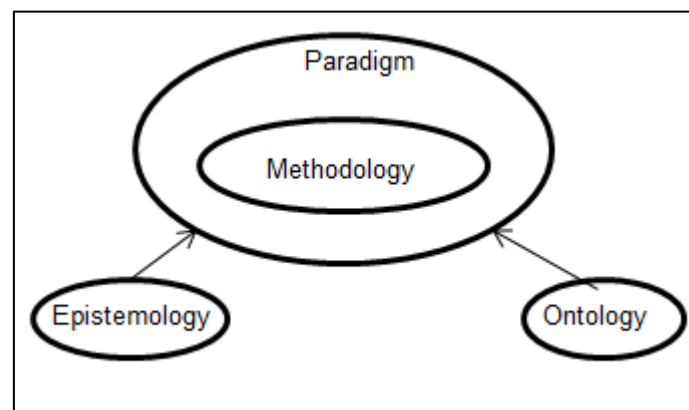


Figure 2-2: Paradigm framework

The paradigm framework is made up of philosophy, ontology, epistemology, and methodology (Joubish *et al.*, 2011:2084, Cagdas & Stubkjaer, 2011:78), as shown in Figure 2-2. Epistemology refers to our theory of knowledge (i.e. how knowledge developed) and ontology refers to our view of reality; these underpin the researcher's theoretical perspective and methodology. The ontological and epistemological aspects are concerned with a person's perception with regard to aspects of reality.

On a philosophical level "a paradigm is the underlying assumptions and intellectual structure upon which research and development in a field of inquiry is based" (Kuhn, 1970:1). A research paradigm can also be thought of as a guide because it is an approach to think about and do research (Wahyuni, 2012b:72). Based on this background, research paradigms inherently reflect our beliefs about the world we live in. Methodologies and methods used in a research study must be used under the philosophical umbrella of a paradigm that relate to the nature of knowledge (Ballarat, 2010b). The relevant methodology used to develop knowledge in research must be within the context of a paradigm (Wahyuni, 2012a:72).

Positivism, interpretive, critical social theory, and design science paradigms are presented next.

2.4.1 Positivism

There are many research topics within the scientific realm that have no connection to personal meaning. Such topics are characterised by findings derived from empirical data that can be verified and which, using the results (quantitative data), form a single body of knowledge about the world (Hammersley, 2012). Such research studies fall under the positivism paradigm; the aim is to study the relationships between variables, rather than explain behaviour subjectively.

Positivism as a research paradigm dominated the research world for the first half of the twentieth century (Hammersley, 2012). The term positivism was coined in the 18th century by one of the founders of sociology, i.e. Auguste Comte (Martiau, 1896:39). He argues that human cognition must pass through three historical phases: the theological, the metaphysical, and the scientific. Spiritual forces are used as a reference point to explain the natural and social phenomena in the theological stage. In the metaphysical stage, the original causes are used to explain the phenomena. In the scientific stage, scientists seek to discover correlations amongst phenomena rather than explaining it (Jakobsen, 2012). This means that events are explained by scientific principles. For this discussion attention is given to the scientific stage.

The ontological assumptions of positivism consider reality as independent of social construction; observation and reason explain human behaviour (Kaboub, 2008b:786). This means that the researcher should not subjectively get involved. Positivism studies the cause of behaviour, not the motivation thereof. The deep social meanings that may be associated with root causes are not a prerogative for positivism. The positivism paradigm measures the results with the precept that the world remains the same (Krauss, 2005a:758). Reliability is measured by repetition of results (Kaboub, 2008a:786). This is achieved by scientific and experimental methods that aim to measure the influence of a specific variable in a situation (Krauss, 2005b:759). The researcher seeks to discover, describe, explain facts, and

replicates the results given the same set of parameters within the unchanged environment (Yeasmin & Rahman, 2012a:155).

The epistemological assumptions of positivism are based on the notion that data, and the relationships between data, explain human behaviour (Kaboub, 2008a:786). The object of study has no direct or indirect influence from the researchers. Direct observation is used to discover knowledge of a phenomenon (Krauss, 2005b:579). Deposition of phenomena into small parts leads to fact finding. Figure 2-3 illustrates a summary of the positivism paradigm.

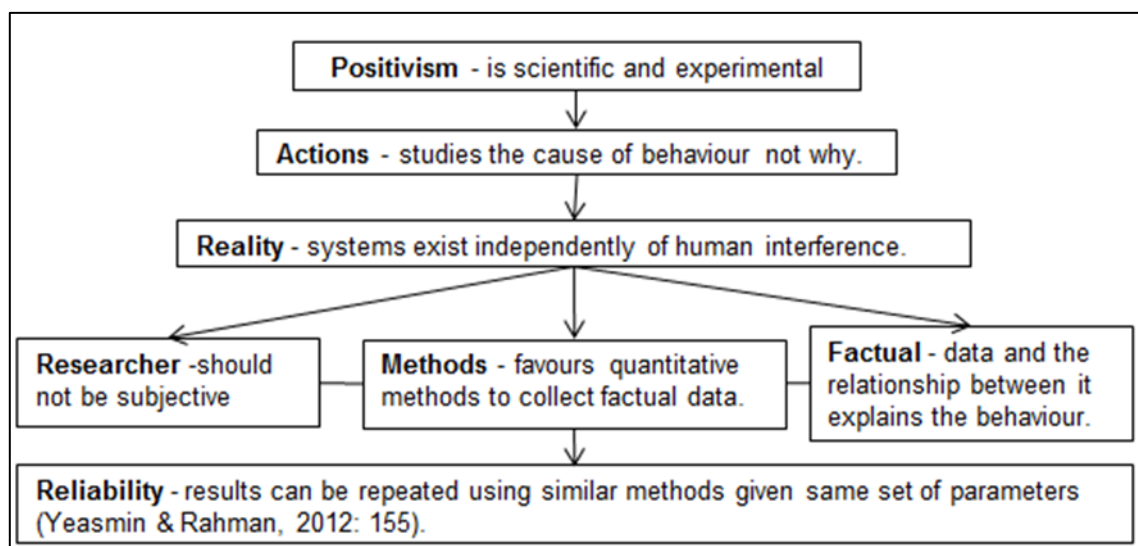


Figure 2-3: Illustration: summary of positivism

Positivistic research is objective; on the other hand, the interpretive paradigm is more subjective. Interpretive research is discussed in the next section.

2.4.2 Interpretive research

The interpretive research paradigm developed out of critique for positivism. While positivism attempts to eliminate bias and subjectivity; interpretive research reorients subjectivity and conceives data collection as a process whereby a researcher is engaged in the problem environment (Mottier, 2005:1). Kaboub (2008a:786) argues that positivism is a prescription to treat the symptoms of the underlying mechanisms that are invisible, rather than the root cause of the problem. There are elements of our world that cannot be measured objectively because they are not quantifiable;

however, the truth about them can be understood (Rowlands, 2005b:82). Interpretive research is a paradigm used to study the behaviour by looking at consciousness, action, and unpredictability (Livesey, 2006). This paradigm seeks to understand social or human problems, forming a holistic picture with words, and reporting detailed views of informants that are conducted in a natural setting (Carruthers, 1999:82). Researchers become part of the culture they are studying through observation (being actively involved and being part of the world), verification, and conclusions (Yeasmin & Rahman, 2012b:155).

The ontological assumptions of interpretive research are based on the notion that social phenomenon, or the world, is not objective; rather, it is experienced. Social phenomenon is subjective because reality changes and people's perceptions also change (Mingers, 2001:244). Human beings are part of a multidimensional world that does not remain the same and is characterised by unpredictability (Mingers, 2001:244). Uncovering the truth in a social construct requires a good approach that is methodological. The goal of interpretive research is to uncover the truth about a phenomenon in order to bring understanding of why things are the way they are (Joubish *et al.*, 2011:2083). For example, a question like '*why are students not doing well or coping in the industry after leaving the university?*' is a descriptive interpretive study question. This study requires uncovering the truth instead of analysing quantitative data. Quantitative data may give a certain percentage or count value of students that are not doing well. But the follow-up question that is, '*why are they not doing well?*' constitutes a descriptive interpretive study enriched by qualitative data.

The epistemological assumptions of interpretive research are grounded in the belief that experience brings knowledge and understanding. The social world consists of, and it is constructed through, the meanings that humans attach to it (Joubish *et al.*, 2011:2083). A researcher is not limited to the idea that there is only one view of the world that is realistic (Krauss, 2005b:758). Interpretive research methods seek to understand the world in the view of the involved (Rowlands, 2005a:82). A traceable consolidation of human perceptions of the world is a key activity in interpretive research. Data collected are not separated from context that adds meaning to it. Variables and their interrelations are not predefined in interpretive research (Krauss,

2005b:758). The aim is to produce understanding of social contexts (Walsham, 2006:321). Table 2-2 gives a summary the characteristics of the interpretive paradigm. Characteristics are presented in first (left) column with a short description on the second column. Figure 2-4 illustrates a summary of the interpretive research paradigm.

Table 2-2: A summary of interpretive research characteristics - Adapted from Joubish *et al.* (2011:2084)

Characteristics	Description
Purpose – to understand	Seeks to understand peoples' perceptions and their interpretations.
Reality – is dynamic	Peoples' perspectives are influenced by reality changes.
Viewpoint – requires insider or active participant	Reality is not always what is it but what people perceive it to be.
Values – value bound	An account of values should be taken into consideration when conducting and reporting research
Focus – is holistic	A total sum view of all perspectives is sought.
Orientation – discovery	Data collected serve as the basis for theories and hypothesis.
Data – is subjective	Peoples' perceptions serve as input data in the environment.
Instrumentation – human	The researcher and participants are the primary data collection instruments.
Conditions – naturalistic settings	Research investigations are conducted under natural settings.
Results – valid	The main focus is of the research is on the design and procedures to gain deeper meaning.

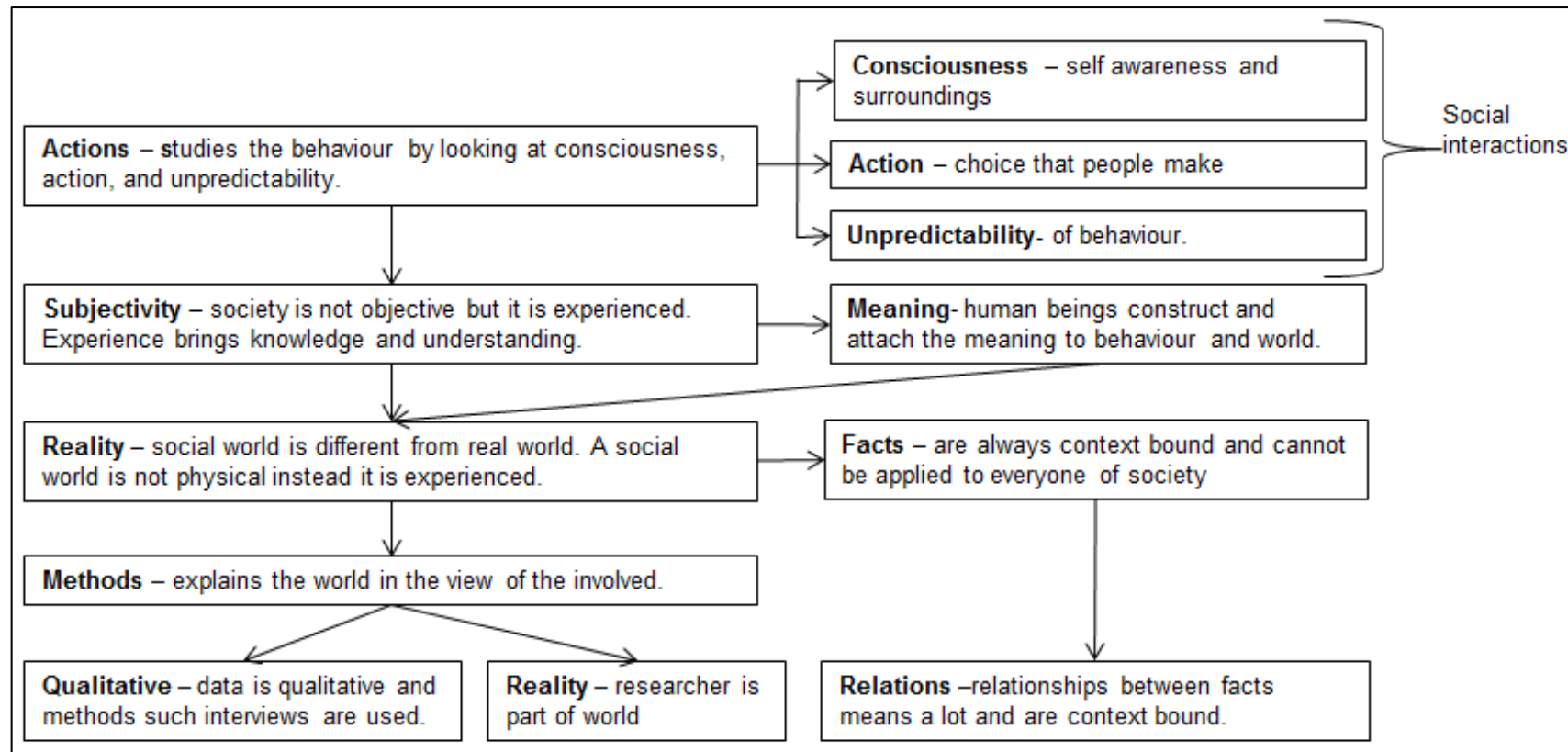


Figure 2-4: Illustration: summary of interpretive research

2.4.3 Critical social theory

According to Leonardo (2004b:11) critical social theory (CST) as a research paradigm can be traced back to philosophy, literature, and critical theory roots. In terms of philosophy, he refers to the works of Hazard Adam's critical theory since Plato "and the Frankfurt School's programmatic study of a Kantian theory of knowledge coupled with a Freudo-Marxist theory of modern society" (Leonardo, 2004b:11). Critical theory is related to social theory. Ngwenyama (1990:2) argues that the differences between social theory (ST) and CST can be seen by looking at the intention of the researcher. ST seeks to understand; CST seeks to change, improve, or find a better way around the current social conditions. Solving the problem is the ultimate goal of CST.

Critical thinking engages in a problem evaluation process with deep scepticism; it is totally focused on deciding what to believe or do (Facione, 2000:61, Facione, 1990:3). This is a self-regulatory judgment that brings analysis upon which that judgment is based (McPeck, 1990:8). Ulrich (1983:1) played a pioneering role in the critical thinking approach; he applied it to systems thinking when he developed the critical systems heuristics (CSH) method. CSH is "a framework for reflective practice based on practical philosophy and systems thinking" (Ulrich & Reynolds, 2010a:243). This critical dimension is important because social phenomenon does not yield a single right answer (Ulrich, 2005a).

The CST paradigm shares the inquiry process with the interpretive paradigm (Joubish et al., 2011:2084). Critical social theory is immersed with the critical analysis dimension that became its established policy and practice (Taylor & Medina, 2011). The epistemological assumptions of critical social theory are such that the critical researcher enables the participants to articulate their perceptions' of the world and identify the negative effects of being in it (Chermack & Lynham, 2004:47). Knowledge is gained from expressed views during the process of emancipation.

The ontological assumptions of critical social theory are based on the notion that a critical researcher assumes that social reality is historically constituted, and that it is produced and reproduced by people with inherently conflicting views (Chermack &

Lynham, 2004:47). The researcher investigates and exposes the power of relationships in the study environment. The role of researcher is that of an advocate of change who promotes a more balanced, fair, and sustainable society for all (Harvey, 1990:1). Emancipation, which is the goal of CST, necessitates a root cause analysis of the observed phenomenon (Harvey, 1990:1). Factors contributing to the current situation must be identified through a knowledge development process. Incorrect diagnosis of these factors may lead to failure of the desired change or emancipation. The understanding of the phenomenon and its social structures lead to the development of knowledge (Harvey, 1990:1). Knowledge of social structures is not static; rather, it is dynamic because the social world does not remain the same. Knowledge is a key aspect to CST as it is also intertwined with critique. Interpreting and developing knowledge is coupled with developing an effective alternative solution to the problem environment. The critical researcher has to take direct action to make a difference and evaluate the impact of recommended transformative actions.

The CST paradigm reflects (keeps one honest) on social practices and theories that attempts to understand domination and inequality with the goal of emancipating or creating equality (Ngwenyama, 1990:2, Harvey, 1990:5, Leonardo, 2004a:11). The result is a transformation process that is harnessed by a researcher creating new knowledge. A change that empowers (without creating imbalances elsewhere) the previously disempowered is critically assessed and this new knowledge contributes to the knowledge body of CST.

Critical social theory uses action research (AR) as a method to achieve emancipation in a problem environment (Klein & Myers, 2011:23). Action research is a cyclical method that has four iterative processes: initial diagnosis, planning and implementation of the intervention, analysis of the success of intervention, and advice on improvement (Zuber-Skerritt & Fletcher, 2007:149). Figure 2-5 below illustrates the summary of CST. A change (action) that is tightly intertwined with reflection and understanding (research) is pursued in AR (Altrichter *et al.*, 2002:131, Zuber-Skerritt & Fletcher, 2007:416).

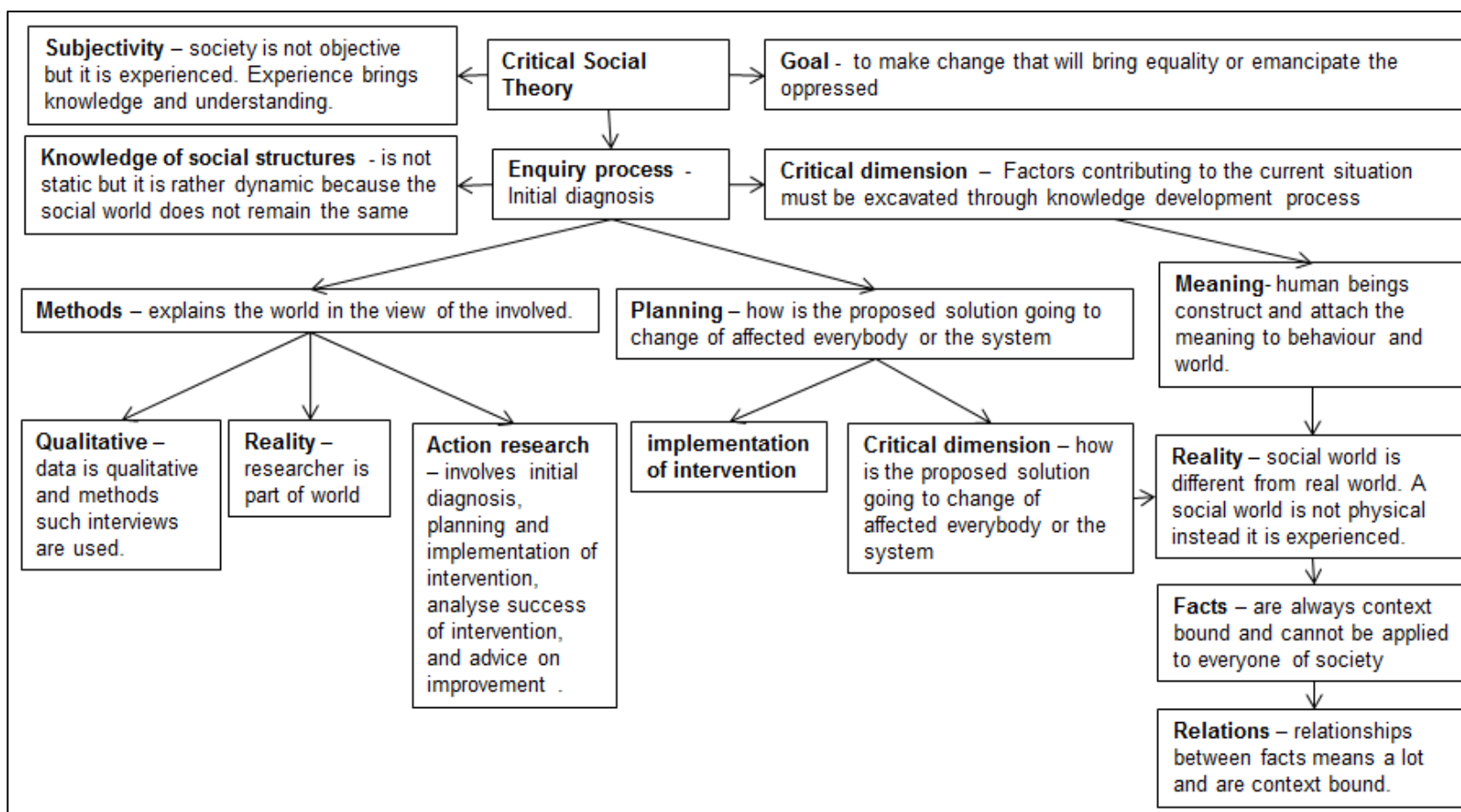


Figure 2-5: Illustration: summary of critical social theory

2.4.4 Design science

In recent years fields like ICT and bio-engineering had an exponential increase in knowledge as a result of scientific experiments that produced different kinds of artefacts (Gregor & Hevner, 2013:337). The artefacts changed how humans live and work. For example, bio-metric devices used for access, touch screen phones, and more, affected how security and communication is conducted today. These artefacts are examples that fall under the body of knowledge that is known as design science. Design science is a paradigm that is used in the information systems (IS) research field to design artefacts (such as computer software) (Gregor & Hevner, 2013:337, Keith *et al.*, 2013:229). It is about a change intended to alter how things work by extending the boundaries of human and organisational ability (Cagdas & Stubkjaer, 2011:78, Hevner *et al.*, 2004:74). This change is achieved through scientific methods of investigation (Cross, 2001:47).

Design science has an inclination of engineering in the sense that it is concerned with how things ought to be (Simon, 2006:4). This inclination is goal orientated. However, a different view warns against suggesting that only engineering is change orientated. “The intellectual activity that produces material artefacts is no different fundamentally from the one that prescribes remedies for a sick patient or the one that devises a new sales plan for a company or a social welfare policy for a state” (Simon, 2006:111). Figure 2-6 below illustrates an abstract view of a design science.

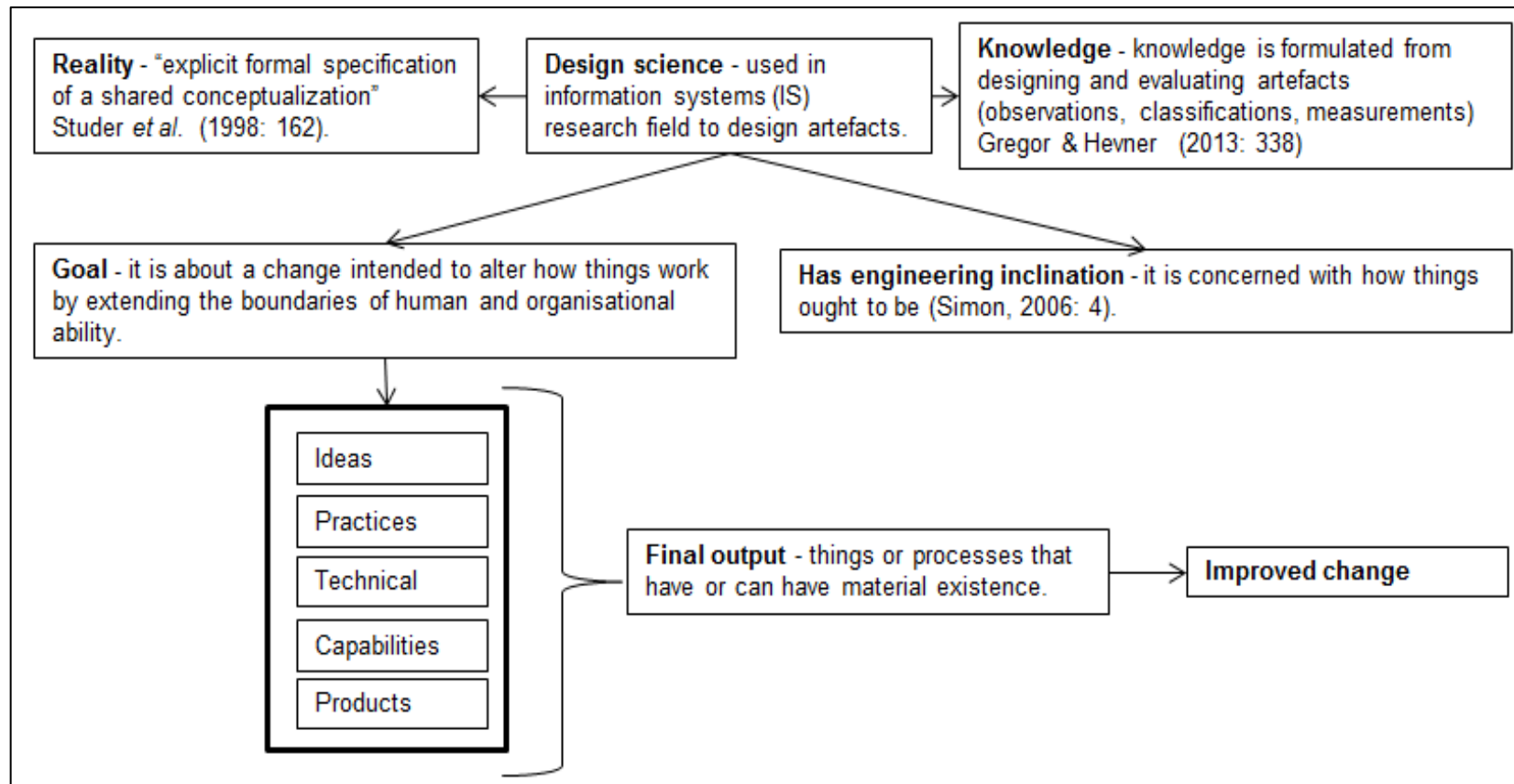


Figure 2-6: Illustration: summary of design science

Design science's epistemological assumption is that knowledge is the base for the design of practical artefacts (Gregor & Hevner, 2013:337). According to Gregor and Hevner (2013:338) design science artefact phenomena's knowledge is formulated from designing and evaluating artefacts. The body of knowledge is increased as a result of evaluation methods including: experiments, testing, observation, analytics, descriptive analysis, and action research (Baskerville & Myers, 2004:329, Hevner *et al.*, 2004:85). This knowledge or theory is about how to do something and the impact of its implementation (Keith *et al.*, 2013:229). Studer *et al.* (1998:162) define the ontology of design science as "explicit formal specification of a shared conceptualization". This definition mandates the explicit concepts and constraints of the artefact.

2.4.5 Position of the study

The focuses of the four paradigms' discussions are on the ontological and epistemological assumptions aspects. Therefore, the ontological and epistemological assumptions are used as paradigm selection criteria for this study. The current study is not objective (real world) and thus positivism is not a paradigm of choice. Design science and critical social theory are also not chosen because the objective of the study is neither to create an artefact, nor to bring change or emancipation to a problem area. The ontological stance of interpretive research is based on the notion that the world is subjective and it must be experienced in order to understand it. The current research study shares ontological assumptions with the interpretive research paradigm. The objective of this study is to understand the needs of the participants and to develop guidelines for DW/BI modules in information technology programmes at universities based on an understanding of the needs of various stakeholders, using critical systems heuristics and soft systems methodology.

The paradigm of choice is based on the ontological argument presented in the paragraph above and on its epistemological characteristics. Experience brings knowledge and understanding; this forms the basis of the epistemological assumptions of interpretive research. Participants' views are used to construct the world and forms data input. Data are gathered in order to understand perceptions, and guidelines are developed (grounded) in the qualitative data. Therefore, this

study is done according to the interpretive paradigm. The rest of the chapter provides more detail of the interpretive paradigm.

2.5 Interpretive research methodology

The goal of interpretive research is to understand human interpretations about social phenomenon (Myers & Klein, 1999:67). This “understanding” can be obtained through a social construction such as language, e.g. questionnaires, consciousness, shared meanings through interviews, and focus groups (Myers & Klein, 1999:67). Understanding and reporting participants’ interpretations, i.e. perceptions and/or attitudes, are the main goals of interpretive research. The next section discusses the principles of Myers and Klein (1999:67) that are used to guide the interpretive research methodology. This discussion is followed by discussions on data collection and analysis. The section concludes with a discussion on interpretive research evaluation.

2.5.1 Principles for interpretive research

Interpretive research is not objective; it involves an active researcher that is collecting empirical data that needs systematic analysis and processing. A form of direction is needed on how to conduct such research. Myers and Klein (1999:72) developed a set of interpretive research principles on how to conduct interpretive research in the field of information systems.

The first principle of Myers and Klein (1999:72) is the hermeneutic circle. This principle’s emphasis is on the enquiry process to understand meaning attached in communication (Ulrich, 2005a). This is achieved by implementing hermeneutics, which is defined as the theory of text interpretation. Communication entails decoding the small meanings that are attached and that form the whole message between two or more people. It is imperative for a researcher to understand the actual meaning attached to small blocks of communication, so that the actual meaning is not lost (Myers & Klein, 1999:72). The hermeneutic circle suggests that the whole body of relevant literature for a specific phenomenon is made up of multiple small texts that need to be understood from the perspective of single words, as well as from the

perspective of the entire text (Checkland, 1997b:668, Boell & Cecez-Kecmanovic, 2010:5). In turn, the individual texts are seen as parts of the whole body of relevant literature. The understanding of small texts is never isolated. It is interpreted in the context of other small texts from the literature. Human understanding is achieved by iterating between the small parts and the whole formed by them. The understanding “is built up through the understanding of individual texts” (Boell & Cecez-Kecmanovic, 2010:5).

The second principle (contextualisation) suggests that understanding of interpretive research must be put in context so that it can be relevant (Myers & Klein, 1999:72). Context gives clarification and relevance about the history of the situation and the current state as to how it came to be (Myers & Klein, 1999:72). The third principle relates to interaction between the researchers and the subjects. Active participation in research can influence the quality of data collected (Myers & Klein, 1999:72). The data collection process depends on how the researcher and the participants construct perceived worldviews. The researcher gains better understanding of the social phenomenon when he/she is actively involved in the same natural settings where the participants are in. The emphasis here is the ‘critical reflection’ aspect that “data is socially constructed between researcher and participants” (Myers & Klein, 1999:72).

The fourth principle is that of abstraction and generalisation. Interpretation of text meaning (hermeneutics) and contextualisation are the most important principles of interpretive research. Interpretive research is not characterised by generalisation, as is the case with positivism, because the research environment does not remain the same (Yeasmin & Rahman, 2012b:155). Myers and Klein (1999:70) point out that, once the meaning is understood within the context, logical links between related issues can be used for generalisation. Logical reasoning that is applied to abstraction and generalisation is important; it must be implemented critically (Walsham, 2006:325).

Principle five is concerned with dialogical reasoning. For example, a court of law reviews evidence presented; the judge can then either declare the suspect free or

sentence him/her. Myers and Klein (1999:72) argue that the same principle must be applied to interpretive research; the researcher must present, analyse, and review the evidence and actual findings that will either prove or contradict the theoretical preconceptions guiding the research design. This process is called reasoning ideologically (Walsham, 2006:325). Ideological reasoning results from the discussion of ideas, interviews, beliefs, viewpoints of participants, and the role played by a researcher. An interpretive researcher must have an open mind in a study; he/she must accommodate contradicting view points and build them in rather than discard them.

The sixth principle is about multiple interpretations. A research problem can produce different perspectives from different participants from the same sequence of events under study. Participants attach their meanings, which may not be exactly similar. An interpretive research world is a multidimensional world where opposition, critique, and conflict enrich the understanding of a researcher; thus, it requires an open mind (Myers & Klein, 1999:73).

The seventh (and last) principle (suspicion) suggests that the active role of a researcher can possibly distort data or show some biases on how narrative data was collected from participants (Myers & Klein, 1999:73). This principle implies that the researcher must also be attentive not to be misled by participants.

2.5.1.1 Position of the study (Interpretive principles applied)

A number of research strategies may be used to ensure that data are credible. The set of interpretive research principles of Myers and Klein (1999:72) on how to conduct and evaluate interpretive research are applied in this study. Table 2-3 gives a summary on how the principles are applied in the study.

Table 2-3: Interpretive research principles

Principles	How it is applied
The fundamental principle of hermeneutics circle.	The researcher's aim is to understand the participants' views about current DW/BI education, and what they think it ought to be. The understandings of interview answers (text) are never isolated. It is interpreted in the

	context of other answers that were provided to questions that preceded them. The understanding is developed through the understanding of individual questions and the answers that are provided.
The principle of contextualisation	This principle is used to gain context, clarification, and relevance about the history of DW/BI education in institutions of higher education. Context plays a very important role in understanding the current participants' viewpoints. Specific questions are asked to provide contextual knowledge.
The principle of interaction between the researchers and the subjects	Data is collected using active and passive communications methods. Care is taken to ensure that active participation in research does not influence or compromise the quality of data collected. Open- ended and follow-up questions are used to ease the discussion.
The principle of abstraction and generalisation	The meaning of data is contextualised within the context of DW/BI education; logical links between related issues building up to BI education are used to develop guidelines with a wider application.
The principle of dialogical reasoning	The evidence is presented, analysed, reviewed, and the actual findings prove or contradict the theoretical preconceptions guiding the research design. No data are disregarded.
The principle of multiple interpretations	Participants attach their meanings that may not be exactly the same as that of the researcher. The multidimensional aspect of this study is in such a way that opposition, critique, and conflicting views are used to enrich the understanding of the DW/BI educational needs of different stakeholders.
The principle of suspicion	The researcher applies the principles of suspicion throughout the study to ensure consistency. Critical thinking is also applied in conceptualising the study, collecting the data, interpreting the findings, and reporting results processes.

Participation selection methods are discussed in the next section.

2.6 Interpretive research methods

Research methods are used to conduct the actual research. Methods are classified into data collection methods and data analysis methods. Prior to data collection, decisions need to be taken on whom to involve as participants in the study. The applications of research methods are guided by the chosen research paradigm of the study; it is, in this instance, interpretivism.

2.6.1 Participants selection

Before data collection can take place decisions need to be taken regarding participant selection. Participant selection in interpretive research is based on purposeful selection of participants; the number of participants is not generally

predetermined (Sargeant, 2012:1, Turkett, 2004:47). According to Sargeant (2012:1) purposeful participants refer to those who can best inform the research questions and enhance the researcher's understanding of the phenomenon; the total number of "participants depends upon the number required to fully inform all important elements of the phenomenon being studied". The research elements are fully informed (rich) when the data saturation point is reached (i.e. no further concepts are identified) (Turkett, 2004:47). It aims to provide a wide variety of perspectives; the more perspectives you have the better is your understanding (Ulrich, 1996:288).

2.6.1.1 Position of the study

For the purpose of this study, the required information is the understanding of the former students, lecturers, and industry experts with regard to the skills graduates should possess when they enter the industry, and the process of developing a DW/BI module. Participants for the interview process are identified by the roles they play in DW/BI education – refer to Chapter 7. Groups that represent specific participants, i.e. former students, BI/DW professionals, and faculty members are viewed as different stakeholders. These groups are chosen to achieve more perspectives on the problem situation, in order to better understand it. This is in line with the ideas of Ulrich who argues that more "conditioned realities" are required to improve our understanding of a system; he refers to "conditioned realities" because our perspectives are conditioned by our prior experiences and knowledge of reality (Ulrich, 1983:259).

2.6.2 Data collection methods

A research study requires a detailed outline of how data will be collected. Data collection methods must be relevant to the research study, and within the context of a particular paradigm (Wahyuni, 2012b:72). The goal of interpretive research is to understand human thought as a social phenomenon (Myers & Klein, 1999:67). Therefore, the data collection period must be long enough to allow for better understanding and re-collection if necessary; it must be coupled with multiple data collection methods. Analysis involves thinking about data and ideas, writing memos, exploring etc. To incrementally build theory, more data must be collected. Thus, analysis guides data collection further; each step of data collection feeds into

analysis. When applied in a grounded theory method, this iterative process between data collection and analysis continues until saturation is reached (Glaser & Strauss, 1967:61).

2.6.2.1 Grounded theory as data collection method

Grounded theory was first published in 1967 by Glaser and Strauss (1967:61). It is a research method used to generate theories regarding social phenomena (Lingard *et al.*, 2008:459). Glaser and Strauss (1967:61) define grounded theory as “the discovery of theory from data systematically obtained from social research”. It is a method used to build the base of a research study, because it a method used to develop theory (Jones *et al.*, 2005). Grounded theory is discussed as both a data collection and a data analysis method. A researcher can use grounded theory to formulate knowledge or understanding that is based on systematic analysis of data (Bitsch, 2005:77). Grounded theory is characterised by continuous interplay of data collection and analysis until at least saturation of concepts is reached. This means that full development of a category is exhausted, not the data itself (Glaser & Strauss, 1967:61). Coding techniques are presented in the Section 2.6.3.1 on the analysis aspects of grounded theory.

2.6.2.2 Interviews

Interviews are used to gather the meanings that people attach to the world. Seaman (1999:558) describes three types of interviews that can be used in interpretive research. Firstly, structured interviews are conducted with pre-defined questions in hand. Structured interviews are good to focus on specific topics and avoid deviations. Secondly, unstructured interviews have open ended questions. Open ended questions let the interviewee go beyond the asked question; it allows them to bring their understanding and experience without being channelled. Thirdly, semi-structured interviews combine structured and unstructured interviews; the intention is to combine the strengths and overcome the shortcomings of the one methods. An interview consisting of open-ended questions is designed to guide the discussion between researcher and participant.

2.6.2.3 Ethnography

Ethnography is a data collection method that permits the researcher access to peoples' meanings and views without removing them from their natural settings (Atkinson, 1998:248). A researcher becomes an active participant and part of environment for a longer period of time. According to Atkinson (1998:248), ethnography explores phenomena instead of proving it; it works primarily on unconstructed data and the explicit interpretations of the meanings people attach thereto. Elliot and Jankell-Elliot (2003:216) say ethnography is a method that requires natural settings, allows one to see the world through the eyes of participants, extends presence in the field, and enables active participation of a researcher (walk a mile in their shoes).

2.6.2.4 Collection of relevant documents

One of the research starting points is existing documents that can be used as a first set of data in interpretive research. A documentary analysis is conducted on this data to understand substantive content or to illuminate deeper meanings, and understand the relevance of historical events/experiences (Crinson & Leontowitsch, 2011). For example, in this study, document data are gathered on DW/BI modules at universities.

2.6.2.5 Position of the study

Data collection from documents is used in this study since data from university programmes are used in the initial phases of the guideline development process. The literature review guides the development of semi-structured interviews with stakeholders by providing context. More detail is provided in the following subsections.

2.6.2.5.1 Primary data collection

Interviews with participants are the main data collection method in this study (refer to Chapter 7). Semi-structured interviews are conducted using questions grounded in literature of DW/BI (refer to Chapter 3) and guided by critical systems heuristics (CSH) and soft systems methodology (SSM) (discussed in Chapter 6). The tools

provided by SSM are used to represent the perspectives of different stakeholder groups.

2.6.2.5.2 Secondary data collection

Document data is gathered on DW/BI modules at universities. In the preliminary phase the researcher identifies which institutions offer DW/BI modules. A selection of the top universities in South Africa are identified according to university ranking by academic performance (URAP) (performance, 2015). This process is discussed in Chapter 5.

2.6.3 Data analysis

Data are collected and then organised and analysed. Interpretive research data enquiry processes mostly result in qualitative data that necessitate different analysis methods and techniques. This analysis describes facts, patterns, and trends; it also detects abstract thought, understanding, self-awareness, reasoning, learning, planning, and problem solving (Krauss, 2005b:758).

2.6.3.1 Grounded theory as analysis method

Grounded theory as an analysis method is used to understand social phenomena that are not abstract (Glaser & Strauss, 1967:61). It is used to uncover basic social processes such as integral social relationships and the behaviour of groups. The rigor of grounded theory comes with a continuous theory development process. A boundary of proper scope of the area of investigation must be set. Continuous sampling of data and analysis goes to a point that no more theories can be developed in the defined scope. The concepts that are identified from the theory and their relationship must be confirmed. Data analysis is achieved by using a coding method.

Coding is used to extract patterns from qualitative data for quantitative analysis (Seaman, 1999:557). According to Glaser and Strauss (1967:80) there are three different level types of coding a researcher can use for data. Level one is open coding where conceptual categories are formulated from sorted raw data. On level

two axial coding is applied to find and verify the relationship between the identified categories. In level three selective coding is applied to connect categories that build the story.

Since the researcher chooses the concepts to focus on, Walsham (2006:325) argues that grounded theory has an element of coding that is subjective. He suggests a loose approach where the researcher writes impressions during the research; after each interview he/she then organises them into sets of themes and issues. The emerging new data are used to develop, enhance, confirm, or even discount theories. The open, axial, and selecting coding are the three types that can be used in this process (Glaser & Strauss, 1967:80). This process is called the constant comparison method.

2.6.3.2 Content analysis with interpretive coding

The artefacts of social communication, such as written documents or transcripts of recorded verbal communication, must be examined. Data that are in text, visual, and verbal communication form carry literal meanings as well as context of use to it. Harvey (1990:8) argues that “data is meaningful only in terms of its theoretical context, reliability and validity are functions of the context and the epistemological presuppositions that the researcher brings to the enquiry.” This type of data that represent social communication requires content analysis that goes beyond counting words and extracting objective meanings attached to it. Content analysis is “any technique that is used to make inferences by systematically and objectively identifying special characteristics of messages” (Holsti, 1968:608). Content analysis is a method used to infer validity from context (Elo & Kyngas, 2007:108, Schadewitz & Timothy, 2007:2). The main objective is to exhaust all content analysis to account for any variation of message content; other researchers must obtain the same result when looking at the same message (Cagdas & Stubkjaer, 2011:173, Holsti, 1968:608). Coding methods similar to those used in grounded theory are used to extract patterns from the data.

According to Zhang and Wildemuth (2009a:5) “qualitative content analysis uncovers patterns, themes, and categories important to a social reality.” Zhang and Wildemuth

(2009a:5) recommend eight process steps for content analysis. Data are gathered, units of analysis are identified, and categories of schemes are developed. The development of categories involves developing coding schemes for inductive and deductive content analysis. Deductive reasoning that generates concepts or variables from previous studies or theory is largely used. The coding scheme is about testing on sample data for consistency before all text is coded. On completion of all text coding, conclusions are drawn from the coded data.

The relationship between theory and research becomes evident when they are both looked at in inductive and deductive approaches (Blackstone, Elo *et al.*, 2014:3). In an inductive approach, collecting substantial amounts of data that are relevant to the research study, checking for patterns, and working to develop a theory that could explain those patterns are done. Figure 2-7 illustrates the inductive content analysis approach.

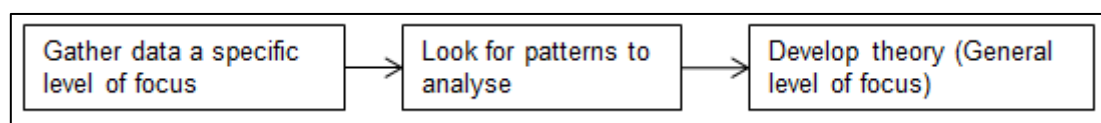


Figure 2-7: Illustration: inductive approach

A deductive approach begins at a general level, and moves to a more specific level. The compelling social theory and their implications are tested with data (Elo *et al.*, 2014:3). Figure 2-8 illustrates deductive content analysis approach.

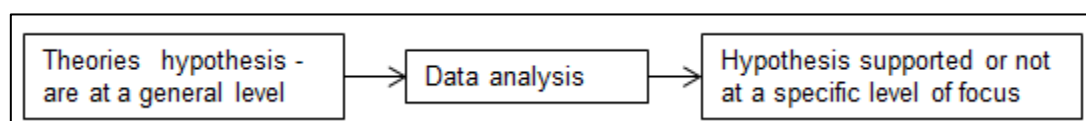


Figure 2-8: Illustration: deductive approach

2.6.3.3 Cross case analysis

To develop a theory (understanding) from qualitative data, cross case analysis is used to analyse stories and views that are treated as cases (Darke *et al.*, 1998b:273). This method is applied to analyse data across all of the cases in order to identify similarities and differences between them (Meyer, 2001:331).

Cases are compared and differences are highlighted. Codes are added after analysis of each interview. The researcher makes notes from reading through text and generates codes and patterns between them. Direct interpretation is used to develop naturalistic generalisations and present an in-depth picture.

2.6.3.4 Position of the study

Interpretive coding and cross case analysis are selected as applicable methods for this study. Inductive coding are also used where no theories are available and they only emerge through data examination and constant comparison (Zhang & Wildemuth, 2009a:3). The eight process steps of Zhang and Wildemuth (2009a:5) is applied to the study – refer also to Chapter 7. These steps include: preparing data, define units of analysis, develop categories of schemes, test your coding schemes on a sample of text, code all the text, assess the coding consistency, draw conclusions on the coded data, and report your findings. This process is detailed in Chapter 7.

2.6.4 Research evaluation: trustworthiness of the study

A research evaluation process should be in place in order to audit the events, influences, and actions of the researcher in a study (Carcary, 2009:15, Darke *et al.*, 1998a:273). Lincoln and Guba (1985:290) recommend six prerequisites that need to be documented to inform the audit trail: First, raw data must be collected. Second, data reduction and analysis notes must be kept safe for reference purposes. Third, data reconstruction and synthesis products must be kept. Fourth, process notes, such as procedures, designs, strategies, and rationales must be made available. Fifth, materials related to intentions and dispositions, such as inquiry proposal and personal notes must be traceable. Last, preliminary development information must also be traceable.

The credibility of a research study depends on many aspects; one of them is the evaluation and processing of data. The evaluation process must have an element of trustworthiness. Trustworthiness is the quality of an investigation and its findings that makes it noteworthy to audiences. According to Lincoln and Guba (1985:290)

credibility, transferability, dependability, and conformability are key trustworthiness elements that interpretive studies must satisfy.

Credibility is concerned with the truthfulness of the data collected (Lincoln & Guba, 1985:300). Measures must be taken to ensure that collected data are not fabricated, and that the researcher does not influence the research outcome (neither positively nor negatively). Conformability of the findings is the degree to which the data and data interpretation are based firmly on evidence collected, rather than the personal construction of the researcher. To increase the integrity of research, a negotiated outcome where the researcher collects data and the participants agree or disagree with a researcher, can be used to ensure quality and confirm findings.

The transferability element refers to the generalisability of the study (i.e. the results of research in one situation relate to other similar situations or environments) (Lincoln & Guba, 1985:300). For example, in positivistic work the researcher seeks to demonstrate that the results of the study at hand can be applied to a wider population. However, findings of interpretive research are specific to a particular environment and specific people; thus, transferability can be difficult to achieve. The researcher should keep detailed notes and document the analysis of data collected from interviews, observations, and documents. Decisions taken by a researcher must be traceable; evidence must come from the research. For example, the researcher should be able to trace final data presented back to themes, codes, and the actual scripts.

Dependability pertains to the rigour associated with the process of inquiry (Lincoln & Guba, 1985:301). Consistency of the researcher in the study is assessed. The effort and care of the researcher is checked to see if there are no careless mistakes in processes pertaining to conceptualising the study, collecting the data, interpreting the findings, and reporting results. Clear presentation of processes as well as logic used to select participants, events to observe, interviews, and everything included in the study should be presented.

Accuracy, and elimination of biases, is increased by use of a triangulation method. It entails combining multiple observers, theories, methods, and empirical materials to avoid intrinsic biases and the problems that come from single-method, single-observer, and single-theory studies; it convergences different perspectives (Yeasmin & Rahman, 2012b:153). Triangulation offers a researcher a validation process of comparing different methods to collect data; replication of result outputs are not possible in interpretive research because of natural settings.

2.6.5 Position of the study

Care is taken in this study to demonstrate an auditable process to develop guidelines. Code labels are assigned to codes in different chapters to ease linking of findings to literature and data. Initial guidelines are identified and labelled in Chapter 4 from DW/BI literature; these guidelines are then enriched from data collected from university programme documentation in Chapter 5. The empirical data collected from interviews Chapter 7 are linked to these guidelines; the origins of the data processed into guidelines are identified by code labels to improve traceability.

The items of information proposed by Lincoln and Guba (1985:300) are stored and will be available from the researcher on request. The interpretive principles that are developed by Myers and Klein (1999:72) are used as an instrument to evaluate the trustworthiness of the study; this is to achieve trustworthiness in this study as – refer also to Section 2.5.1.1 of this chapter.

2.6.6 Ethical considerations

All participants were advised about the research process; their participation was with their consent. A clear statement of the research goals was communicated so that intervention and advocacy could be implemented if the participants identified elements of harm or illegal acts.

The integrity of participants is protected by not disclosing information that may lead to their identification. Participants were given the opportunity to verify their input as a way of checking that there is no misrepresentation.

The data collection strategy of this study was approved by the ethics committee of the School of IT of the North-West University where the study is undertaken.

2.7 Concise research plan for this study

The aim of this chapter is to provide justification for the selection of the interpretive research paradigm as well as for the selection of specific interpretive research methods. This is achieved in the section labelled “Position of the study”; it is after the general discussions and based on research methodology literature. Together, these sections form the research plan of the study. The research plan of this study is depicted in Figure 2-9.

The aim of the study is to develop guidelines for DW/BI modules in IT programmes at universities; the guidelines are based on an understanding of the needs of various stakeholders. The process to develop the guidelines is presented in a traceable manner. A literature review of DW/BI concepts (Chapter 3), together with a well-known mapping of skills for IT professionals, is used to develop of a list of skills required by DW/BI professionals (Chapter 4). Since many universities in South Africa currently present some form of DW/BI programme content (presented in Chapter 5), the researcher decided to incorporate this current knowledge in the guidelines. Programme documentation from universities is viewed as secondary data in this study. The initial set of skills from literature, enhanced by the programme content from universities’ programme documentation, is presented in Chapter 5 in terms of: prerequisite skills; DW/BI specific skills; and advanced general IT skills that are required by DW/BI professionals.

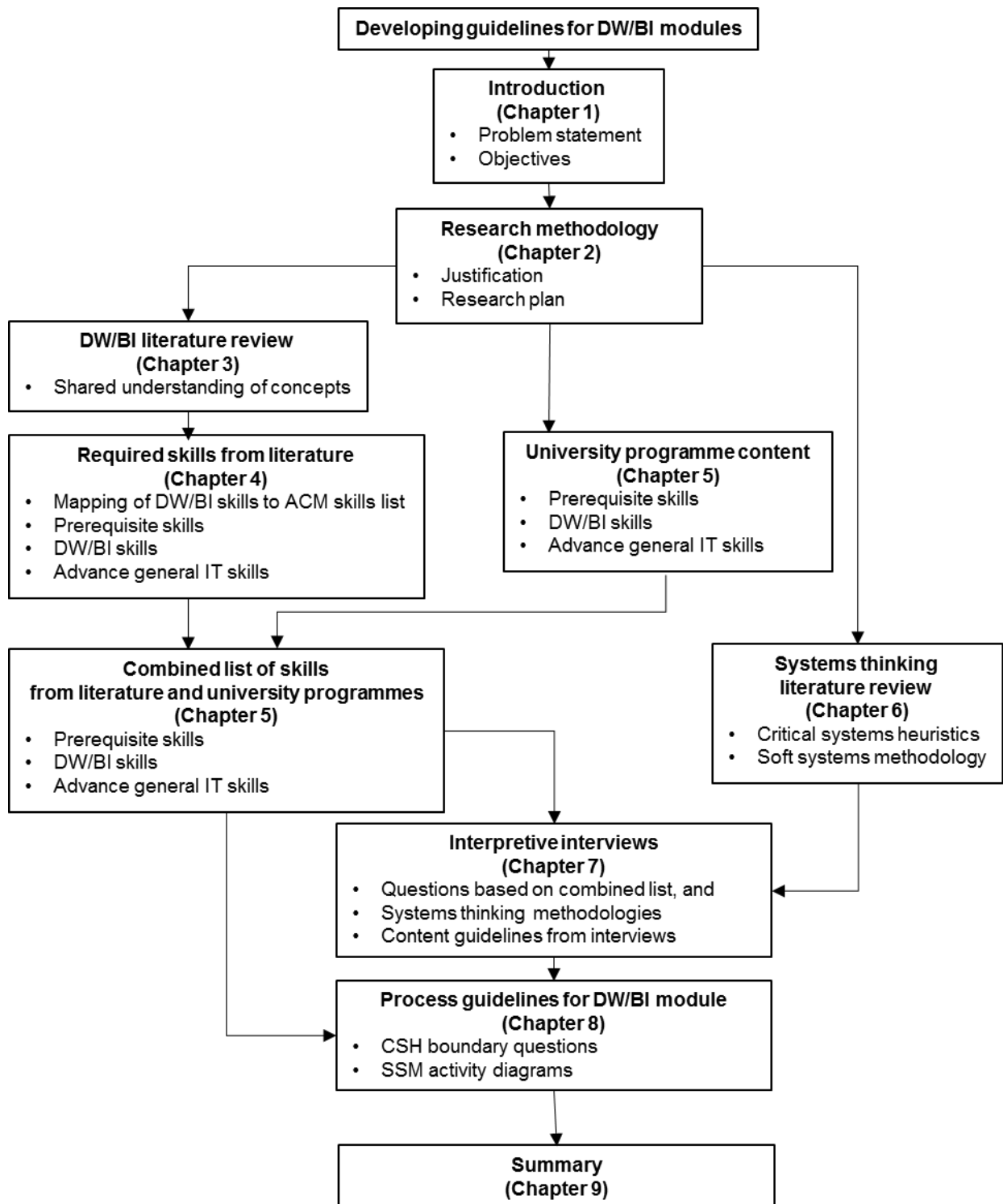


Figure 2-9: Research plan

Critical systems thinking methodologies (CSH and SSM) are used to guide the identification of the stakeholders, and to present the worldviews of these groups of stakeholders. Stakeholders identified in this study includes: former students currently employed in the DW/BI or IT industry; experienced DW/BI professionals; and faculty members (lecturers) at universities. CSH provides a set a questions that can be used to articulate a worldview in terms of sources of: motivation, control, expertise required, and legitimation (Ulrich, 1983:258). SSM, as described by Checkland and Poulter (2006:1), provides tools for participants to present their different worldviews in order to articulate the course of action they believe should be required to achieve improvement in a problem situation. Both CSH and SSM are discussed in Chapter 6 and are used, together with the results of Chapter 5, to develop an interpretive interview in Chapter 7.

Chapter 7 provides the interview development as well as data analysis of the interpretive interviews conducted with participants that are representative of the identified stakeholder groups. Semi-structures interviews are conducted and analysed with content analysis and cross case analysis. Findings presented are supported by coding tables used in data analysis.

Analysis of the interview data can be divided into: findings in support (enhancement) of the content of the DW/BI module based on the set of skills developed in Chapter 5; and also findings that guide the process of developing a DW/BI module at universities. Analysis of the interview data in terms of module content is summarised after the presentation of the data in Chapter 7 while analysis of the findings of the interview data in terms of the process of developing a DW/BI module is presented in Chapter 8.

An updated list of skills required is presented as a set of guidelines for DW/BI content, together with a set of guidelines for the process of developing a DW/BI module at universities in Chapter 9.

This process is depicted in Figure 2-9, which serves to guide the reader throughout the study

2.8 Summary

In this chapter the principles for good guidelines are presented in order to bring clarity on what is a guideline is, and how it is used. The position of the study on how the principles of guideline development process are applied is presented to guide and justify the research plan of the study.

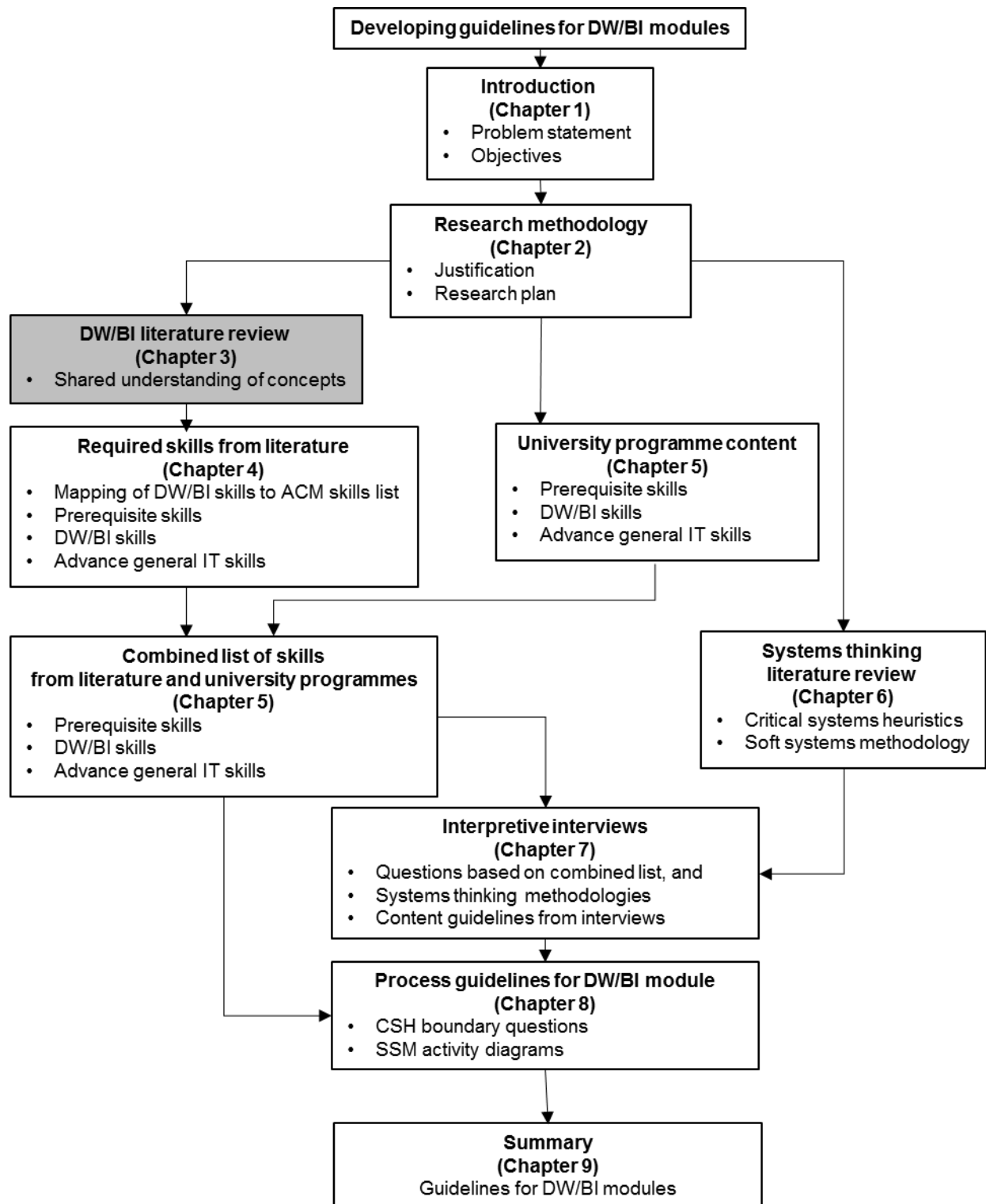
Four paradigms, i.e. positivistic, interpretive, critical social theory, and design science are discussed. The interpretive research paradigm is adopted as the paradigm for this research study. Background on the interpretive research methodology is presented.

Interpretive research methods provide information on the research methodology (participant selection, data collection, and analysis). A position of the study, or how the methodology is carried out in the study, is also presented at the end of each section.

The chapter concludes with a diagrammatic view of the entire study that shows the development of the guidelines, for both the content and development process, for DW/BI modules in IT programmes at universities.

A literature review on DW/BI is presented in the next chapter, in order to achieve a shared understanding of DW/BI concepts used in the guidelines.

This page is left open intentionally



Chapter 3 Business intelligence and data warehousing

3.1 Introduction

Business intelligence is continually being implemented across many industries; however, “the one thing that is not being addressed is where the employees are going to come from to fill those analytical jobs” (Bradbourne & Rouse, 2011). These future employees must be educated appropriately. The objective of this study is to develop guidelines for data warehousing/business intelligence modules in information technology programmes at universities; the guidelines are based on an understanding of the needs of various stakeholders. Two methodologies, i.e. critical systems heuristics and soft systems methodology, are used to understand the various stakeholders’ perspectives. The purpose of this chapter is to discuss the fundamental aspects of business intelligence and data warehousing, so as to ensure shared understanding of these underlying concepts for the remainder of the study.

The potential of computers in business was realised in the early 1960s (Hedgebeth, 2007b:415). Business intelligence is a discipline that emerged out of the use of computers in business to improve decision support functions and/or processes. Immense quantities of available data in businesses necessitated efficient extraction, storage, transformation, sharing, and disposal processes so that the data could be used effectively for decision making (Delen & Demirkan, 2013:359). Business intelligence is rooted in other disciplines; decision support systems, executive information systems, data warehousing, and online analytical processing are all predecessors of business intelligence (Bucher *et al.*, 2009:408, Power, 2007, Zeng *et al.*, 2006:4722). The discipline of business intelligence has a rich history; however, as it stands today, it will not necessarily serve future users of business intelligence efficaciously.

This chapter is structured as follows: Section 3.2 discusses the evolution of business intelligence. Section 3.3 focuses on BI and analytics. Section 3.4 describes business intelligence components such as data, processes, tools and technologies, as well as

human talent and education. Section 3.5 discusses data warehousing and Section 3.6 presents the data warehouse life cycle. Maintenance and growth of a business intelligence solution is presented in Section 3.7. The future of business intelligence is discussed in Section 3.8. This chapter concludes with a summary in Section 3.9.

3.2 Evolution of business intelligence

A decision support system (DSS) is a computer-based system, used in semi-structured or structured situations, to provide support for decision makers (Cassie, 1997:293). These systems were created in the 1960s, followed by the theory around them in the 1970s (Hedgebeth, 2007a:415, Power, 2007). The theory development and the application of DSSs led to their spanning into specialised types namely text-oriented, database-oriented, spreadsheet-oriented, solver-oriented, and rule-oriented DSSs (Power, 2007).

Challenges, opportunities, and growth in industry created the need for a better way to deliver information to decision makers. Executive information systems (EISs) evolved from single-user, model-driven DSSs (Zimlichman & Meilek-Weiss, Sharmila & Subramani, 2013:3). EISs are tailored for executives' information-related requirements and are therefore based on requirements from them (Zairi *et al.*, 1998:50). An EIS is referred to as a subset of business intelligence (BI) software; it is built to support managerial decisions (Taleghani, 2011:1037).

The growing dependence of organisations on data created a need for data management (Munawar & Roliana, 2011:1119). Data uploaded in EISs necessitated better data management approaches. Consequently, due to high demand of data stored in multiple sources to provide insight to decision makers, data warehousing (i.e. processes to collect, transform, store, and present data to users for decision making) evolved (Hua *et al.*, 2012:53). A data warehouse (DW) is a central repository of data that provides a single, yet comprehensive, source of current and historical information; it is mostly used to support DSS and EIS (Chountas *et al.*, 2003:3076, Hajmoosaei *et al.*, 2011:290, Ou & Peng, 2006:760, McFadden, 1996:120). Online analytical processing (OLAP) tools leverage data stored in the DW to execute different types of analysis operations for users (Zhao, 2008:97).

It is evident that BI is rooted in these other disciplines and theories that preceded it. Organisations must decide how to adopt and implement BI in business process modelling. Data can be an afterthought to business process modelling. This implies that BI is a supporting function instead of being part of the core processes (Ou & Peng, 2006:780). Parallel execution of data extraction, storage, retrieval, and analysis, as well as retrieval of information, are separated from business process activity workflows (Chen *et al.*, 2012:1166).

The need to provide information to all levels of management and supporting staff, brought about an enhanced business requirement to incorporate BI into activity workflows in business process modelling. This implies that BI evolution can therefore not be separated from the understanding of the current industry. Business processes bring context, value, and understanding of the business into information produced from data. Experts in the field advocate that BI evolved into two types, i.e. data-centric BI or process-centric BI (Ou & Peng, 2006:780). Operational and real-time BI surfaced to close this gap between processes and data (Bucher *et al.*, 2009:408). (Sahay & Ranjan, 2008:30) describe operational and real-time BI as two interlinked terms to mean “a true up-to-the-moment view in which the information gleaned is actually current enough to be useful in managing and executing business processes”. This approach favours integration of BI and operational process execution. Process-centric BI brings context to data.

Business decisions must be motivated (Marshall & Harpe, 2009:2). The best way to motivate a business decision is by having reliable and consistent supporting data that show historical trends that can be used to predict the future (Few, 2005). Many organisations use BI to enhance their competitiveness in the market (Kapoor & Sherif, 2012:1626, Ranjan, 2009b:63). The purpose of implementing BI is to enable business users to collect, store, transform, process, analyse, and convert data into information that informs decisions (Marshall & Harpe, 2009:2, Dalrymple, 2011:41).

3.3 Business intelligence and analytics

The terms BI and analytics are used together or interchangeably; however, when singled out, analytics usually supports BI (Hostmann *et al.*, 2009:17). For example, a

telecommunication company may want to know what causes network usage fluctuations during working hours. This question prompts for an analytics project that will attempt to provide answers to this question. The BI tools are then used to provide analytical output to the users in a format that is meaningful them. Analytics is taking things apart in a specific domain using analytical techniques such as predictive models and analytical applications (Hostmann *et al.*, 2009:2). Data analytics is grounded mostly in data mining and statistical analysis (Chen *et al.*, 2002). Figure 3-1 below depicts levels of analyses that can be accomplished using BI and analytics.

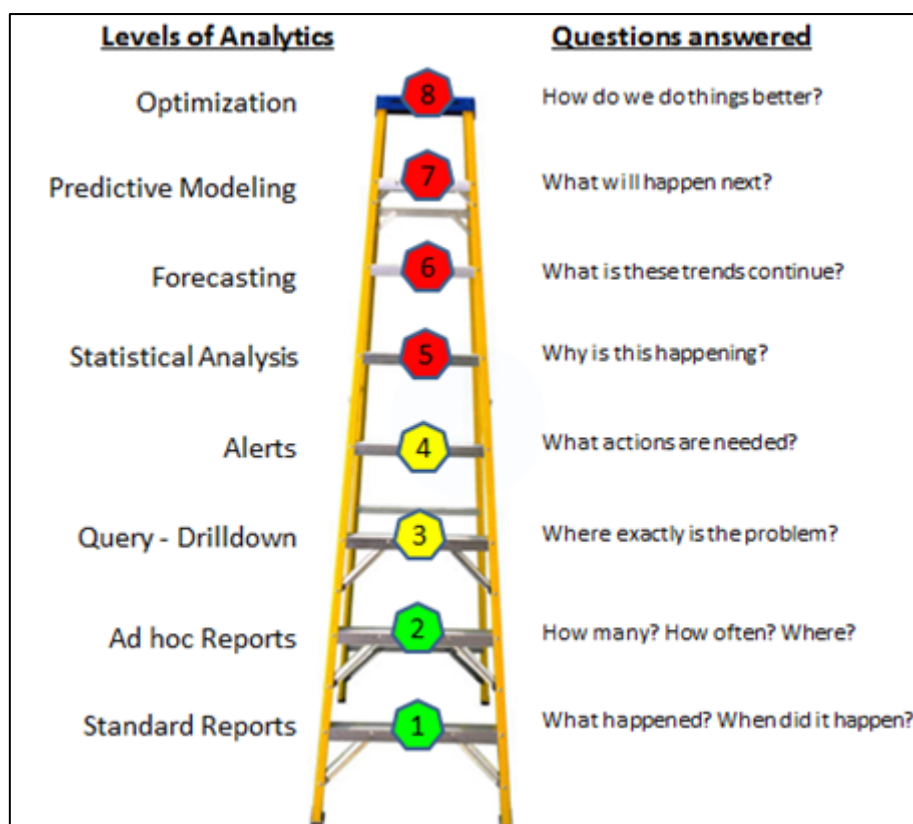


Figure 3-1: Business intelligence and analytics usage (Mirji, 2011)

Data are meaningful if transformed into usable information; thus, analytics enables organisations to move from having data towards having information (Barlow, 2013:8). Analytics can be used by an organisation to achieve business objectives through reporting of trends, predictive modelling, as well as analysing business processes and opportunities. Descriptive, predictive, and prescriptive are three popular categories of analytics (Delen & Demirkan, 2013:361).

Descriptive analytics includes standard, ad-hoc, periodic, and dynamic reporting; OLAP reports answer daily operational questions such as “what happened?” or “what is happening in the business”? As an example, consider a case where a telecommunication company wants to investigate the total call charges for subscribers at a specific point in time. The report answering this question warrants a descriptive answer of an event or process that already took place.

Predictive analytics seeks to identify patterns, associations, affinities etc.; it uses mathematical and discovery techniques and methods. Predictive analytics is future-orientated and it gives us the probability of different outcomes (Payandeh, 2013). For example, predictive modelling can be used to identify the most profitable customers, those with the most profit potential, or even those that will most likely stop being customers (Davenport, 2006:4). Telecommunication service providers can attempt to make use of existing data to prevent losing its customers and also to increase profits from its existing customer base. Predictive modelling uses algorithms and machine-learning techniques that are ideal for analysing large data sets in order to make connections between bits of information (Wolniewicz, 2003). Additionally, predictive modelling may also suggest specific actions that may aid to retain a customer (Davenport, 2006:4).

Prescriptive analytics uses mathematical and discovery techniques to discover alternative courses of action that can yield greater value in comparison to others (Payandeh, 2013). Prescriptive analysis is usually applied in complex problems with more than one variable, which add to the complexity of the problem. As an example, a telecommunication service provider can seek to increase its revenue by applying models to see if free or cheaper short message services (SMSs), multi-media messaging services (MMSs), data bundles, airtime, or contract incentives may attract more (new) customers. This process involves simulations with variables to determine the best possible option.

Although analytics can help most organisations to realise their goals, Davenport (2006:4) reports that only a handful of companies achieve the expected level of proficiency using analytics. Davenport (2006:3) suggests that the key success

factors in analytics are: (1) the widespread use of modelling and optimisation (i.e. looking beyond basic statistics), an organisational approach (i.e. rolling out analytics solutions in the entire business), and (2) senior executives advocate. Figure 3-2 below depicts a summary breakdown business analytics types.

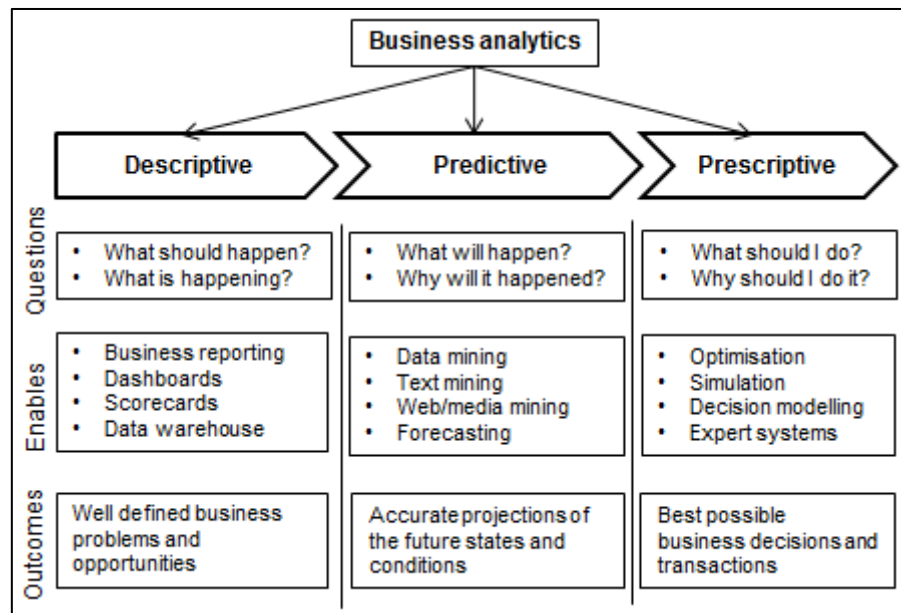


Figure 3-2: Analytics categories

3.4 Business intelligence components

Business intelligence is an umbrella term used to describe concepts and methods to improve business decisions that are based on facts and derived from data (Al-Qirem, 2012a:93). BI includes data, technology, analytics, and human knowledge. The implicate acceptance of data warehousing as part of term “BI” motivated the researcher to use the extended term: “DW/BI” to include the notion of DW in BI explicitly. This is in accordance with Kimball *et al.* (2008a).

3.4.1 Data

Data are stored and retrieved in different formats that are either structured or unstructured (Ranjan, 2009b:62, Negash, 2004:170). Depending on the context, unstructured data can mean different things, i.e. conversations, graphics, images, news items, Excel spreadsheets, text, videos, web pages and others (Blumberg & Atre, 2003:42). Structured data is in a format that can be stored in OLAP cubes,

database management systems (DBMSs), data mining systems, enterprise resource planning (ERP) systems, EISs, and DSSs (Negash, 2004:178).

The term data has many definitions, which matured with time and context (Zins, 2007:479). Hey (2004:1) defines data as “discrete, atomistic, tiny packets that have no inherent structure or necessary relationship between them”.

Both structured and unstructured facts types are used for reporting in BI. Discarding data is not a goal of BI, since there can be semi-structured data that does not fit neatly into relational or flat files (Negash, 2004:180, Papakonstantinou *et al.*, 2002:1). There is high demand for reporting on unstructured data, which resulted in a lot of effort and innovation for tools that can perform this task (Blumberg & Atre, 2003:42). Most BI tools’ functionality are tied to structured relational databases, which limits the resulting experience of users for analysing unstructured facts (Papakonstantinou *et al.*, 2002:1).

3.4.1.1 Data Hierarchy

An opportunity or problem in a business necessitates the gathering of data, which will be used in the process leading to decision making (Ou & Peng, 2006: 780). A BI solution’s purpose is to know and understand the capabilities, opportunities, and challenges presented by the data in an organisation (Few, 2005). Quantitative information, i.e. numbers to measure performance, identify opportunities, and forecast the future, is what drives businesses. Data, information, and knowledge are the building blocks of information science; these are therefore the focus of BI (Zins, 2007:479). Figure 3-3 below depicts the linear or sequential representation of data, information, knowledge, and the extension of wisdom; the extension of wisdom is at the top of this hierarchy. The sequential interrelations of data, information, and knowledge are still debatable to many scholars (Zins, 2007:479). However, this is beyond the scope of this research study. The interrelations can, instead, be reasoned or logically linked.

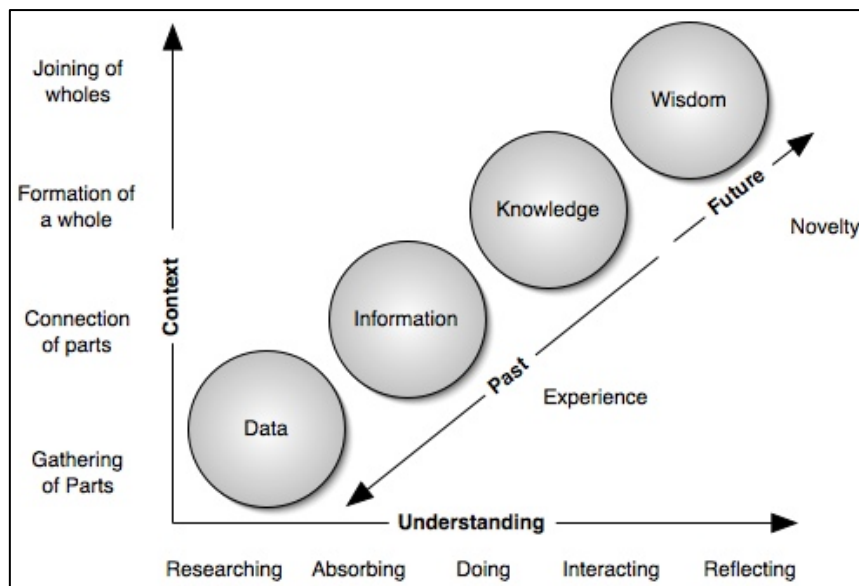


Figure 3-3: Data hierarchy (Clark, 2004b)

Business intelligence projects usually emerge as a result of problems/opportunities that need to be explored, analysed, and understood; dependant on their nature, they may be solved (researching the problem). This process requires gathering of data. Data that are stored must be absorbed and understood in terms of logical links between them. Through understanding and logical links, information can potentially be formulated. Information adds context to data for it to be meaningful (Hey, 2004). Information and knowledge are derived based on research or experience (Stephens & Plew, 2001:19). Data carry the meaning of information if it can also be presented (Clark, 2004a). Knowledge materialises when human beings have the capacity to understand, explain, and reason about concepts and present it with the information gained (Zins, 2007:479). Wisdom tells us when and why to use information. The value of data in an organisation is not in owning the data, but the value lies in the interpretation of the information.

3.4.1.2 Data storage and retrieval

Databases are built to manage formatted data that are organised in patterns, structures, and formats; data are then stored and retrieved (Negash, 2004:177). Data organisation includes indexing, sequential filing, direct accessing, indexing of files, storing, and retrieving (Chiodo *et al.*, 1994:433). Models such as fuzzy set, indexes, search trees, utility, and other probabilistic models are used to retrieve data.

The greatest challenge about unstructured data is based on the need for the data to be searchable (Blumberg & Atre, 2003: 42). The challenge is not only limited to search, but also includes how to store, index, and maintain data.

The efficiency of BI in the organisation is attributed to many factors; one of them is the education and knowledge BI users (Owei *et al.*, 2006:77). Information users must know what to search for and how to search using BI tools. The speed at which data can be stored and retrieved is an important factor for information workers and decision makers. These data sources must be easy to use in order to allow users to store and retrieve data. Hardware also plays a very important role in terms of processing speed and power, as well as random access memory allocated.

3.4.2 Processes

Daily activities that take place form business processes (Stephens & Plew, 2001:19). BI is an end-to-end process that has other processes in itself (Sharma & Djiaw, 2011:114). This means that there are dependencies or interdependencies amongst BI processes. For example, the extraction and transformation processes must be completed before a data warehouse load. The success of a BI solution depends highly on its well implemented processes involving technology and users (Kaniçlıdes & Kimble, 1994).

Processes define continuous flows of activities; hence, Ross (2009) advises that a BI solution must always be built around business processes and not the other way round. Taking an order, and billing a customer, are activity examples of business processes. BI and especially data warehousing operational processes, such as data extraction, data transfer (and loading), data transformation, data cleaning, and computation, are then built in support of business processes (Vassiliadis *et al.*, 2001:210). These processes are usually represented in a workflow of activities (Karagiannis *et al.*, 2013:927).

3.4.3 Tools and technologies (hardware and software)

Market changes, industry competition, technological advancements, and innovative thinking led to the development of various BI tools (Thoo *et al.*, 2013:1). The term BI was coined in the 1980s; since then, spreadsheets, OLAP, ETL, and many others tools have been developed (Sharmila & Subramani, 2013:2). Both commercialised and open source tools such as RapidMiner, SpagoBI, Eclipse BIRT Project, and others entered the market.

Business intelligence and data warehousing tools and technologies are components that make reporting possible (Ma *et al.*, 2000:127). BI tools have been linked to powerful use of Microsoft Excel, only with more extended capabilities that deliver powerful analysis and knowledge discovery capabilities (Nadeem & Jaffri, 2004:1). These tools enable knowledge sharing by humans in an organisation (Sharma & Djiauw, 2011:113). However, the effectiveness of BI tools depends on their purpose and use (Hedgebeth, 2007a:415). Effective BI tools should empower managers and professionals to spend only 10% of their time using them (Sharmila & Subramani, 2013:2).

Hardware selected for BI project depends mostly on how much data are to be processed and how scalable the platform is. The number of users that must access the DW, and the measure of real-time to business, help to select the hardware's processing power, random access memory (RAM), hard disk capacity, as well as the input and output (IO) capacity that will satisfy business requirements (Inmon, 2000). The recent trends of BI/DW tools have what (Kimball *et al.*, 2008c:207) refers to as "best-of-breeds"; for example, these ETL tools incorporated some dimensional intelligence. Microsoft Integration Services (MIS) is an example of a tool that has built-in slowly changing dimension components.

3.4.4 Human talent and education

The nature of BI/DW requires human talent that can build and define solutions. The BI tools can be off-the-shelf products that are customised by organisations; however a BI solution is not an off-the-shelf product. A BI/DW solution is developed based on

organisational data and business requirements (Seufert & Schiefer, 2005:919). For example, a DW is one of the key components of a BI solution that is developed based on existing organisational data, business requirements, and processes.

Business intelligence is a discipline with diverse professional roles that requires the necessary set of skills to implement (Marjanovic, 2011:151). This discipline requires a certain set of skills in order to work like a well-oiled machine (Chen *et al.*, 2012:1182). Professionals in BI/DW acquire education and skills from different educational institutions, private colleges, and some in-house organisational training. Some of the roles that fall within the scope of BI include database administrators, system analysts, business analysts, standard query language (SQL) developers, DW architectures, and report developers.

Human talent is one of the components that is crucial for BI implementation (Kapoor & Sherif, 2012:1626). The evolvement of information communication and technology (ICT) brings with it different opportunities and challenges. One of the challenges driven by ICT is for higher education institutions to design and package innovative learning activities around BI/DW education programmes (Marjanovic, 2011:151). The personnel required in the industry necessitates both industry and higher education institutions to prepare the skills for future BI/DW professionals (Chen *et al.*, 2012:1165).

It is important that graduates of higher education institutions receive proper academic preparation before they enter workplaces as BI/DW professionals. Competition among ICT vendors, economic competition, and rapid technological advancements are some of the factors that force industry to demand graduates of higher quality and that are ready to add value in the workplace (Ramakrishnan & Yasin, 2011:1445).

3.5 Data warehousing

A DW depends upon a variety of already existing databases; therefore, a DW discussion begins with on-line transaction processing (OLTP) systems (Nilakanta *et al.*, 2008:263, Rifaie *et al.*, 2008:58). The rapid growth of organisational data

generated by OLTP systems created the dilemma of how to analyse large volumes of historical data (Chen *et al.*, 2012:1165). An OLTP system is “a system that captures, enters, stores, retrieves and process relevant details of business events and generates the information and documents necessary for running the business” (Rahmatian, 2003:479). Traditionally, OLTP systems were built to fulfil specific requirements for specific customers, i.e. sales, orders, call centres etc. Typically, OLTP systems process data at the individual transactional level (lowest granularity) for operational purposes, whilst decision makers need summarised data (Datta & Thomas, 1999:290). The original intent of OLTP was computerisation (Devlin & Murphy, 1988:60).

Typical OLTP design techniques attempt to eliminate data redundancies by breaking data into multiple smaller tables that are related (Ballard *et al.*, 1998:33). These techniques serve two major purposes: transactional speed and transactional simplicity (Kimball *et al.*, 2008b:205). OLTP systems’ speed requirement led to older data being purged from these systems to open up space for newer/more recent data (Rodgers, 1996). For example, systems owners may decide to keep data only for the most recent two quarters, and move all other data into operational data stores. Querying historical data that is older than the most recent two quarters will then produce no results as the older data are stored in operational data stores. This shortcoming created the need for DWs that can retain historical data for analytical purposes.

Organisational leaders depend on data to make good decisions; their dependence on data cultivated and broadened the realm of DWs, which dates back to the late 1980s (Hua *et al.*, 2012:53, Maqboolrao *et al.*, 2013:157). An article published in 1988 in The International Business Machines (IBM) systems journal gives evidence to the fact that OLTP was not entirely sufficient to cater for storage and retrieval of large-scale volumes of data (Devlin & Murphy, 1988:65). A key idea within data warehousing is to integrate data from multiple sources, i.e. spreadsheets, DB2 databases, IDMS records, and VSAM files. So, “[i]t is now apparent that architecture is needed to draw together the various strands of informational system activity within the company” (Devlin & Murphy, 1988:65).

The concepts and principles around data warehousing were first discussed in the 1970s by Inmon; he is considered the father of the DW (Inmon, 2007, Williams, 2012). Inmon (2002:389) defines a DW as “a collection of integrated, subject-oriented databases designed to support the decisions support systems function, where each unit of data is relevant to some moment in time”. So, according to Inmon (2002:389), a DW is subject-oriented, integrated, time-variant, and non-volatile. A DW is traditionally based on the representation of business areas as facts with dimensions around the facts (Schneider, 2008:309). For example, business areas (i.e. sales, orders, distribution, etc.) are subject areas of focus in a DW. Data are extracted, transformed, integrated, and loaded into these subject areas from source systems. Once data is stored in a DW it is in a non-volatile state; it does not change (Inmon, 2002:389). Inmon (2002:389) says that data must always be associated with a specific data/time in a DW. Trend analyses, projections, patterns, and many other forms of data reporting are possible if pieces of data are associated with a specific date/time (Rahmatian, 2003:480). As an example, the total calculation of sales every month allows for trends to be analysed in order to investigate sales performance; in this case month is the time factor associated with the sales data.

In the first edition of the data warehouse toolkit book, Kimball *et al.*, (1998, 19) define a DW as “a union of all the constituent data marts” from multiple sources. However, since then Kimball distanced himself from the term “data mart”; he now defines a warehouse as “the query-able data in your DW/BI system” (Kimball *et al.*, 2008d:523). A DW in itself it is not BI; rather, it is a major component of BI (Hua *et al.*, 2012:53, Sahay & Ranjan, 2008:31). A DW is a storage base for many BI platforms (Bucher *et al.*, 2009:411, Chen *et al.*, 2012:1182). Decision makers use BI tools and processes to leverage from this central repository. BI tools are implemented to utilise and report on data stored in a DW.

3.6 The BI/DW life cycle

Various activities comprising of processes and involving stakeholders are required to deliver BI solutions. Kimball *et al.* (2008a:3) developed a life cycle methodology, which can be followed to tie together these activities in order to successfully deliver BI/DW projects. Although this methodology is named after Ralph Kimball; he still

maintains that many other experts have contributed towards the development of the methodology (Kimball *et al.*, 2008a:2). This framework comprises of thirteen control flow activates with dependencies that can be executed in a cyclic form. This methodology is used to guide the BI/DW discussion in the next sections. The aim is not to discuss these activities in a sequence, as presented in Figure 3-4, but to cover all the activities in this chapter.

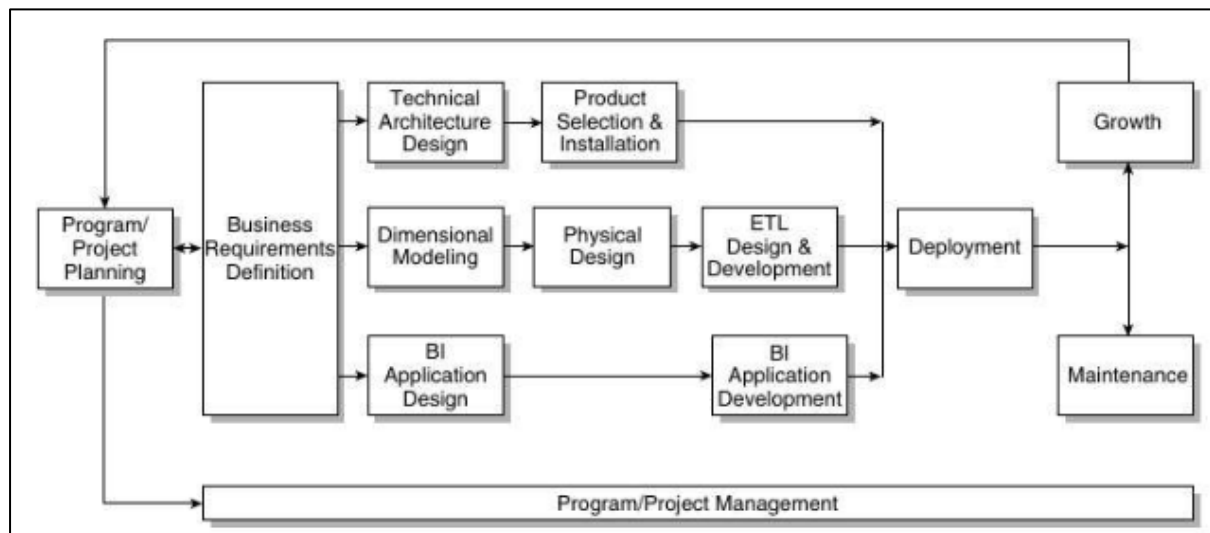


Figure 3-4: Kimball DW/BI life cycle diagram (Kimball *et al.*, 2008a:3)

3.6.1 Project planning and management

The success of a BI/DW solution depends heavily on the appropriate planning, integration, and management of the other twelve activities depicted in Figure 3-4 above. Planning involves: determining the organisational readiness, sponsorship, and other stakeholders' levels of commitment; a compelling business motivation and feasibility of a solution; financial implications; resourcing; and time management (Kimball *et al.*, 2008a:17). The project manager must ensure that all activities are on track and synchronised (Kimball *et al.*, 2008a:4). Continuous awareness of project status, resource allocation, tracking, and change control management are the main activities in project management.

3.6.2 Requirements collection

Business intelligence solutions are not ready-made to be picked up from shelves in shops; instead they are mostly designed based on specific business requirements.

Business data and user requirements' differ from one organisation to the other, and this makes requirements collection a very important activity. Kimball *et al.* (2008a: emphasise strongly that business requirements impact on all aspects of the project (i.e. programme/project planning and management, technical architecture, dimensional modelling, physical design, ETL design/development, BI application design/development, deployment, and maintenance). Requirements gathering methods (e.g. interviews and facilitated sessions) are used to collect requirements.

3.6.3 BI/DW architecture

Data warehousing is not about tools; rather, it is about a strategy to plan, design, and construct a data store to enable answering business questions (Nelson, 2007:2). A DW must be designed in such a way that it eliminates many of the inconsistencies inherited from its various source systems (Rifaie *et al.*, 2008:58). This integration requires careful thought and planning in terms of the architectural design (Levene & Loizou, 2003:225). The construction and integration of organisational data must follow the architecture that is best suited for the business needs and environment (Seufert & Schiefer, 2005:919, Oracle, 2013). According to Kimball *et al.* (2008d:100) architecture answers the question related to “how will we do it?”. Architecture provides an answer by identifying parts of components, their purpose, and the relationships among them. The BI/DW architecture addresses aspects such as data allocation over different structures, how to handle users based on their priority, and how to manage performance and speed (Maqboolrao *et al.*, 2013:157). Reporting, data management, security, band width, and storage requirements are all included in the BI/DW architecture (Inmon, 2000). Kimball *et al.* (2008a:211) emphasise that the security manager should be involved in the BI/DW architecture design process to ensure information security around the DW.

3.6.4 Data warehouse architecture types

Kimball and Inmon are known to be the two catalysts of DW philosophies that are influencing the DW world (George, 2012, Anupindi & Coady, 2005). Respectively, they advocate for dependent DW with linked dimensions and centralised DW architectures. However, there are also three other architectures models that have fundamental differences and strong advocates – these are discussed next.

3.6.4.1 Centralized data warehouse

Centralised data warehouse (CDW) advocates for a single database that serves the entire business in terms of, for example, department, branch, or division (Watson & Ariyachandra, 2005:4). This architecture serves the needs of more than one business unit. Data extracted from a database that was developed using CDW can be used across the business units because it represents a single version of truth (Hajmoosaei *et al.*, 2011:293). This architecture is founded on extensive enterprise-level analyses of data requirements (Ariyachandra & Watson, 2010:200). Figure 3-5 below illustrates CDW.

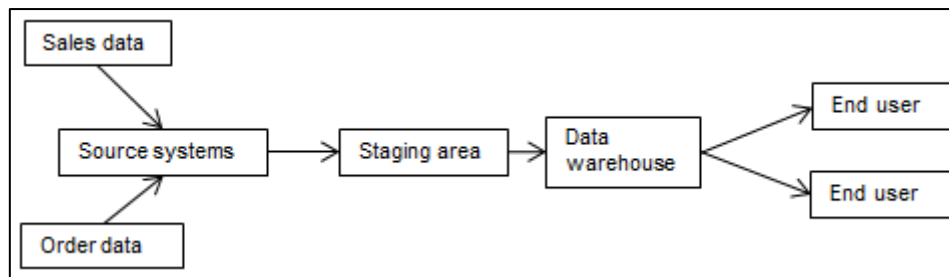


Figure 3-5: Centralised DW architecture

Inmon (2002:389) coined the term “corporate information factory” to refer to a centralised DW with dependent data marts. He champions the large centralised DW approach leveraging on solid relational design principles. He also advocates for a top-down development philosophy of designing a DW (Inmon, 2002:389). For the best decisions to be made, all the relevant data has to be considered. That necessitates the presence of all available data in a DW.

3.6.4.2 Independent data mart

It is common for an organisation to have business units with different functions and, consequently, legacy independent data marts (IDM) (Ariyachandra & Watson, 2010:201). This architecture allows business units to build their own data marts that are specific to their units; these do not have conformed dimensions. Independent marts cannot be used across the business because they don’t represent a single version of truth. Each data mart is built and loaded individually; there is no common or shared metadata. Kimball *et al.* (2008d:25) highlight the problems associated with this approach and refer to this as a stovepipe solution. The measures and

dimensions used in this architecture are also inconsistent across the business. Figure 3-6 below is an illustration of IDM.

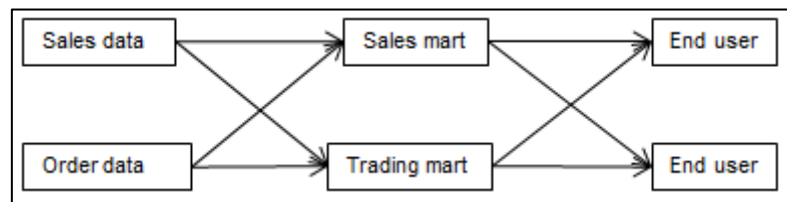


Figure 3-6: Independent data mart architecture

3.6.4.3 Dependent data warehouse

A dependent data warehouse (DDW), which is linked with dimensions' data marts, is advocated by Kimball (Ariyachandra & Watson, 2010:200). Kimball's philosophy is a bottom-up data warehousing methodology where individual data marts provide small/specialised views of organisational data (Kimball & Margy, 2002:14). Kimball's original focus was on data marts and conformed dimensions. However, in his later writing, Kimball *et al.* (2008d:15) say the term 'mart' has been relegated to mean summarised departmental and independent data sets. Common requirements from different business units is the major reason why Kimball distances himself from the term mart (Kimball *et al.*, 2008d:118). Kimball now uses 'business process dimensional models' to endorse the scope and purpose of data. Conformed dimension is a term Kimball uses to describe data elements that have similar meaning across all business units. For example, a customer key can be used to locate the same customer across sales and order systems. Figure 3-7 is an illustration of a DDW architecture model with sales and orders as business processes inputs.

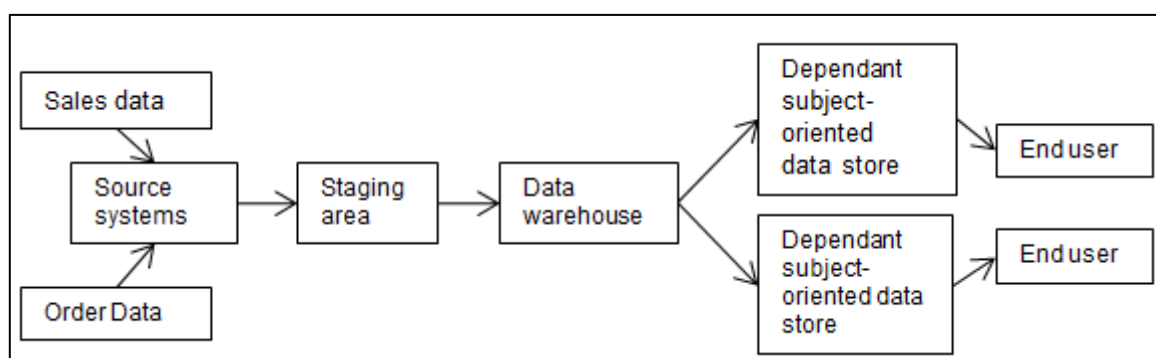


Figure 3-7: Dependent DW architecture

3.6.4.4 Homogeneous data warehouse

A homogeneous data warehouse (HDW) is a type of distributed DW architecture where data marts are physically distributed across the nodes of a network (Watson & Ariyachandra, 2005). A key factor of a distributed DW is speed. Providing users with access to only what is relevant to them, optimises performance. There is local dependence (interdependence) in a homogeneous DW because data from multiple sources interact. “Data found in the data warehouse is homogeneous in the sense that all of the types of records are the same” (Inmon, 2005:62). Data is collected from various sources and organised into a DW for easy access and understandability, thus giving the homogeneous nature. Figure 3-8 is an illustration of HDW.

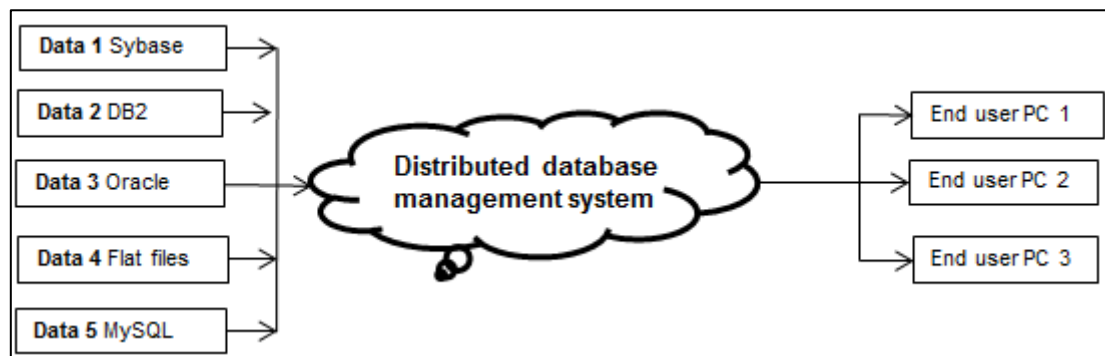


Figure 3-8: Homogeneous DW architecture

3.6.4.5 Heterogeneous data warehouse

Diverse sources may result in diverse data content and thus in a heterogeneous DW (HDW). “The hetero-homogeneous data warehouse is homogeneous with respect to a common base schema, whereas different sub-cubes and dimension branches may introduce additional measures and aggregation levels” (Schutz, 2010:12). This architecture uses different database management systems (DBMSs) at a local site. Independent data marts store interdependent data (Watson & Ariyachandra, 2005). Data from external sources can still be included in these different marts. Figure 3-9 illustrates the HDW architecture model.

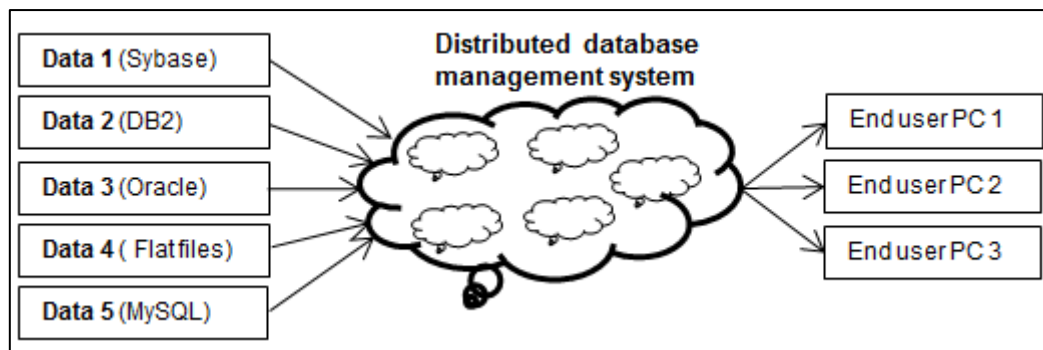


Figure 3-9: Heterogeneous DW architecture

Data warehouse architecture models are composed of components that are common to all. For the purpose of this study the Kimball architecture model is used from here onwards to guide the discussion.

High level BI/DW architecture components include: source systems, an ETL system, a presentation server, and the BI applications – these are represented in Figure 3-10 below; it depicts the high level DW/BI system architecture model. Each component is characterised by interrelated activities and processes. For example, the ETL system component includes extract, clean, and other process.

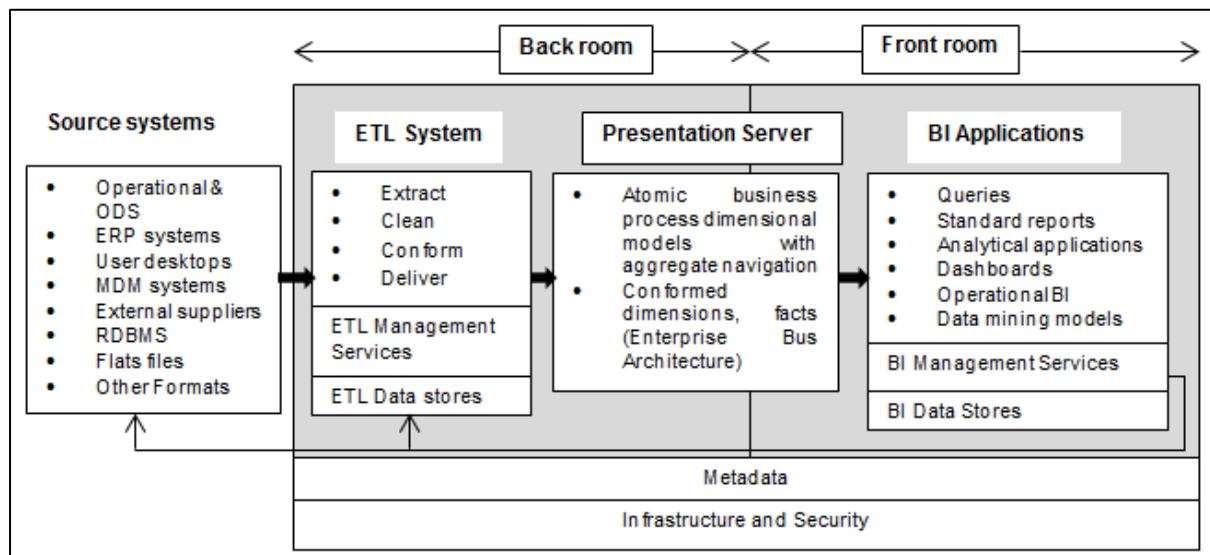


Figure 3-10: High level DW/BI system architecture model (Kimball *et al.*, 2008d:101)

The next sections focus on the illustration of DW concepts, and the four components on the DW/BI architecture model.

3.6.5 Illustrative example of data warehouse concepts

A telecommunication example will be used as an illustration going forward in this chapter to demonstrate and explain concepts and understanding of entity relationship diagrams and the star schema models. Telecommunications companies are currently operating in an extremely challenging business environment. The telecommunication industry owns large quantities of data, generated from numerous operational systems, and is confronted with many business problems that require urgent handling. The most common challenge experienced is that decision makers do not have access to accurate, relevant, consistent, and timeous information to support decision making for their respective roles and areas. This negatively impacts the ability of telecommunication service providers to remain competitive within the industry.

The illustrations in subsequent sections focus on on-line transaction processing (OLTP) systems and DW data presentation models. The purpose is to demonstrate the translation of an ERD model to a dimension model, and the data transformation processes that are involved. Different requirements (refer to Table 3-1 below) between OLTP systems and OLAP systems dictate different data models for these systems (Datta & Thomas, 1999:291). The ERD is the base model for designing the DW because it provides the structure and data requirements (Ballard *et al.*, 1998:35). Dimension modelling is used to conceptualise the OLAP systems (Datta & Thomas, 1999:290). Other data models (ERDs and recently data vaults) are used for data warehousing in industry. A discussion of these models is outside the scope of this study.

Table 3-1: Comparison of OLTP vs. OLAP (*Chen et al.*, 2002)

	OLTP	OLAP
Usage	Application specific	Decision support
Workload	Predefined	Unforeseeable
Access	Read / Write	Read only
Query structure	Simple	Complex
Records per operation	Tens / Hundreds	Thousands / Millions

Number of users	Thousands / Millions	Tens / Hundreds

The main purpose of dimensional modelling is to logically organise data in a format suitable for data warehousing requirements. For example, Figure 3-11 below depicts how data is represented in the OLTP versus the DW respectively.

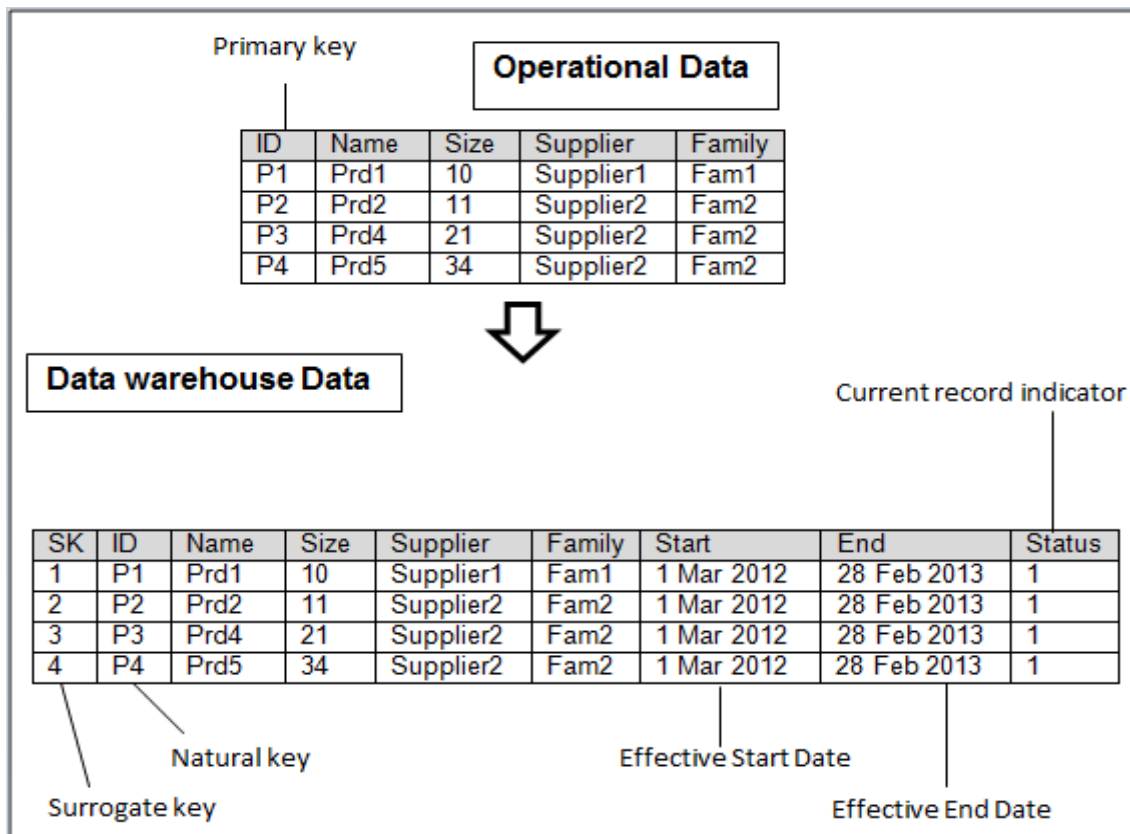


Figure 3-11: OLTP vs OLAP logical organisation of data

Dimensional models are populated with data that is extracted from OLTP source systems using the ETL system. Source systems are discussed next.

3.6.6 Source systems

Data comes from multiple systems that are built with multiple data stores and hosted on multiple platforms (Kimball *et al.*, 2008d:108). These systems are typically transaction processing applications which are a first point of capturing data (Oracle, 2002). For example, one of the source systems for a sales analysis data warehouse might be an order entry system that records all of the current order activities.

The OLTP system's data is modelled using the ERD technique, which breaks data into smaller atomic related tables (Datta & Thomas, 1999:290). Ballard *et al.* (1998:37) define a model as “an abstraction and reflection of the real world”. Figure 3-12 below shows an example of the ERD that a telecommunication company used to capture subscribers' transactional calls.

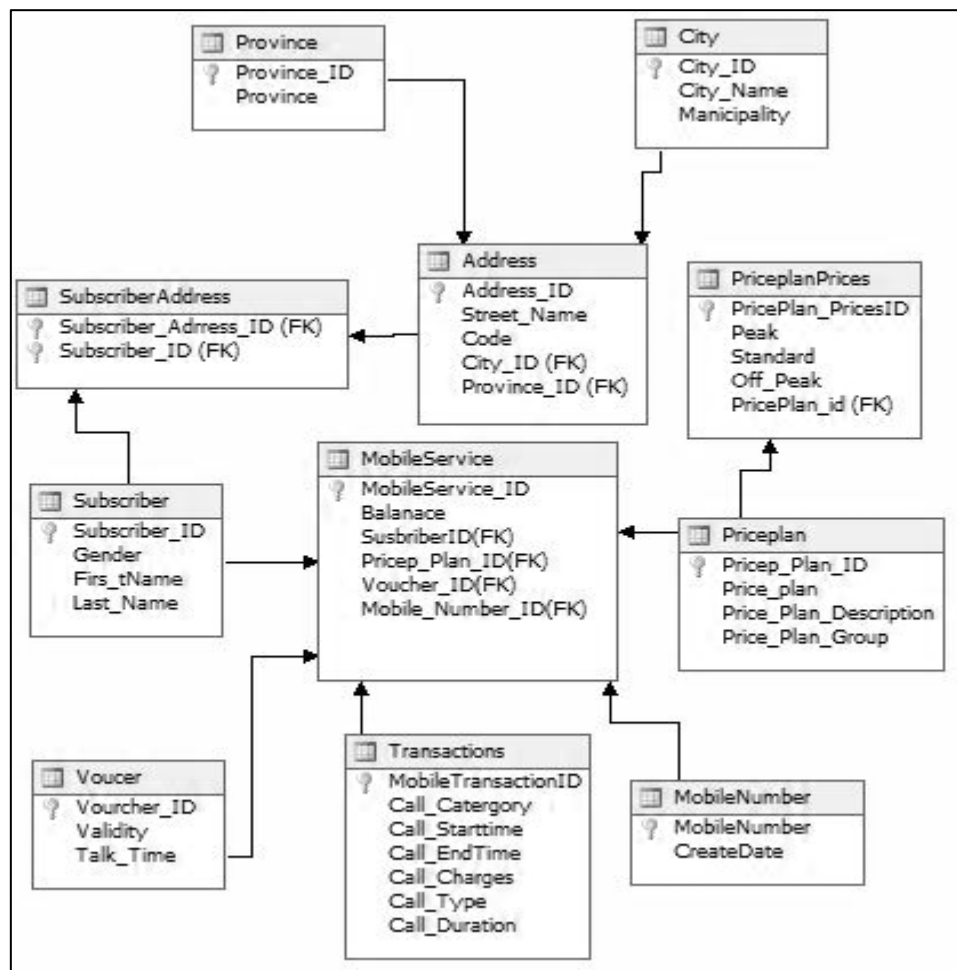


Figure 3-12: Illustrative Example: Entity relationship diagram

The ERD in Figure 3-12 is characterised by entities, relationships and constraints, attributes, as well as primary keys. An entity represents a real world phenomena, i.e. a person, place, thing, or event of value to the business or the organisation (Ballard *et al.*, 1998:37). Attributes define the characteristics of properties of the entities, i.e. gender, first name and last name define a subscriber in a SUBSCRIBER entity (Ballard *et al.*, 1998:37). Real-world entities are related; these relationships are shown on the ERD. A relationship type is a named association between entities. Figure 3-13 shows an example of a single relationship type 'Received' and its

inverse relationship type 'Provided to' between the two entity types SUBSCRIBER and MOBILE_SERVICE.

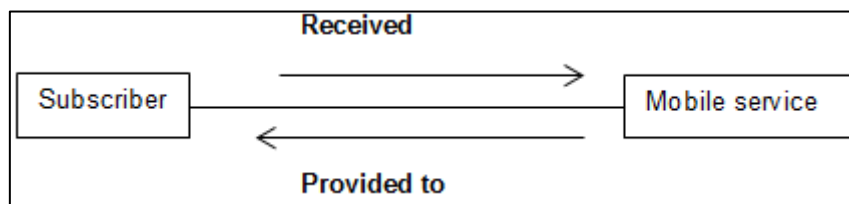


Figure 3-13: Representing a relationship on an entity-relationship diagram

The logical explanation of what is represented in Figure 3-13 is that a subscriber *received* mobile service and mobile service was *provided to* a subscriber. How many services mobile subscribers might have received (the maximum number and the minimum number) and how many mobile service might have been provided to, is shown by the *degree* of the relationship type. The 'degree' of relationship types is defined below.

In Figure 3-14 (below) three different ways of illustrating the existence of a relationship type are shown: Diagram (a) shows SUBSCRIBER and MOBILE_SERVICE entity types represented by index cards; it can be seen that there is a 'received' relationship type between subscriber number 2 and mobile service numbers 7 and 9. So, subscriber number 2 has 'received' these two invoices. These two mobile services have been 'provided_to' subscriber number 2. Diagram (b) shows the same information using a set notation with the relationship type 'received' and inverse relationship type 'provided to' linking subscriber entities and mobile service entities. Diagram (c) shows the ERD version; information about individual entities as well as which entity is linked to which cannot be displayed. The reason for this is simply that in a real database there would be hundreds of subscribers and mobile services entities; it would be impossible to show each one on the ERD.

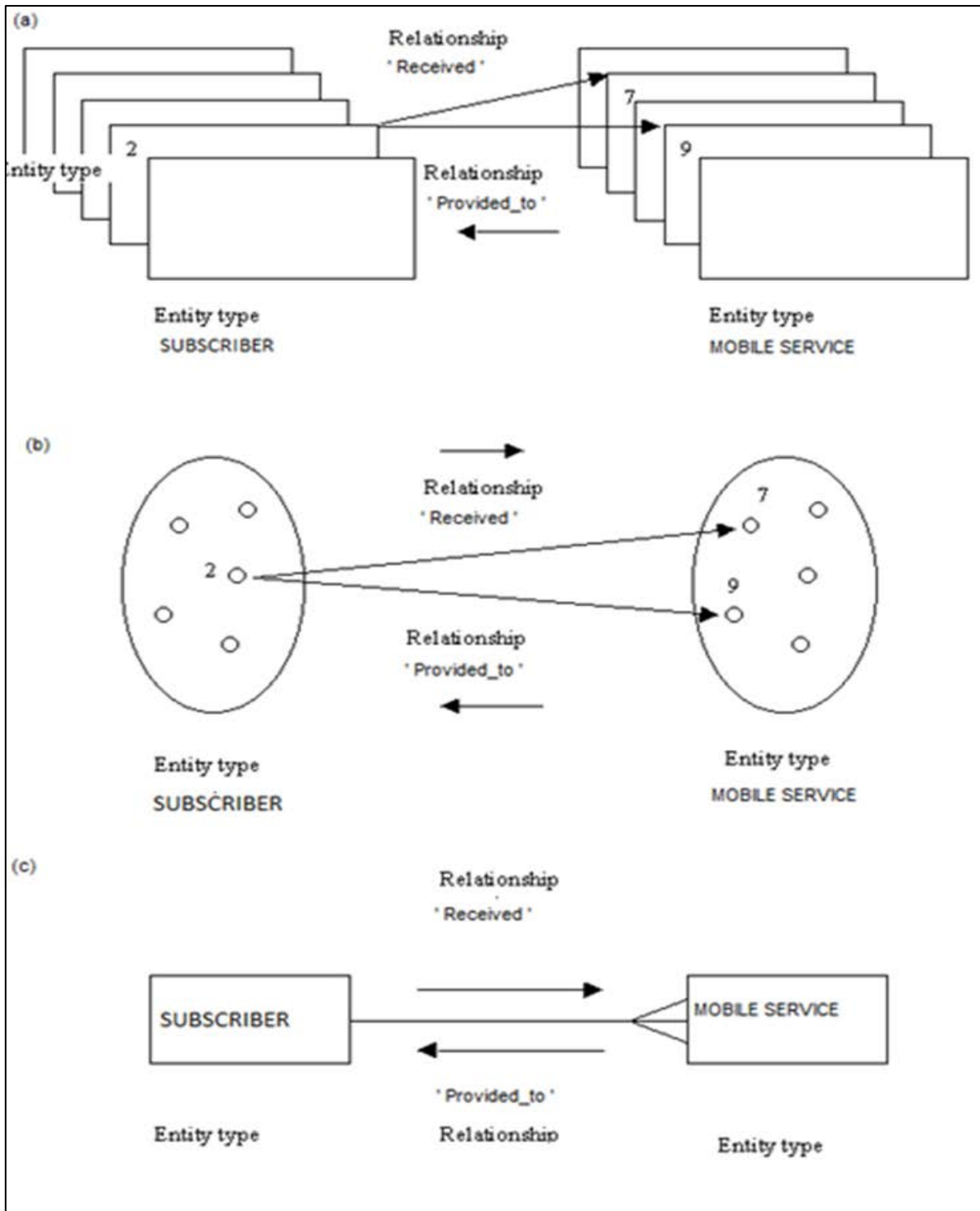


Figure 3-14: Three ways of thinking of a relationship

Figure 3-12 (earlier in the chapter) follows the above types of named associations among all eleven entities, i.e. the ADDRESS, PROVINCE, CITY, and SUBSCRIBER are examples of entities. A relationship is represented with lines between entities. Subscriber entity is related to SUBSCRIBER_ADDRESS entity. This relationship between the two entities is defined by SubscriberID. ETL is discussed next.

3.6.7 Extract, transform and load system

Data moves from a source system through the extraction transformation and load (ETL) process into the presentation servers (Kimball *et al.*, 2008d). This process does not have a straightforward approach because of its complexities, continuity, and hidden semantics (Simitsis & Vassiliadis, 2008:22). The complexities of the ETL process include selecting the right tools, access to source data, platform incompatibilities, scheduling, connectivity, transformation and load speed, etc. (Vassiliadis *et al.*, 2001:210). The ETL process is mostly characterised by the fragmented data flow using traditional ETL tools. Figure 3-15 illustrates a fragmented ETL data flow process. Consequently, integration of vast amounts of heterogeneous data from disparate sources requires a well-designed architecture solution (Hanlin *et al.*, 2012:120).

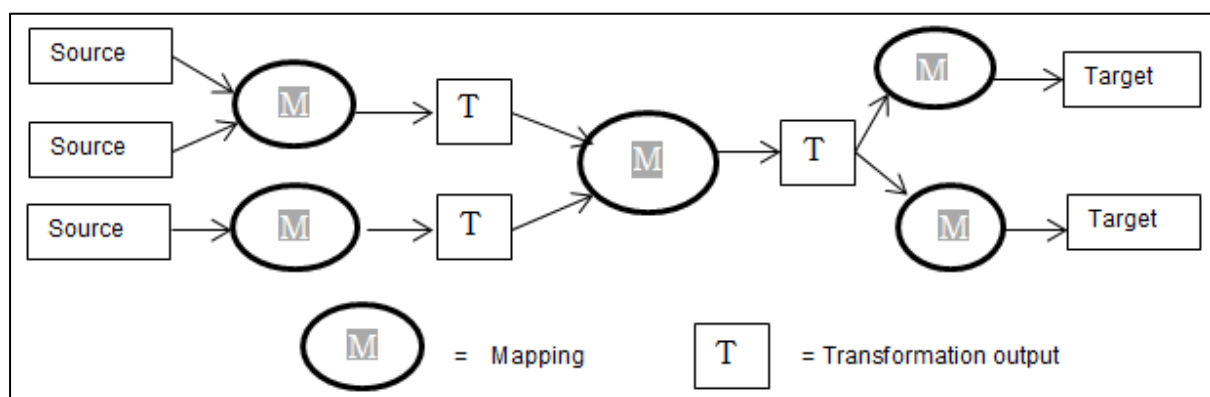


Figure 3-15: Fragmented ETL data-flow

The purpose of the ETL process is to transform data into the format needed for a DW through manual and automated extraction and transformation; it must also allow new content from new sources to be incorporated easily through fast and simple execution processes (Karagiannis *et al.*, 2013:718). The ETL process can take various activities, which are organised into workflows, to depict logical flow of data from source to DW (Karagiannis *et al.*, 2013:927). Consider the ERD model in Figure 3-12 (earlier in this chapter) as an example. Figure 3-16 below shows the ETL workflow on how data flows from the ERD into the star schema model.

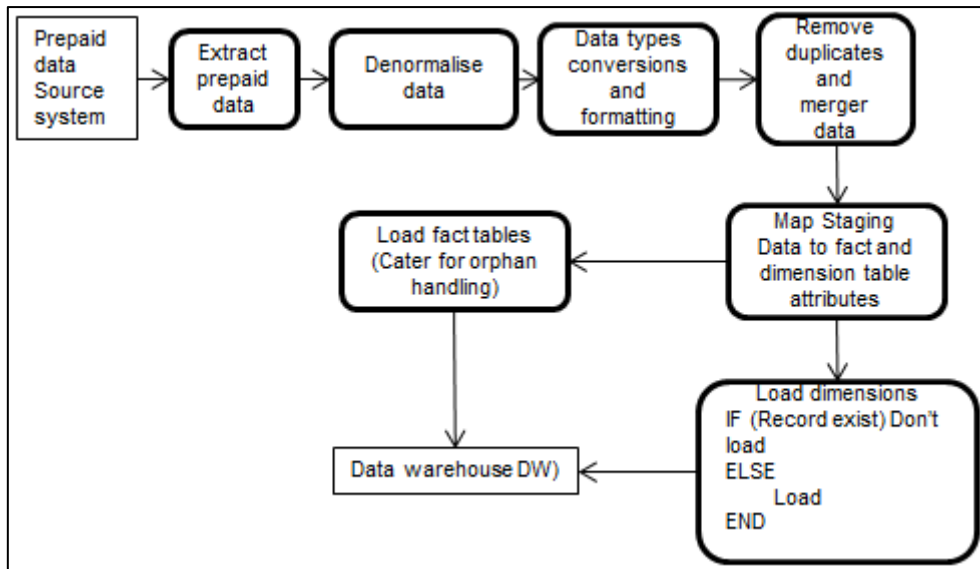


Figure 3-16: Illustrative example: ETL workflow

3.6.7.1 Source to target mapping

The need for data warehouses to report correct information is increasing (Zeng *et al.*, 2006:4722). There are processes such as extraction, staging, transformation, loading, and storing between the systems, where data are captured to be presented via the BI reporting tool. All these processes can compromise the accuracy of reporting data if not executed appropriately. Therefore, data mapping is used to describe on a high level how each target table and column in the DW will be populated (Kimball *et al.*, 2008d:282). Data mapping is fundamental for data cleaning, data integration, and correct reporting (Fletcher & Wyss, 2006:95).

Figure 3-4 (earlier in this chapter) serves as a DW/BI architecture and as source to target because it depicts an abstract high level source to target mapping view of where data is being sourced, how is it being transported until it is finally reported. The low level details of this mapping should include source names (servers), databases and schemas, staging databases and schema names, table's names and columns, and target reports.

Source to target mapping requires a full understanding of sources, destination area, and the mapping area because it is imperative that information should be traceable back to the source (El-Sappagh *et al.*, 2011:92, Kimball *et al.*, 2008d:26). Discovering the relationships between the sources and the targets depend entirely

on good business requirements (Perkins, 2003). Source to target mapping aids the design of transformation rules, i.e. source to target identifies the source system, data to be extracted, the output format, and the final destination system and format. Figure 3-17 below illustrates a high level source to target mapping: (a) depicts the sales data table that is in an Oracle source system; the sales data is extracted, transformed, and loaded by the SQL ETL system into the DW model residing in SQL presentation server.

Consider the telecommunication illustration where the requirement is to extract sales data from an Oracle system and enter the data into a Microsoft SQL based DW, in order to get the sales summaries per customer. Figure 3-17 (b) is a visual representation of the mapping in Microsoft SQL server. The ETL process includes extraction of the sales data from the Oracle source system into the SQL server destination (presentation server); the final output is aggregated sales totals per customer (of the SUBSCRIBER). The ETL system handles the transformation processes using different business rules to archive the required sales output.

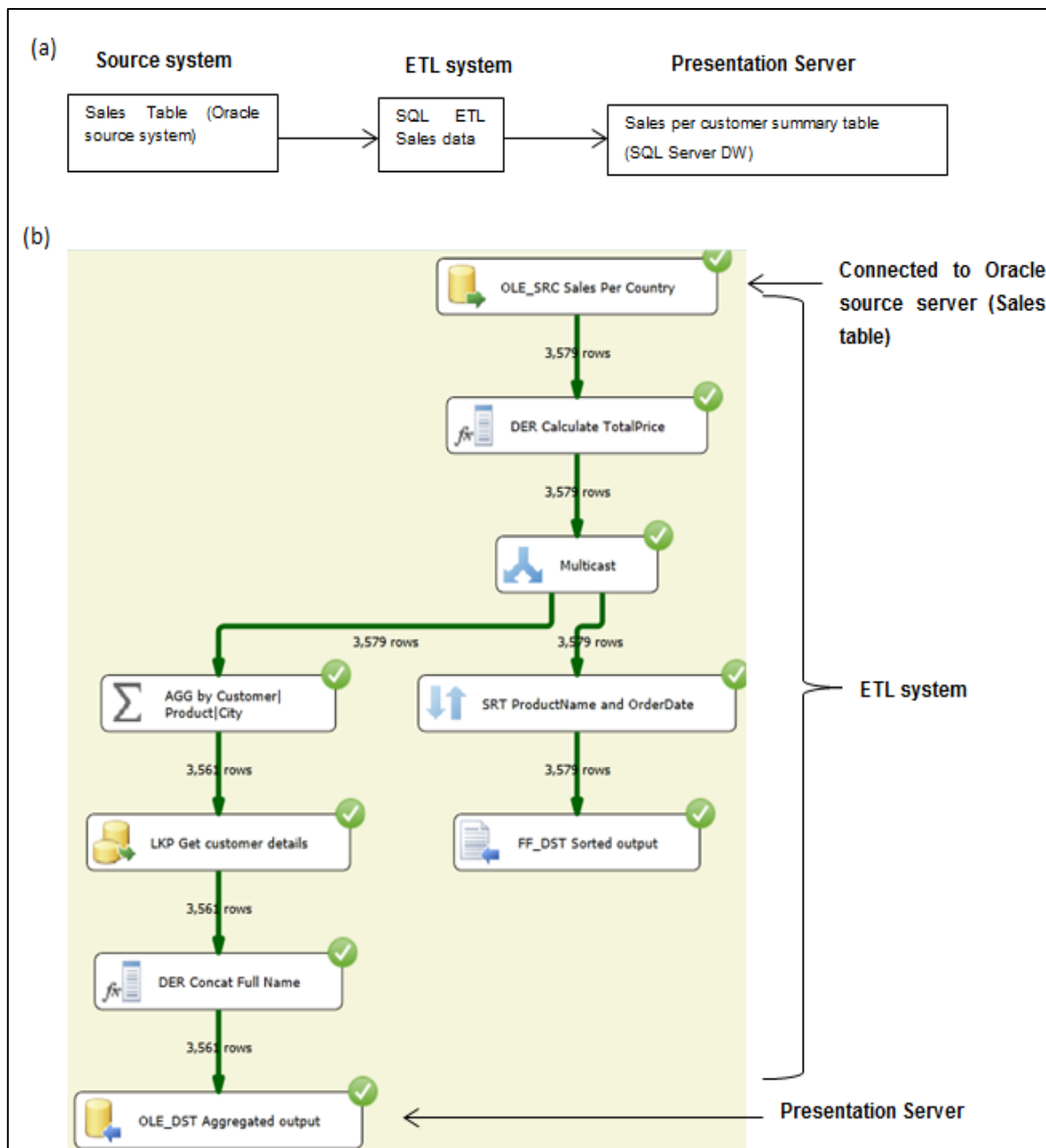


Figure 3-17: Illustrative example: visual representation of source to target mapping

3.6.7.2 Extraction, transform, and load architecture

Karagiannis *et al.* (2013:937) say that the ETL process is important for the accuracy, operability, and usability of the DW; hence, the ETL process forms the backbone of a DW. The ETL architecture entails four major operations: data extraction from the sources, running it through a set of cleansing and conforming transformation processes, storage in prepared format, and delivering it to the presentation server (Kimball *et al.*, 2008d:109). Figure 3-18 illustrates the ETL architecture, which is referred to as the backroom of the DW architecture.

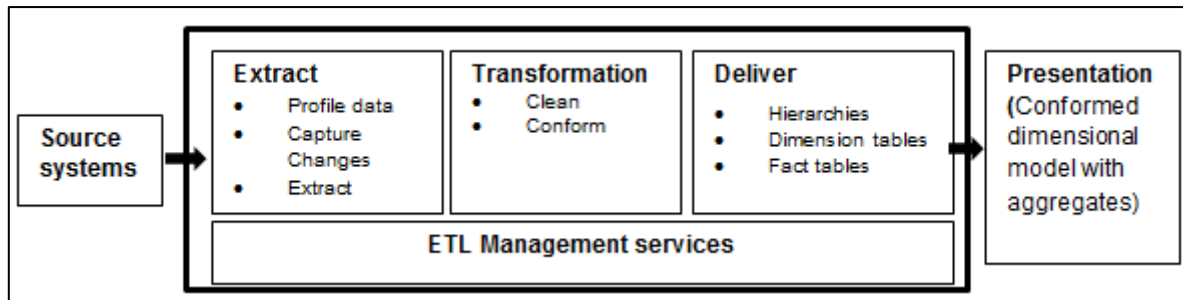


Figure 3-18: Illustrative example: ETL architecture

Data extraction, transformation, and loading are the three main processes of the ETL architecture that are discussed next.

3.6.7.3 Data extraction

Building the ETL process is considered the first comprehensive task in a DW development; it is thus imperative to have a solid understanding of the source, destination, and the mapping areas (El-Sappagh *et al.*, 2011:92, Breslin, 2004:15, Bergamaschi *et al.*, 2011:717). The purpose of the data extraction process is to identify and extract the correct data subsets that will be integrated through the ETL workflow and loaded into a DW (Thoo *et al.*, 2013:3). Different data sets are extracted from multiple sources using custom or off-the-shelf extraction tools (Matthias & Yannis, 1997). InfoSphere DataStage, Integration Services (part of SQL Server), Oracle Warehouse Builder (OWB), and Informatica are some of the leading ETL tools (Thoo *et al.*, 2013:3). These tools establish connectivity using direct connection or drivers (e.g. Open Database Connectivity (ODBC) and Java Database Connectivity (JDBC)) (El-Sappagh *et al.*, 2011: 93). Figure 3-19 illustrates connectivity between data sources and the ETL process.

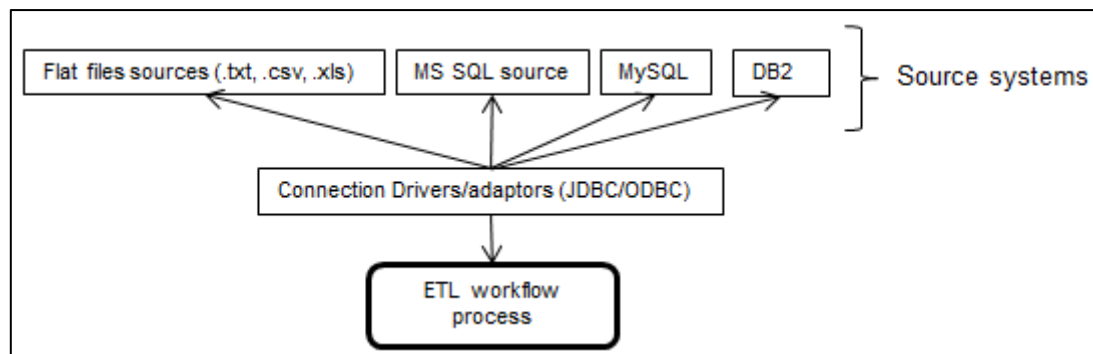


Figure 3-19: Illustrative example: Connectivity

Data warehouse development includes two extraction processes, i.e. full or initial extraction; and incremental data extraction (Breslin, 2004:15). Initial extraction is extracting the data for the first time from sources; whilst incremental extraction is done continuously to update the DW afterwards. The incremental data loading is continues and as a result of changes in source systems, updates to data, and changing user requirements (El-Sappagh *et al.*, 2011:92).

3.6.7.4 Data transformation

Recent database and DW developments are mostly based on existing data from other databases such as source systems (Parent & Spaccapietra, 2000:1). Data must be re-used and redefined for logical and structural unification to fit into a DW for storage and reporting purposes. This unification is referred to as data integration (Inmon, 2002:3). Data integration combines data residing at multiple sources and schemas into one uniform integrated global schema (Parent & Spaccapietra, 2000:1). All possible inconsistencies and redundancies should be resolved to provide a reconciled integrated data view (Calvanese *et al.*, 1999:1)

The transform step applies a set of rules to transform the data from the source to the target, i.e. transformation rules include resolving domain conflicts, dealing with missing elements, correcting wrong spelling, and implementing standard formats (i.e. extract only what is needed). This process takes place in the ETL system (Kimball *et al.*, 2008d:109). Data transformation processes aim to produce data that are correct, complete, accurate, consistent, and unambiguous (El-Sappagh *et al.*, 2011:92).

Multiple source systems may result in inconsistent data feeds; these must be cleaned, i.e. during data capture users make typing and spelling mistakes; then, data lack consistency and cleanliness. Some of the data inconsistencies are caused by corruption in transmission or storage, or by different data dictionary definitions of similar entities in different stores. These inconsistencies must be cleaned before data are loaded in a DW. Figure 3-20 below illustrates attribute, case, and spaces transformation.

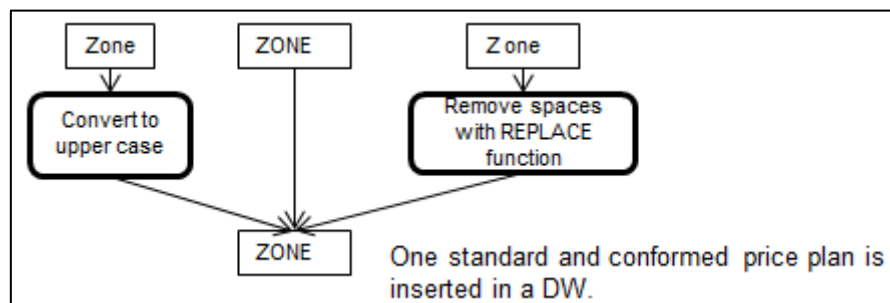


Figure 3-20: Illustrative example: Attribute case and space transformation

The ETL system is responsible for converting data into consistent and standardised formats. Part of the data structuring process is to standardise the format for storage and output of data. Figure 3-21 is an illustration of cell phone number formatting with 3-3-4 digits and spaces in between.

Nonstandard format	Standard format
0784565711	078 456 5711

Figure 3-21: Illustrative example: Cell phone number standardisation

Since data are sourced from heterogeneous systems in different formats, it must be converted into a format that is anticipated by the target DW/BI system. Data type conversions can be made using tools or scripting. Consider, for example, the case where the business users want to see the monetary value in a balance column, instead of a number; Figure 3-22 shows the Microsoft SQL script for the conversion of the number to a monetary value in.

```

SELECT CONVERT (money, Balance) as Balance_Amount
FROM Sales
  
```

Figure 3-22: Illustrative example: SQL scrip to convert number to monetary value

Another form of data transformation is data cleaning (scrubbing), detecting and correcting (or removing) corrupt or inaccurate records from a record set, table, or database. The cleaning step is one of the most important steps because it ensures high quality data in the DW. Data from source systems are rarely clean. Data must be cleaned and pruned to cut off unnecessary spaces and characters, e.g. a customer name ('Steven ') can have a space at the end. This space will cause user requests not to return information when they query the name 'Steven'. This problem can be resolved by removing spaces in a name. Figure 3-23 illustrates a MS SQL statement that removes unwanted spaces in front and at the end of customer name.

```
SELECT LTRIM ( RTRIM ('Steven ') ) as Name
FROM Customer
```

Figure 3-23: Illustrative example: MS SQL statement to remove unwanted spaces

When building dimensions and fact tables, data from multiple source systems must be merged and loaded into one conformed dimension or fact table. This process ensures that all records from source systems are represented in a DW. Figure 3-24 illustrates how data are merged from two source systems (post-paid and prepaid) into one table.

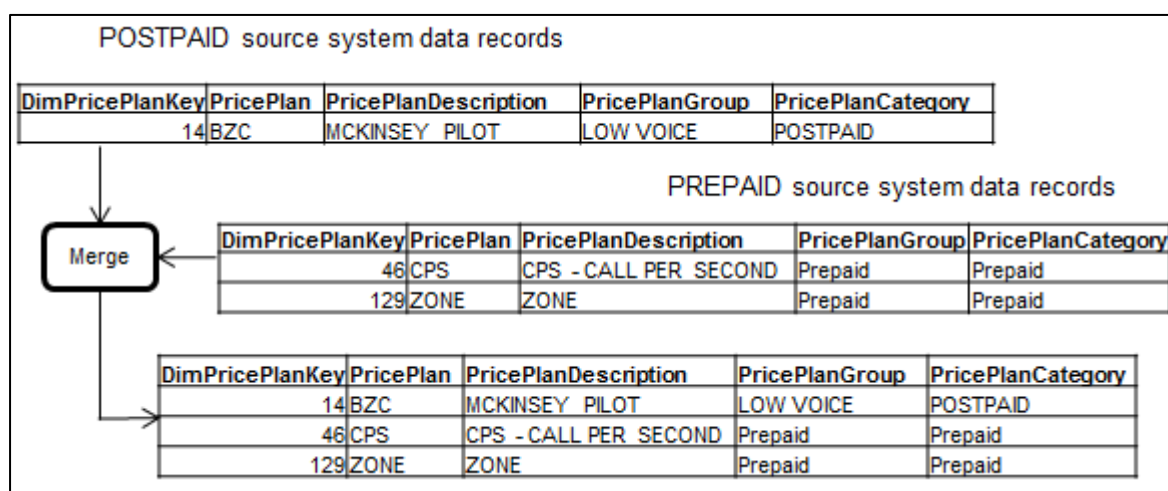


Figure 3-24: Illustrative example: Merge data records

3.6.7.5 Data loading

The final step in the ETL process is to load the data into the DW (Breslin, 2004:15). Dimensions and facts are populated as per business requirements, i.e. the slowly changing dimensions (SCD) rules. A DW has two types of loads: initial data loads; and incremental data loads (Breslin, 2004:16). An initial load is a first data load from sources; incremental loads are done continuously to update the DW. The general rule for DW data loads is that they occur in the background during idle times of source systems. The ETL loads are scheduled to load data automatically in the DW. The presentation server is discussed next.

3.6.8 Presentation server

The presentation server is the third level component of the BI/DW architecture model. Data are extracted from source systems and loaded into a dimensional model on the presentation server using the ETL. This process takes place before data is packaged into its final query able format (Kimball *et al.*, 2008d:113). Data modelling that involves dimensions and facts in the star schema is discussed next.

3.6.8.1 Illustrative example: Star schema

The purpose of this section is to give an illustrative example of a star schema model, which is used to guide the discussion in subsequent sections. Consider the ERD presented earlier in this chapter; the phone calls made by subscribers are actual transactions that can be viewed in many perspectives (dimensions), i.e. time of call, call category, call type, the mobile number that made the call, the voucher used in a transaction, the price plan, and the address of the subscriber. Figure 3-25 below is a graphical representation of transaction perspectives.

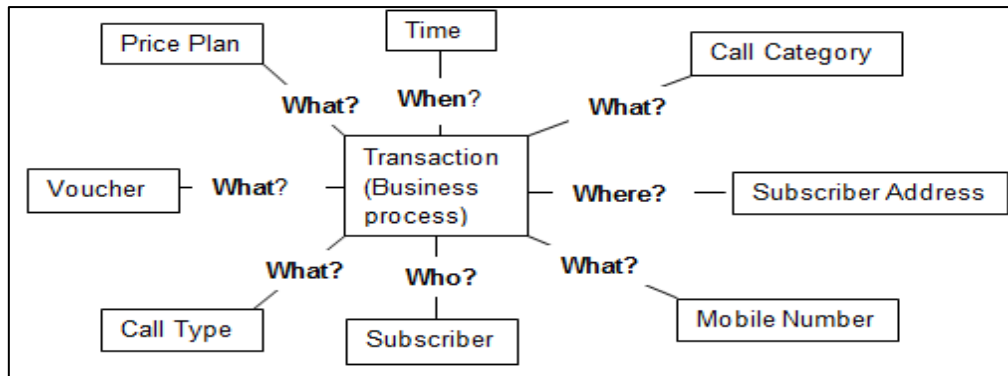


Figure 3-25: Illustrative example: Transaction perspectives (dimensions)

Table 3-2 contains a list of terms, descriptions, and short names used in the running example in this chapter.

Table 3-2: Terms and definitions

Term	Description
Price plan	Recharge method such pay as you go or top up.
Voucher	Airtime voucher used to recharge mobile number.
Call type	Indicate if the call is business or consumer type
Mobile number	The actual cell number such as 078 4565711.
Call category	Describes if the call is national or International.
Subscriber address	Residence of a subscriber.

Figure 3-26 below depicts the star schema. The formation of the star schema from the ERD is explained in the subsequent section.

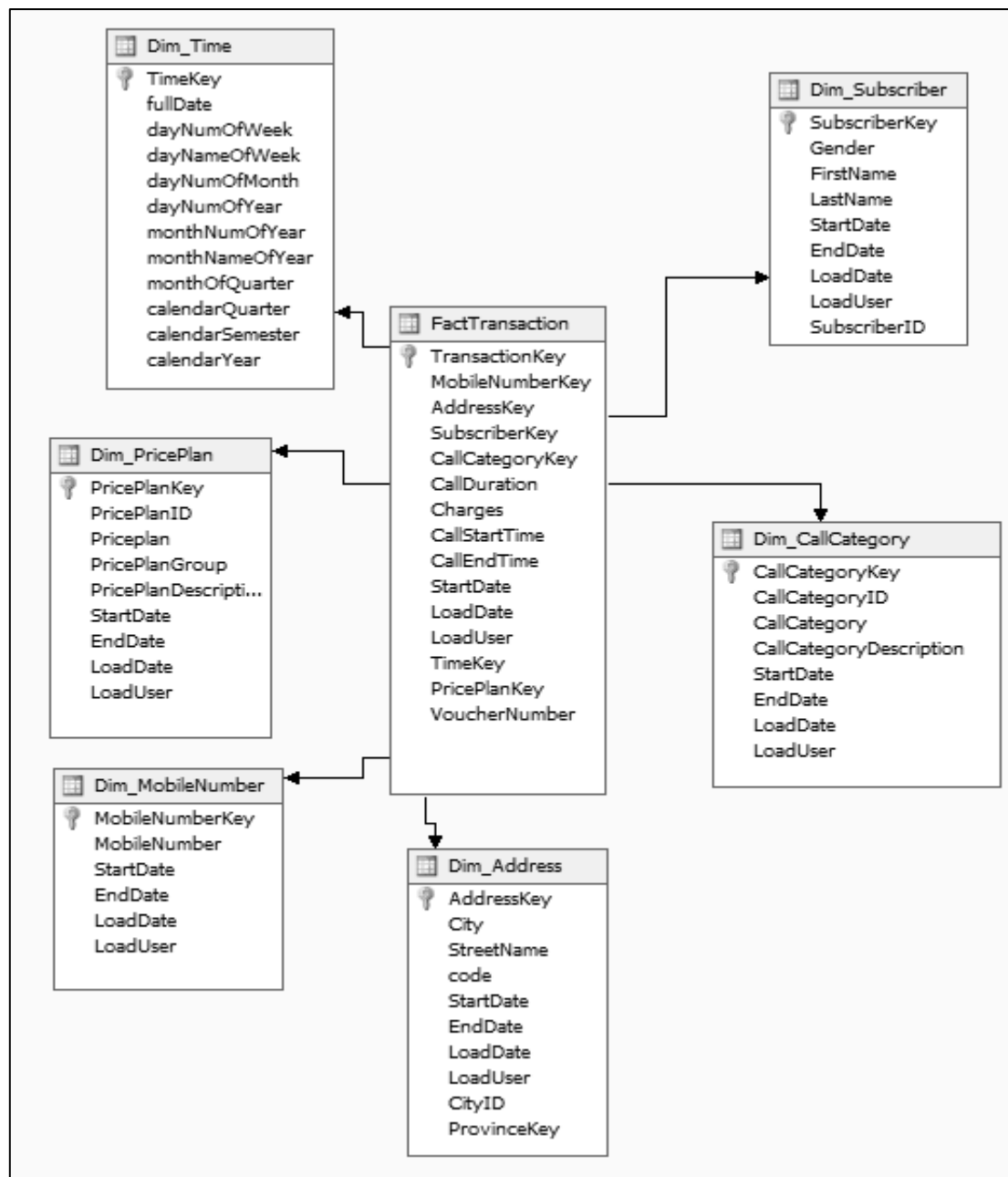


Figure 3-26: Illustrative example: Telecommunication transactions star schema

3.6.8.2 Dimensional modelling

Dimensional modelling is a technique that was invented by Kimball to conceptualise a business as a set of entities, i.e. dimensions and facts tables, and the relationships amongst them (Kimball, 2014, O'Leary, 1999:49). Even though Kimball is considered to be the inventor of dimensional modelling, he rejects taking credit for it (Kimball, 2014). According to Kimball *et al.* (2008d:204), dimensional modelling is an approach that was not invented by a single person; instead, it matured with time and

through designers' understanding. It has performance as the highest goal. Kimball argues that the bulk of the effort required to build a proper DW/BI system should be invested in modelling the system (Kimball *et al.*, 2008d:204).

Understandability and query performance are the main drivers of dimensional modelling (Kimball *et al.*, 2008d:206). Data must be presented in an understandable format. ERDs are complex to understand for data users (Firestone, 1998). Dimensional modelling simplifies data relationships by grouping data into coherent business categories (or dimensions) that make sense to business people (Chen *et al.*, 2002, Ballard *et al.*, 1998:42, Kimball *et al.*, 2008d:124). Less joins between data groupings improve query performance.

The dimensional design process involves identifying business processes, choosing the level of data detail in a business process (referred to as a grain), and identifying dimensions and facts that are applicable to a business process (Kimball *et al.*, 2008d:205). Dimensional modelling allows data to be structured into fact tables and dimension tables, to reflect the abstract real world referred to as a star schema (Ballard *et al.*, 1998:35, Soler *et al.*, 2008:341). A star schema is a type of a logical DW model; it represents the relationship between facts and dimensions by dividing data into measures (facts), and their "who, what, where, when, why, and how" descriptive context (dimensions) (Kimball, 2008). A star schema is made up of dimension and fact tables, which are linked with keys to maintain referential integrity.

The extended version of a star schema is a snowflake schema, where dimensions are in a normalised form (Datta & Thomas, 1999:293). A snowflake dimension is a normalised table; its aim is to reduce use of space, by avoiding repetition of values (Kimball *et al.*, 2008d:233). The need for snow flaking in a DW arises when a dimension attribute has a significantly different grain than the primary dimension and is administered on a different frequency. For example, DIM_ADDRESS and DIM_PROVINCE represent snow flaking in Figure 3-27. This normalisation reduces repetition of province names in the address table.

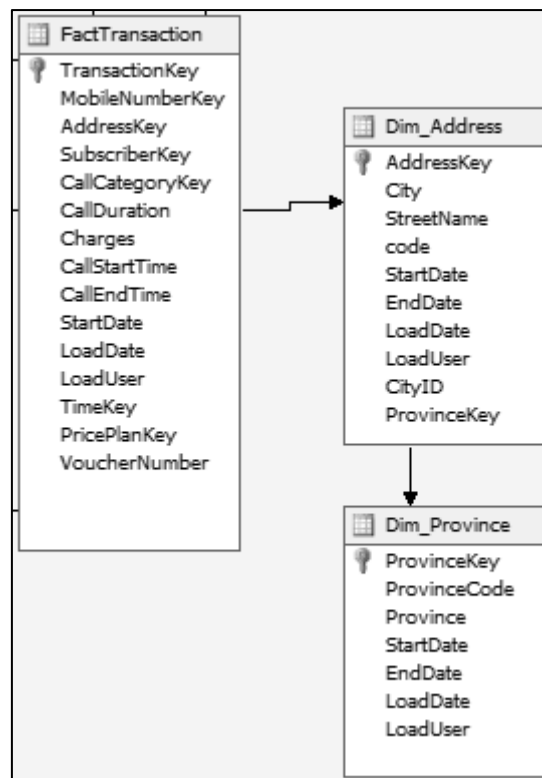


Figure 3-27: Illustrative example: snow flake schema

3.6.8.3 Dimensions

Numbers mean nothing to business users if they are not associated with certain attributes that attaches meaning to them (Ralph, 2008). Dimensions are tables that store textual descriptions of the business and thus they add meaning to numeric numbers by describing the "who, what, when, where, why, and how" (Kimball *et al.*, 2008d:204). Dimensions are represented by axis and should always be built around business processes which constitute facts (Ross, 2005, Ballard *et al.*, 1998). For example, Figure 3-26 (earlier in this chapter) depicts dimensions that are built around a business transaction.

A dimension is considered a static entity in a multidimensional model (Kimball & Margy, 2002:95). This implies that data are anticipated to grow faster in a fact table, whilst the dimensions remain static or grow at a very slow pace. Dimensions are very small in terms of the number of records stored in them when compared to fact tables (Kimball *et al.*, 2008d:337). A dimension stores a single unique record that is associated with more than one record in a fact table. Figure 3-28 illustrates the abbreviated FactTransction and dim_PricePlan (only fields required for the

illustration are included) where PricePlanKey 1 in Dim_PricePlan is associated with 2 records (transaction key 1 and 2) in a fact FactTransaction.

FactTransaction													
Transaction Key	Time Key	MobileNumber Key	Address Key	Subscriber Key	CallCategory Key	PricePlan Key	CallDuration	Voucher Number	Charges	CallStartTime	CallEndTime	StartDate	LoadDate
1	1462	1	1	1	3	1	45.12	1	28.51	7-May-14	7-May-14	7-Jan-14	7-May-14
2	1462	2	2	2	2	1	43.42	2	18.76	7-May-14	7-May-14	7-Mar-14	7-May-14
3	1462	3	3	3	1	2	15.39	3	10.02	7-May-14	7-May-14	2-May-14	7-May-14
4	1462	4	4	6	2	2	28.56	4	47.95	7-May-14	7-May-14	7-May-14	7-May-14
5	1462	5	5	5	1	3	16.21	5	28.78	7-May-14	7-May-14	7-May-14	7-May-14

Dim_PricePlan								
PricePlanKey	PricePlanID	Priceplan	PricePlanGroup	PricePlanDescription	StartDate	EndDate	LoadDate	LoadUser
1	1	Anytime	Prepaid		5/6/2014	1-Jan-00	6-May-14	Admin
2	2	Off Peak	Prepaid		5/6/2014	1-Jan-00	6-May-14	Admin
3	3	TopUp Option	Prepaid		5/6/2014	1-Jan-00	6-May-14	Admin

Figure 3-28: Illustrative example: FactTransaction and Dim_PricePlan Relationship

Hierarchies can be built based on dimensions to allow the users to navigate and zoom in at different levels of granularity. In this case dimension attributes are not only used for labelling, but they are also used for filtering data in a DW (Kimball *et al.*, 2008d:210). For example, in Figure 3-29 the time dimension (Dim_Time – the ‘dim’ prefix indicates a dimension table) can have year, quarter, month, and day hierarchies to allow users to get aggregations associated with a specific time position.

Time Key	fullDate	dayNumOf Week	dayNameOf Week	dayNumOf Month	dayNumOf Year	monthNum OfYear	monthName OfYear	monthOf Quarter	calendar Quarter	calendar Semester	calendar Year
1	1/1/2010	6	Friday	1	1	1	January	1	1	1	2010
2	1/2/2010	7	Saturday	2	2	1	January	1	1	1	2010
3	1/3/2010	1	Sunday	3	3	1	January	1	1	1	2010

Figure 3-29: Illustrative example: Dim_Time

Figure 3-30 below shows a time hierarchy that is built based on the DIM_TIME structure shown in Figure 3-29 above (note that only fields required for the illustration are included). The period hierarchy is used to navigate through the 2014 year by quarter, month, and day of month.

Calendar Year ▼	Calendar Quarter	Month Num Of Year	Full Date	Charges
2014	1	January		4442.66
		February		4304.83
		March		4522.25
		Total		13269.74
	2			14338.29
	3			13297.61
	4			14148.18
	Total			55053.82
	Grand Total			55053.82

Figure 3-30: Illustrative example: Time hierarchy

According to Inmon (2002:389) time is one of the basic characteristics of a DW. When associated with data, time separates the past from the future. The time attribute enables a conclusion that the state of data is true at a point in time (Nguyen *et al.*, 2007:28). Continuously changing business requirements (associated with time) necessitates structural (e.g. adding new column) and or data changes (corrections of mistakes in the data) in a DW (Nguyen *et al.*, 2007:29). Dimensions require a special approach because of their slow evolving nature; these changes must thus be handled and controlled carefully to preserve the integrity of the entire DW. Kimball introduced slowly changing dimensions (SCD) in 1996 as an approach that he recommends can handle operational attribute value changes in a dimension (Ross, 2013a).

According to the Kimball group there are eight types (types zero to seven) of SCD (Ross, 2005). The primary purposes of a DW are to keep historical data for analysis and reporting, as well as to enable future projections. Accordingly, the first SCD type is SCD type zero, where data do not change once it is stored in a DW (Ross, 2005). Data must be preserved and not changed once it moves from operational systems into a DW (Inmon, 2002:389). Inmon (2002:389) refers to this as a non-volatile state.

Types one to seven allows for changes to dimensional data. This is needed where there are business requirements that dictate that stored data be overwritten with new data. Certain types of changes defeat the purpose of a DW of keeping history (Ross, 2005). Type one requires that a record is updated and hence history is lost. Because history is lost, a fact is only associated with the current value that is in the dimension column, via a surrogate key (Nguyen *et al.*, 2007:29). The decision to overwrite data

must come from the business if it is to be implemented (Ross, 2005). Figure 3-31 shows that the surname of a subscriber (ID 3) changed from Pillar to Sanders. The old surname will be overwritten; history will not be kept in the DIM_SUBSCRIBER table.

Dim_Subscriber								
Subscriber Key	SubscriberID	Gender	FirstName	LastName	StartDate	EndDate	LoadDate	LoadUser
1	58256445	Male	Steve	Richards	5/6/2014		6-May-14	Admin
2	58653566	Male	Thys	Davids	5/6/2014		6-May-14	Admin
3	965454	Femaile	Mary	Pillar	5/6/2014		6-May-14	Admin
5	9659554	Female	Nosipho	Ndulula	5/7/2014		7-May-14	Admin
6	822564	Male	Jomo	Sono	5/7/2014		7-May-14	Admin
7	856523	Female	Kali	Nko	5/7/2014		7-May-14	Admin
New record from source with new LastName								
Subscriber Key	SubscriberID	Gender	FirstName	LastName	StartDate	EndDate	LoadDate	LoadUser
3	965454	Femaile	Mary	Sanders	5/6/2014		6-May-14	Admin
Dim_Subscriber with new LastName and startDate (Old surname is overwritten)								
Subscriber Key	SubscriberID	Gender	FirstName	LastName	StartDate	EndDate	LoadDate	LoadUser
1	58256445	Male	Steve	Richards	5/6/2014		6-May-14	Admin
2	58653566	Male	Thys	Davids	5/6/2014		6-May-14	Admin
3	965454	Femaile	Mary	Sanders	5/20/2014		20-May-14	Admin
5	9659554	Female	Nosipho	Ndulula	5/7/2014		7-May-14	Admin

Figure 3-31: Illustrative example: SCD type one

Type two proposes that keeping the original version of the data, as well as the new one, is important (Ross, 2005). A fact table is associated with the old and current record in a dimensional column via a surrogate key (Nguyen *et al.*, 2007:29). Figure 3-32 illustrates that the 'Anytime' PricePlan has a new PricePlanGroup from 'Prepaid' to 'Top-Up'; a new record was created with a new surrogate key 4 and new PricePlanGroup (price plan in this case is a source key that is used during loading to ensure that no duplicates are loaded). For integrity and accuracy the maximum loadDate column is used to retrieve the latest updated record from this dimension.

Dim_PricePlan								
PricePlanKey	PricePlanID	Priceplan	PricePlanGroup	PricePlanDescription	StartDate	EndDate	LoadDate	LoadUser
1	1	Anytime	Prepaid		5/6/2014		6-May-14	Admin
2	2	Off Peak	Prepaid		5/6/2014		6-May-14	Admin
3	3	TopUp Option	Prepaid		5/6/2014		6-May-14	Admin
PricePlanKey	PricePlanID	Priceplan	PricePlanGroup	PricePlanDescription	StartDate	EndDate	LoadDate	LoadUser
1	1	Anytime	Prepaid		5/6/2014	7/17/2014	6-May-14	Admin
4	1	Anytime	Top-Up		7/17/2014		7/17/2014	Admin
Dim_PricePlan updated with new PricePlanGroup								

Figure 3-32: Illustrative example: SCD type two

A SCD type two does not allow simultaneous views of the historical and current status' of the record (Ross, 2013b). Type three attempts to eliminate this shortfall; it adds a tracking column. This SCD change has a limitation – it keeps limited history because it can only keep the latest history. Figure 3-33 below is an illustration of type three.

Effective						History			
PricePlan Key	PricePlanID	Priceplan	PricePlan Group	PricePlanDescription	StartDate	EndDate	PricePlanGroup_History	Effective_Until_Date	LoadUser
1	1	Anytime	Top-Up		7/17/2014		Prepaid	7/17/2014	Admin

Figure 3-33: Illustrative example: SCD type three

The SCD type four is characterised by splitting a group of dimension attributes into separate mini dimension (Ross, 2013b). This technique is a remedy for volatile attributes where historical data are kept in one table and current data are kept in the other (Nguyen *et al.*, 2007). Figure 3-34 below is an illustration of type four.

Dim_PricePlan								
PricePlanKey	Price PlanID	Priceplan	PricePlan Group	PricePlanD escription	StartDate	EndDate	LoadDate	LoadUser
1	1	Anytime	Top-Up		7/17/2014		7/17/2014	Admin
Dim_PricePlan_History								
PricePlanKey	Price PlanID	Priceplan	PricePlan Group	PricePlanD escription	StartDate	EndDate	LoadDate	LoadUser
1	1	Anytime	Prepaid		5/6/2014	7/17/2014	6-May-14	Admin

Figure 3-34: Illustrative example: SCD type four

Types zero to four are the most used SCD types. Table 3-3 below summarises the SCD types.

Table 3-3: SCD summary (Ross, 2013b)

SCD type	Dimension table action	Impact on fact analysis
0	No change to attribute value.	Facts associated with attributes original value.
1	Overwrite attribute value.	Facts associated with current value.
2	A new dimension row for profile with new attribute.	Facts associated with attribute value in effect when fact occurred.
3	Add new column to preserve attributes current and prior values.	Facts associated with current and prior attributes alternative values.
4	Add a min dimension containing rapidly changing attributes.	Facts associated with rapidly changing attributes in effect when fact occurred
5	Add type 4 mini dimension along with overwritten type1 mini dimension key in base dimension.	Facts associated with rapidly changing attributes in effect when fact occurred, plus current rapidly changing values.
6	Add type 1 overwritten attributes to type 2 dimension row; all overwrite all prior dimension rows.	Facts associated with attributes value in effect when fact occurred, plus current values.
7	Add type 2 dimension row with new attribute value, plus view limited to current rows and /or attribute values.	Facts associated with attribute value in effect when fact occurred, plus current value.

3.6.8.4 Fact tables

Fact tables are a focus of dimensional analysis (Firestone, 1998). Fact tables are tables used to store events or transactions via foreign keys of dimensions that carry the detailed information (Kimball *et al.*, 2008d:204). Rahmatian (2003:480) defines a transaction “as a business event whose relevant attributes needs to be recorded internally (in the corporate database) as well as externally (for a customer, a specific supplier, a business partner, etc.) due to the impact this information will later have on other operations of the organisation and its stakeholders.” Kimball (2008) highly recommends these events to be at the lowest, non-summarised level of detail (i.e. lowest grain). Figure 3-35 illustrates a very detailed FactTransaction that is part of the illustrative star schema used in this chapter.

FactTransaction													
Transaction Key	Time Key	MobileNumberKey	Address Key	SubscriberKey	CallCategory Key	PricePlan Key	CallDuration	Voucher Number	Charges	CallStart Time	CallEnd Time	StartDate	LoadDate
1	1462	1	1	1	3	1	45.12	1	28.51	52:51.7	52:51.7	1/1/1900	52:51.7
2	1462	2	2	2	2	1	43.42	2	18.76	52:51.7	52:51.7	1/1/1900	52:51.7
3	1462	3	3	3	1	2	15.39	3	10.02	52:51.7	52:51.7	1/1/1900	52:51.7
4	1462	4	4	6	2	2	28.56	4	47.95	52:51.7	52:51.7	1/1/1900	52:51.7
5	1462	5	5	5	1	3	16.21	5	28.78	52:51.7	52:51.7	1/1/1900	52:51.7

Figure 3-35: Illustrative example: Fact transaction

Fact tables store numerical data on which various kinds of aggregations, such as sum, count, variance, and average, are calculated and stored (Kimball, 2008). There is a second type of a fact table where a summarised snapshot can be created based on the transactional grain of the fact table. This fact table is important where data is rarely queried to the lowest grain (Kimball & Ross, 2002: 17). For example, consider the telecommunication star schema presented earlier in this chapter. Figure 3-36 illustrates a transaction summary table at the granularity of monthly call category.

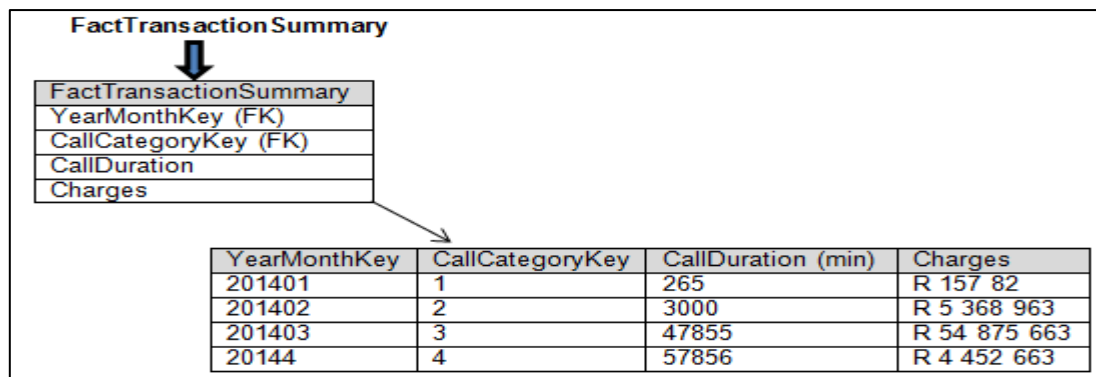


Figure 3-36: Illustrative example: Summary table: granularity: monthly call category

Because of large volumes of data and the need for a higher data presentation level, data aggregation levels are introduced (Kimball *et al.*, 2008d:119). Aggregates are summarised versions of atomic data stored at table level. For example, data can be stored at a transactional level and presented as weekly summaries per department and monthly per department. Drilling down and up using aggregates are different from that of atomic fact tables, because aggregates are summarised versions of atomic data according to specific requirements (Kimball, 2013).

3.6.8.5 Online analytical processing

A DW design follows a multi-dimensional architecture of conformed dimensions and fact tables to store data (Body *et al.*, 2013:581, O'Leary, 1999:54). This architecture supports OLAP systems by giving users a view of a cube, in which a user can choose the subset of his/her interest out of the whole multi-dimensional space (Chen *et al.*, 2002:1182).

The OLAP approach is used for analysing data in multi-dimensional perspectives; it uses complex operations such as rollup, drill down, slice and dice, and pivot (Datta & Thomas, 1999:289, Wang & Guo, 2013:40). These perspectives are built into a cube. A cube contains cells that are referred to as measures in a set of dimensions. For example, consider the telecommunication star schema used for illustrative purposes throughout chapter. Figure 3-37 shows a transaction cube based on the telecommunication transactions star schema. This cube contains cells with values of interest to the business users, i.e. "what is the price plan, for North west prince, for

the month January?” The answer to this is 8. The cell containing value 8 corresponds to Dim_Time Dim_Priceplan, and Dim_Province.

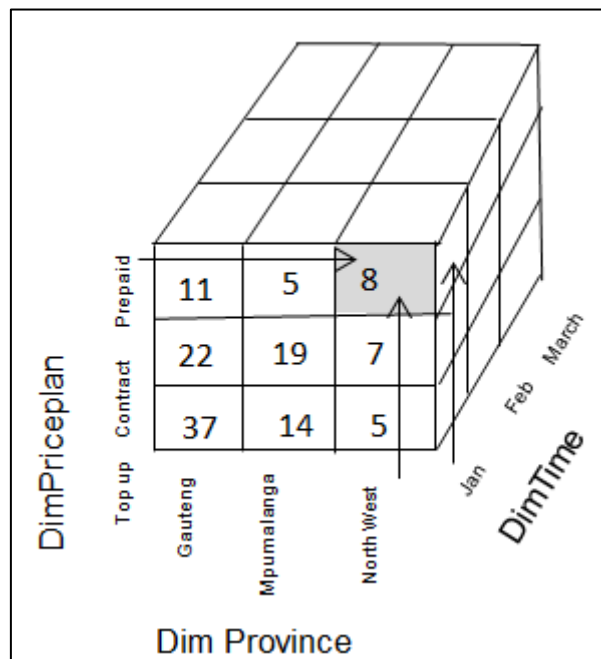


Figure 3-37: Illustrative example: Transaction cube

In OLAP systems, users are allowed to issue analytical queries, as opposed to transactional queries (Loukopoulos *et al.*, 2001:419). Hierarchies defined in dimensions allow aggregations, i.e. SUM (), COUNT (), AVG () etc., to be viewed at different levels of detail in a cube (Yuan & Zou, 2006:1019) Rollup, drill down, slice and dice, and pivot are OLAP operations that are used to navigate through historical data stored in a DW at different granularities (Talwar & Gosain, 2012:460). Considering the illustrative star schema used throughout this chapter, the OLAP cube built from this star model is to sum up the call duration and charges. This OLAP cube offers the capabilities of navigating from less detailed (summarised) to more detailed level (less summarised). Figure 3-38 below illustrates the call duration and charges rolled up and rolled down for year 2014.

Calendar Year	Calendar Quarter	Month Num Of Year	Full Date	Call Duration	Charges
2014	1	January		4584.44	4442.66
		February		4030.45	4304.83
		March		4426.47	4522.25
		Total		13041.36	13269.74
	2			13294.7	14338.29
	3			14278.19	13297.61
	4			14076.34	14148.18
	Total			54690.5900000001	55053.82
Grand Total				54690.5900000001	55053.82

Figure 3-38: Illustrative example: Roll up and down

Dimensions allow the slicing and dicing of data stored in a fact table, i.e. dimensions allow users to see sales figures for a specific time period and based on the demographics of customers. Multiple dimensions can be added to slice the data even deeper and to zoom in at a point in time and from different perspectives. A rotated data view, to give an alternative picture using dimensions, is another strong point of OLAP; it is referred to as pivot. This can be achieved using visual tools. Figure 3-39 below is an illustration of slice and dice. Here, the users can view call durations and charges associated with subscriber name and his/her mobile number (cell-phone) for the months January and February 2014.

		Calendar Year Calendar Quarter Month Num Of Year							
		2014							
		1						Total	
		January		February		Total			
Last Name	Mobile Number	Call Duration	Charges	Call Duration	Charges	Call Duration	Charges	Call Duration	Charges
Davuds	836958547	806.67	715.29	618.33	737.89	1425	1453.18	1425	1453.18
	Total	806.67	715.29	618.33	737.89	1425	1453.18	1425	1453.18
Ndulula	784568215	653.08	594.57	655.08	757.91	1308.16	1352.48	1308.16	1352.48
	Total	653.08	594.57	655.08	757.91	1308.16	1352.48	1308.16	1352.48
Nko	786925364	758.25	759.33	822.78	697.26	1581.03	1456.59	1581.03	1456.59
	Total	758.25	759.33	822.78	697.26	1581.03	1456.59	1581.03	1456.59
Pillar	839565445	584.04	816.8	638.8	653.79	1222.84	1470.59	1222.84	1470.59
	Total	584.04	816.8	638.8	653.79	1222.84	1470.59	1222.84	1470.59
Richards	784568755	838.88	866.08	741.35	674.83	1580.23	1540.91	1580.23	1540.91
	Total	838.88	866.08	741.35	674.83	1580.23	1540.91	1580.23	1540.91
Sono	836548569	943.52	690.59	554.11	783.15	1497.63	1473.74	1497.63	1473.74
	Total	943.52	690.59	554.11	783.15	1497.63	1473.74	1497.63	1473.74
Grand Total		4584.44	4442.66	4030.45	4304.83	8614.89	8747.49	8614.89	8747.49

Figure 3-39: Illustrative example: Cube slice and dice

After data sourcing, extracting, and complex transformation of data for loading into a DW model, data are made readily available using BI front-end user applications. These are the faces of a DW that provide functionality and deliver results in the desired form and format (Kimball *et al.*, 2008d). Data presentation (reporting),

traditional delivery methods coupled with data security issues, standard reports, and performance monitoring are discussed next.

3.6.8.6 Data presentation (Reporting)

Reporting occurs between the presentation server and the BI applications in the architecture model (Kimball *et al.*, 2008d:124). The focus of BI is not limited to the processing of data, but also extends to data distribution to the users; data being distributed must be in the right format and size (ITNEXT, 2013b). Reporting entails delivering of information in the format needed by users. Large volumes of data in a DW are categorised according to business process areas, partitions, dimensions, facts, and measures (Ballard *et al.*, 1998:35, Soler *et al.*, 2008:321). As a result, reporting should be narrowed down to provide specific and relevant reports. Data must be filtered so that a user is shown only the subset that he/she is interested in.

The experience that users have regarding data reported to them reflect on a BI/DW solution; on the other hand, data are useless if it is not used for reporting (Sharmila & Subramani, 2013:1) Organisations have reporting hierarchies that inform the levels of reporting and detail of data. Top-level management (senior management), middle-level management (tactical management), and low-level managers (operational management) are known as the management hierarchy (Beekman, 2015). Top-level management is responsible for the long-range issues related to the business' growth and development. In addition, they are responsible for planning and controlling the overall organisational direction. Middle-level management consists of branch and department managers; they are accountable for the functioning of their branches/departments. They are largely responsible for executing organisational tactics. Low-level managers (operational management) include supervisors or section leads. They are responsible for controlling and directing the day-to-day operations; they also assign tasks to employees, guide employees, and supervise employees' daily activities (Beekman, 2015:14).

3.6.8.7 Traditional data delivery methods

Information must ultimately enrich users and decision makers. The means of delivering information is another fundamental aspect in BI/DW. For example, information can be stored on a shared drive (such as a SharePoint site) for users to access, or it can be emailed directly to users. However, there are also other dynamic ways, i.e. report subscriptions and scheduled SQL jobs, which can be used to distribute information to users. A report subscription is a standing request to deliver a report at a specific time, or in response to an event, and in a specified format (Microsoft, 2014). Microsoft Reporting Services (MSRS) is an example of a reporting tool that can be used to set up a report to be delivered by means of standard or data driven subscriptions. A standard subscription contains static values (delivery time, parameters, and output format) that cannot be varied during subscription execution of the report, whilst data driven values can change at runtime.

The most recent delivery method is allowing users to interact with data in the form of self-service BI/DW. Constantly changing BI/DW requirements, inability of information technology (IT) personnel to meet business requirements, the need to be a more analytical driven organisation, slow access to information, as well as lack of information technology and DW/BI skills, are some of the reasons that led to self-service BI/DW (Imhoff & White, 2011:9). This style of reporting gives users freedom to integrate, prepare, visualise, and share data. Self-service analysis is achieved by having users with appropriate knowledge of data and tools used in the process.

Self-service analysis led to the term self-service BI/DW which is aimed at easy access to data as well as easy-to-use BI/DW tools, and fast-to-deploy, easy-to-manage, and simple user interfaces (Imhoff & White, 2011:9). Microsoft Excel is an example of a powerful front-end tool used for self-service BI. By setting up the data connection to connect to an OLAP cube with the right connection string, users can build pivot views of data, dynamically and without the help of IT personnel. The cube updates in the background, and users simply refresh the connections for updated data to be loaded. Figure 3-40 illustrates self-service BI/DW for the telecommunication transactions cube based on the star model shown earlier in this chapter. Here, the user can connect to the cube and build pivot tables to view the

charges over time (year 2014) and according to a geography (provinces) using Excel. The breakdown of call charges over a period (year, quarter, month, and day) and by province is retrieved instantly by the user from the cube.

Charges	Column Labels		
Row Labels	Guateng	North West	Grand Total
2014	46095.01	8958.81	55053.82
1	11099.77	2169.97	13269.74
January	3683.33	759.33	4442.66
February	3607.57	697.26	4304.83
March	3808.87	713.38	4522.25
2	12059.57	2278.72	14338.29
April	3874.21	832.59	4706.8
May	3860.83	730.16	4590.99
June	4324.53	715.97	5040.5
3	11057.81	2239.8	13297.61
July	3869.05	726.26	4595.31
August	3611.15	723.22	4334.37
September	3577.61	790.32	4367.93
4	11877.86	2270.32	14148.18
October	3942.73	729.65	4672.38
November	3950.32	741.28	4691.6
December	3984.81	799.39	4784.2
Grand Total	46095.01	8958.81	55053.82

PivotTable Field List

Choose fields to add to report:

- Fact Transaction
 - Call Duration
 - ☒ Charges
 - Fact Transaction Count
 - Voucher Number
- Dim Address
 - Hierarchy
 - Province

Drag fields between areas below:

Report Filter

Column Labels

Province

Row Labels

Period hierarchy

Values

Charges

☐ Defer Layout Update

Update

Figure 3-40: Illustrative example: Self-service BI (Microsoft Excel pivot)

Reporting covers both analytical (OLAP) and enterprise (formatted) reporting (Gile, 2006:3). Reporting is a data visualisation process. Visualisation techniques are to surge the scope of our dealings with information (Rao, 2003:32). Data can be reported in different formats, i.e. individual records, summaries, graphs, dashboards etc. Organisations attempt to standardise on a single reporting and analysis platform to deliver analytic, business, and enterprise reporting (Gile, 2006:1).

3.6.8.8 Standard reports

Business reports have a pre-approved, standard format in which data are disseminated (Gile, 2006:1). These types of reports can be requested at any time, they can be scheduled to run at approved times, or they can be parameterised. Parameterised reports are reports that allow users to retrieve from the report only the information that relates to selections as per the report query. This type of reporting saves time and brings only information that the user needs. Figure 3-41 shows an

illustration of report parameters where users can choose starting month, ending month, and the product they wish to be returned via the report.

Figure 3-41: Illustrative example: Report parameters

3.6.8.9 Performance monitoring (score cards and dashboards)

Scorecards combine dashboard metrics and apply them to a strategic map with key performance indicators (KPIs); this is used to measure organisational performance (Schlegel *et al.*, 2013:2). A business intelligence dashboard is a data visualisation tool that displays the current status of an enterprise in the form of metrics and KPIs (Cobbin, 2012). This style of reporting facilitates a status update process and keep managers informed on how business is performing. Figure 3-42 shows a performance dashboard of a telecommunication BI system. The dashboard measures the monthly call charges by province, and the call charges by price plan from January to date. Deployment is discussed in the next section.

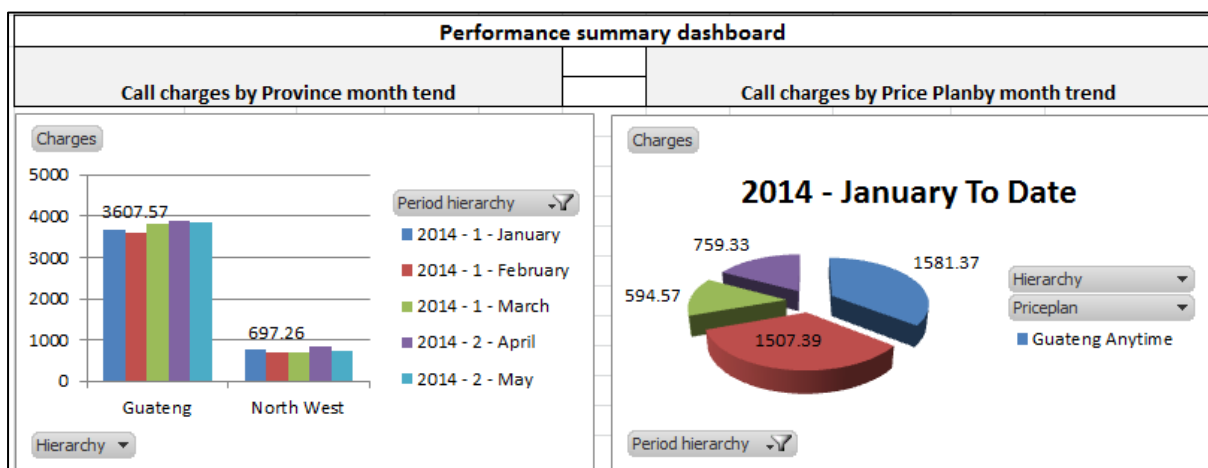


Figure 3-42: Illustrative example: Dashboard

3.6.9 Deployment

After development, a BI/DW solution must be presented to the business and its users. This process is referred to as deployment. System functionality and

performance is now tested against business requirements. Both unit (single) and system tests are conducted. Testing involves testing procedures, data quality, operations processing, performance, usability, deployment testing, and user desktop readiness (Kimball *et al.*, 2008a:341). The environment that will be used to host the systems must also be prepared; the project manager (with the BI/DW team) prepares a deployment plan and schedule. Deployment entails documentation, preparation of training for users, and organisation of user support processes (Kimball *et al.*, 2008a:341). Maintenance and growth is discussed next.

3.7 Maintenance and growth

Business intelligence/data warehousing solutions are forever changing to adapt to rapid data growth and business requirements that change. All life cycle activities require maintenance to allow for growth (Kimball *et al.*, 2008a:7). There are always other business cases that arise after the solution is deployed; this must be catered for as part of maintenance. For example, business may require adding a dimension or fact table (mart) into the DW. This change does not only affect the dimensional modelling activity, but it will also affect the ETL, application design, and maintenance processes. Data velocity and volume becomes a major problem in the DW, and requires continuous maintenance and support. This will affect the database administrator's team. Enhancements to the BI/DW applications, report development, and maintenance are also handled at that stage.

3.8 Future business intelligence

The lineage of BI/DW is a clear trend indicator of the future of BI/DW; it is still evolving and more challenges and possibilities such as real time BI, big data, mobile BI/DW, and in-memory data processing usage are still arising (Bucher *et al.*, 2009:408, Zeng *et al.*, 2006:4722). Traditional BI/DW that is used by company executives, which provides them with reports, is fading away as the workforce becomes more mobile (Stangarone, 2012). The future of BI/DW is also shaped by technology; it determines how we develop and consume data. As an example, resolutions are changing, and direct manipulation through touch interfaces brings users closer to data (Perez, 2012).

3.8.1 Real time business intelligence

Despite the evolution of BI/DW, information users still blame BI/DW solutions for information delays; this may be mainly due to information integration (Sahay & Ranjan, 2008:34). Organisations still have a need for reporting that are as close as possible to real-time. Even though real-time means different times for different organisations, this gap still exists currently within BI. Real-time BI/DW means delivering information as soon as possible after the business event occurred (Sahay & Ranjan, 2008:34). There are many approaches that attempt to minimise the reporting time (e.g. optimising solutions and self-service BI/DW); however, it can still be improved upon.

3.8.2 Big data

The large amounts of historical data accumulated thus far by many organisations, and its current growth rate, resulted in the coining of the term “big data” (Tankard & Pathways, 2012:5). This growth gave rise to volume (amount of data), velocity (speed of data), and variety (range of data). Data variety is a challenge since organisations are trying to derive more meaning from structured and semi-structured data. It is still a challenge to establish value from huge volumes of data, as data are useless if not used in reporting (Sharmila & Subramani, 2013:5). Storing, processing, and analysing big data are challenges; there is a need for new tools and approaches to handle these processes (Cumbley & Church, 2013:603). Even though the sizes of data are increasing, BI/DW users expect faster results from their BI/DW environment; this is valid for standard reports as well as navigation through source data.

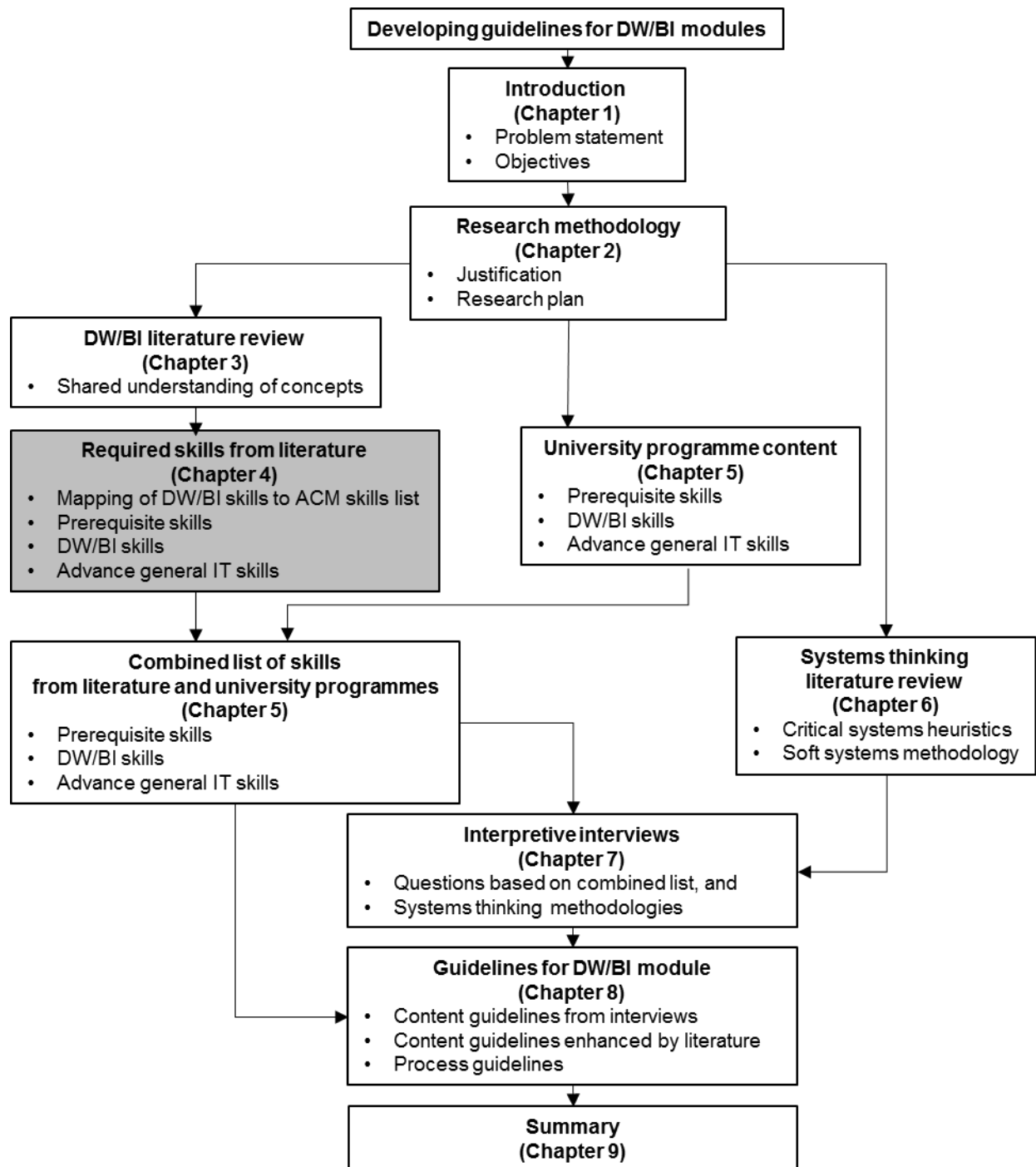
3.8.3 Mobile business intelligence

The existence of a seesaw analogy between technology capabilities and opportunities and business requirements brought about a shift for the report-centric paradigm (ITNEXT, 2013a). Wireless devices removed desktop restrictions and increased accessibility of data (Stangarone, 2012, Cowie & Burstein, 2007:372). Business' requirement to access data in many convenient ways increased and extended reporting from static to dynamic, and most recently, to mobile and social

collaborations; this process is called mobile BI/DW (Verkooij & Spruit, 2013:23). Tablets are, for example, ideal to render mobile BI/DW applications in order to encourage collaborative decision making. This is a more convenient, fast, and portable way of accessing data. The increased use of tablets and smartphones have already become part of the norm in many business environments (Nori, 2013:1). The dynamics of changing between different working environments also influence mobile and strategic decision making needs. Mobile BI/DW is a new trend; it is to be embraced by many organisations and decision makers.

3.9 Summary

The aim of this chapter is to outline the lineage, the components, and the purpose of BI. BI is concerned with storage, processing, and management of varied data volumes, the velocity at which this data can be stored and retrieved, the intelligence gained from data, and how it is reported (Russom, 2001:2). BI and analytics are both used together to provide answers to business. BI objectives are to eliminate data that is gathered but not analysed, reported but not understood, guessed about rather than acted upon. All these BI facets require trained professionals to drive them (Chen *et al.*, 2012:1182). Analytical, information technology (IT) knowledge, business knowledge, and communication skills are a summarised set of skills that BI professionals must acquire (Chiang *et al.*, 2012:4). The new areas of growth within BI, such as real time BI, big data, mobile reporting, and self-service BI, indicate that BI is still evolving. Skills required are identified from DW/BI literature.



Chapter 4 Guidelines from DW/BI literature review

4.1 Introduction

The objective of this study is to develop guidelines for data warehousing/business intelligence modules in information technology programmes at universities; the guidelines are based on an understanding of the needs of various stakeholders. Two methodologies, i.e. critical systems heuristics and soft systems methodology, are used to understand the various stakeholders' perspectives. Chapter 2 presented guidelines (the nature of a guideline is discussed in Section 2.2), and Chapter 3 presented business intelligence concepts, in order to create a shared understanding of the underlying concepts. These two chapters are fundamental to this chapter. The insights gained in Chapter 2 and Chapter 3 are used in this chapter to create the initial guidelines for business intelligence/data warehouse modules in information technology programmes.

Successful business intelligence professionals have a combination of technical skills and other capabilities; this is coupled with a broader frame of reference for the work, how it is carried out, and how it is measured. The pertinent question is: "How to ensure business intelligence students become successful business intelligence professionals?" Different stakeholders may have different viewpoints regarding the expertise that business intelligence professionals have, or ought to have, as well as what counts (should count) as relevant knowledge in this field. In an initial attempt to address the problem, this chapter aims to understand literature on what expertise is (or ought to be) possessed by business intelligence professionals or what counts (should count) as relevant knowledge in this field.

This chapter is organised as follows: Section 4.2 discusses the chapter methodology. Information from the information technology body of knowledge is presented in Section 4.3; this is done from an educational perspective. Section 4.4 provides guidelines based on the business intelligence/data warehouse literature review that was conducted in Chapter 3. Section 4.5 provides a summary for this chapter.

4.2 Methodology

The purpose of this section is to define the process that is used to develop the guidelines for DW/BI modules. As discussed in Section 2.2 of Chapter 2, a guideline is one form of strategy that can be used in the process of trying to achieve the best possible outcome. This strategy is characterised by a list of recommended steps or procedures on how to produce the best possible outcome. The purpose is to achieve a focused outcome. For the position of this study, the guidelines' focus is to better prepare business intelligence/data warehouse (DW/BI) students for industry. This outcome is achieved by creating a list of recommended knowledge areas and skills that students should have as per the literature review (refer to Chapter 3) and data collected from universities' programmes (presented in Chapter 5). Knowledge refers to learning concepts, principles, and information regarding DW/BI; skills refer to the ability to use and apply DW/BI information. However, for the purpose of this chapter, skills are used to refer to both skills and knowledge; good DW/BI skills require a sound theoretical knowledge base. These skills should serve as a guideline rather than a policy or a rule; so, they must allow for flexibility, adaptability in implementation, and regular revision.

The first three principles for developing guidelines (refer to Section 2.2 in Chapter 2) are applied in this chapter: *Principle one* suggests that guidelines must focus on the outcomes. This chapter provides a list of skills that should have been acquired by students who have successfully completed a DW/BI module, so that they are better prepared to add value in the DW/BI industry.

Principle two requires that guidelines are based on evidence. In this chapter the initial guidelines for DW/BI modules are formulated from evidence found in literature and enhanced in the next chapter from evidence found in universities' programmes.

Principle three requires the guideline formulation process to be systematic; it encourages decision makers to carefully consider the entire process and follow through on their choices (NHMRC, 1999:17, Johnson, 2011). In this chapter initial guidelines are systematically formulated from evidence by:

1. Investigating the general field of Information Technology (IT) and its categorisation (Section 4.3).
2. Identifying skills required as per DW/BI literature; and dividing the skills into more general IT skills and specific DW/BI skills (Section 4.4).

These steps are discussed in more detail in the next sections.

4.2.1 Information technology body of knowledge

First, the researcher presents work from the IT body of knowledge, and the categories of its knowledge areas, as described by the Association for Computing Machinery (ACM) (Lunt *et al.*, 2008:5). The aim of this is to use these categories and group the skills that an undergraduate degree must include accordingly; this can then provide a foundation for a DW/BI module.

4.2.2 Identify skills / knowledge according to DW/BI literature review

Second, the researcher identifies skills from the literature; refer to the review of BI and DW concepts in Chapter 3. The DW/BI lifecycle and architecture model of Kimball *et al.* (2008d:101) is used to guide the discussion and to:

1. Identify and list skills and their relevance as per the lifecycle and architecture components – refer to Sections 4.4.1-4.4.10.
2. Identify general IT skills from the list (refer to point 1) as guidelines in Section 4.4.11. This list is then divided into skills that are prerequisites for DW/BI, and advanced general skills in their application in DW/BI.
3. Identify core DW/BI skills from the list (as per the point 1 in this paragraph) as guidelines in Section 4.4.12.

The process must include all participants (*principle four* from Section 2.2 of Chapter 2). According to NHMRC (2011:13) a guideline development process must include all stakeholders; this section thus focuses on the perspectives of the international community (literature review) and that of universities (IT programmes). Views from

other stakeholders such as former students and DW/BI professionals are incorporated in Chapter 7.

So, the first step is to present a categorisation of the IT body of knowledge in order to provide a structure for the identified skills; this is discussed next.

4.3 Information technology body of knowledge

The Association for Computing Machinery (ACM) identified a body of knowledge that relates to undergraduate programmes (Lunt *et al.*, 2008:5). This body knowledge is subdivided into knowledge areas with core topics – refer to Table 4-1 below. Each knowledge area has units which are core learning outcomes (skills) that students obtaining a four year degree in the field must acquire.

Table 4-1: IT body of knowledge with core topics underlined (Lunt *et al.*, 2008:27)

No	The information technology body of knowledge	No	The information technology body of knowledge
1.	Information Technology Fundamentals • Pervasive Themes in IT • History of Information Technology • IT and Its Related and Informing Disciplines • Application Domains	2.	Information Management • Concepts and Fundamentals • Database Query Languages • Data Organization Architecture • Data Modelling • Managing the Database Environment • Special-Purpose Databases
3.	Programming Fundamentals • Fundamental Data Structures • Fundamental Programming Constructs • Object-Oriented Programming • Algorithms and Problem-Solving • Event-Driven Programming	4.	System Integration and Architecture • Requirements • Acquisition and Sourcing • Integration and Deployment • Project Management • Testing and Quality Assurance • Organizational Context • Architecture
5.	Human Computer Interaction • Human Factors • Aspects of Application Domains • Human-Centred Evaluation • Developing Effective Interfaces • Accessibility • Emerging Technologies • Human-Centred Computing	6.	Social and Professional Issues • Professional Communications • Teamwork Concepts and Issues • Social Context of Computing • Intellectual Property • Legal Issues in Computing • Organizational Context • Professional and Ethical Issues and

			Responsibilities - History of Computing - Privacy and Civil Liberties
7.	Platform Technologies - Operating Systems - Architecture and Organization - Computing Infrastructures - Enterprise Deployment Software - Firmware - Hardware	8.	IPT. Integrative Programming & Technologies - Intersystems Communications - Data Mapping and Exchange - Integrative Coding - Scripting Techniques - Software Security Practices - Miscellaneous Issues - Overview of Programming Languages
9.	Information Assurance and Security - Fundamental Aspects - Security Mechanisms (Countermeasures) - Operational Issues - Policy - Attacks - Security Domains - Forensics - Information States - Security Services - Threat Analysis Model	10	Web Systems and Technologies - Web Technologies - Information Architecture - Digital Media - Web Development - Vulnerabilities - Social Software
11.	System Administration and Maintenance - Operating Systems - Applications - Administrative Activities - Administrative Domains	12	Math and Statistics for IT - Basic Logic - Discrete Probability - Functions, Relations and Sets - Hypothesis Testing - Sampling and Descriptive Statistics - Graphs and Trees - Application of Math & Statistics to IT
13.	Networking - Foundations of Networking - Routing and Switching - Physical Layer - Security - Network Management - Application Areas		

The ACM's work is used to organise the skills identified skills from the literature review and other universities. The next section presents guidelines from literature.

4.4 Guidelines from DW/BI literature

Business intelligence (BI) is defined as an umbrella term that incorporates four components, i.e. data, processes, technology, and people (Ranjan, 2009b:62, Negash, 2004:170). This chapter focuses on the fourth component, i.e. the skills that DW/BI professionals must have. Kimball et al. (2008a:3) developed a DW/BI lifecycle methodology that can be followed to tie together activities and to better harness DW/BI project delivery. Figure 4-1 below presents the Kimball lifecycle methodology with its thirteen activities; these activities require specific skill sets. For this study, Kimball et al. (2008a:3) is used to guide the process to identify skill sets that are necessary in DW/BI.

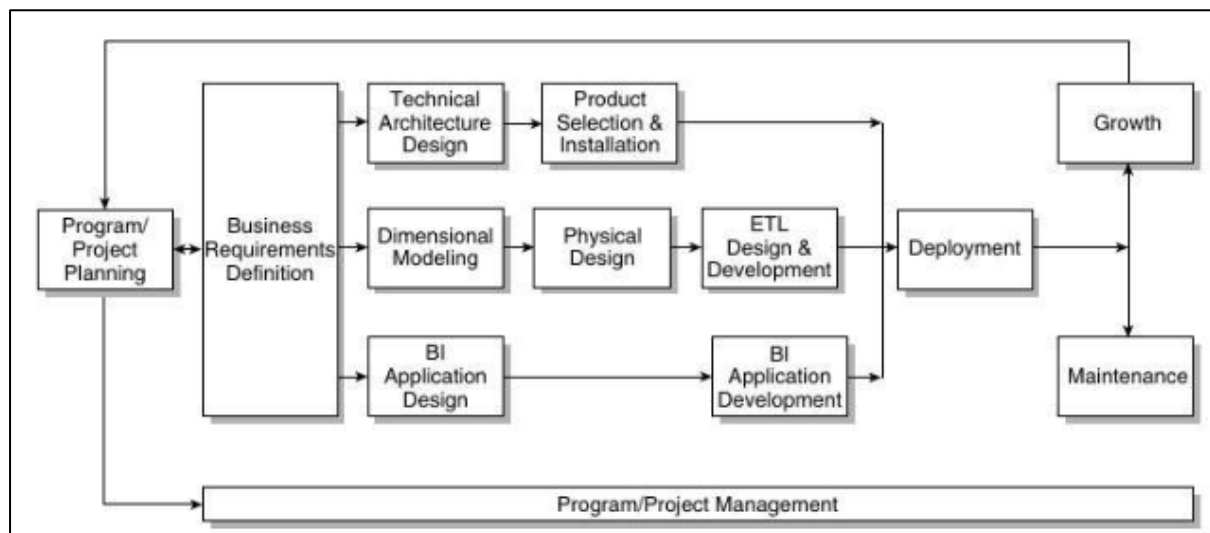


Figure 4-1: Kimball lifecycle diagram (Kimball et al., 2008a:3)

Knowledge and understanding of DW/BI components (data, technology, processes, and people), the data warehouse (DW) architecture, the DW lifecycle extract, load and transform process, and alternative DW methodologies is fundamental for a DW/BI professional. DW/BI project planning and management is presented in the next section.

4.4.1 Project planning and management

Chapter 3 highlighted that the success of a DW/BI solution depends heavily on the appropriate planning, integration, and management of the other twelve activities illustrated in Figure 4-1. Project management is one of the key activities that Kimball

et al. (2008a:4) discuss in the DW/BI lifecycle methodology. The project manager must determine organisational readiness and ensure that all activities are on track and in sync.

Skills embedded in project management include ensuring a continuously updated project status document, resource allocation, tracking, and change control management (Kimball *et al.*, 2008a:4). The skills required of a project manager of DW/BI projects, other than good DW/BI knowledge, are not substantially different from other IT project management skills. Project management skills are used in this study as representative of a group of skills that are required to manage IT projects well. Education regarding project management skills is required by DW/BI professionals.

4.4.2 Business requirements definition

As discussed in Chapter 3, it is essential to properly define requirements to ensure a successful DW/BI project. The requirements represent the needs of the organisation where the DW/BI system is developed. The role of business requirements is one of the main differences between the DW/BI development methodologies of Inmon (2002:389) and (Kimball *et al.*, 2008d:523). Inmon (2002:389) argues that the DW/BI must be developed from all available data; Kimball *et al.* (2008a:523) argue for an iterative process where requirements are used to prioritise the development of DW/BI functionality for specific business processes.

Kimball *et al.* (2008d:65) argue for the importance of an overall understanding of the business environment of the organisation. He divides business requirements into programme requirements that guides the entire DW/BI initiative, and project requirements that focus on a single project within the programme (Kimball *et al.*, 2008d:65). Project requirements are used in the track situated at the bottom of Figure 4-1 to develop specific business applications. Kimball *et al.* (2008d:248) promote a holistic design of the overall DW/BI initiative, rather than iteratively designing it. This is achieved by designing conformed tables for early development efforts that can be reused later. Kimball *et al.* (2008d:255) advocate the use of a matrix consisting of rows representative of business processes and columns of

representative of data entities or dimensions. The entries of the matrix indicate that specific business processes use specific data entities.

The skills mentioned by Kimball *et al.* (2008d:110) to gather business requirements are similar to the skills required for online transaction processing (OLTP) software development; these are therefore treated as a group of skills in this study.

The three central tracks in Figure 4-1 represents the DW architecture model; it is presented by Kimball *et al.* (2008d:101) as per Figure 4-2 in the next section. This figure is a more detailed illustration of the process than Figure 4-1; it will therefore be used to guide the discussion of the skills required by the central tracks of Figure 4-1. The discussion of the skills required for the remaining lifecycle components (deployment and growth) follows after the discussion of the skills required by the architecture model.

4.4.3 Technical architecture

The DW architecture was presented in Chapter 3. According to Kimball *et al.* (2008d:100) architecture answers the question: “how will we do it?”. This is referred to as the blueprint of a DW/BI solution. Architecture provides an answer by identifying components, their purposes, and the relationships between them. DW reporting, data management, security, bandwidth, and storage requirements are all inclusive in the DW/BI architecture (Inmon, 2000). Kimball and Inmon are known to be the two catalysts of DW/BI philosophies that influence the DW/BI world (George, 2012, Anupindi & Coady, 2005). However, there are also other architecture models that have fundamental differences and strong advocates. The Kimball architecture is used to guide the discussion in this study. A high level DW/BI system architecture model by Kimball *et al.* (2008d:101) is used to structure the guidelines’ creation process that follows. Figure 4-2 shows the high level DW/BI system architecture model.

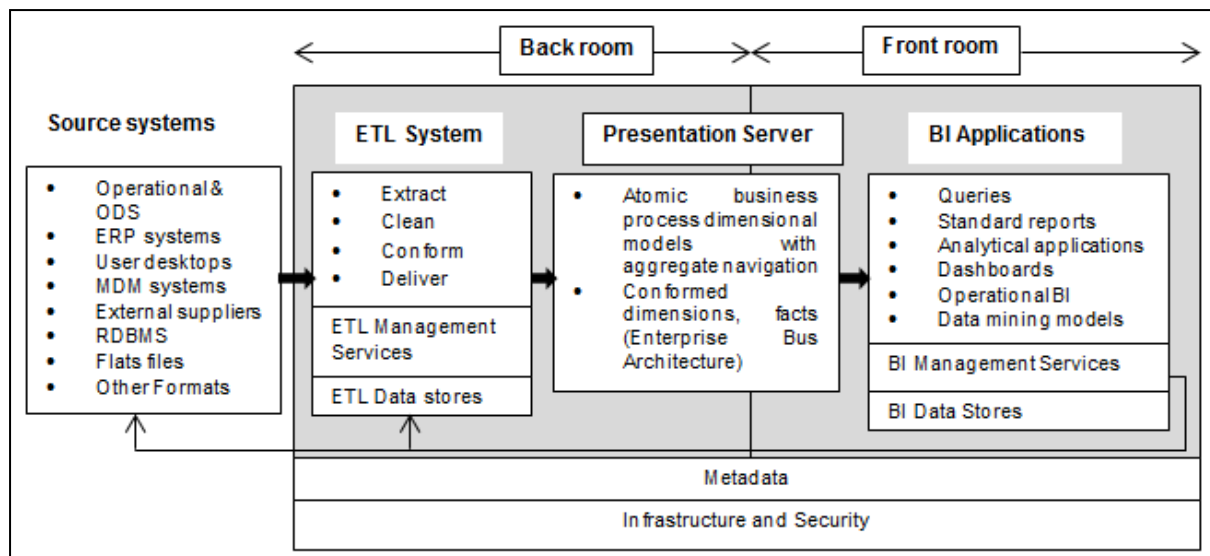


Figure 4-2: High level DW/BI system architecture model (Kimball et al., 2008b:101)

The DW/BI architecture components include: source systems, the extract, transform, load (ETL) system, the presentation server, the DW/BI applications, and metadata. In the following subsections these components are used to identify key DW/BI skill sets that are important in the DW/BI profession. The subsequent sections (Section 4.4.4-4.4.8) discuss these architecture components. The remaining lifecycle activities (deployment, growth, and maintenance) are discussed in Sections 4.4.9 and 4.4.10.

4.4.4 Source systems

The first component of the DW architecture is source systems. This component addresses the question about “how data is generated and organised”. Data are first captured in source systems. Data come from multiple systems that are built with multiple data stores and are hosted on multiple platforms (Kimball *et al.*, 2008d:108). A good understanding of source systems and business processes is a solid foundation for DW/BI (Bergamaschi et al., 2011:717). As discussed in Chapter 3, OLTP systems’ data are modelled using an entity relationship diagram (ERD) technique (Pan et al., 2013:1027); ERDs break data into smaller (atomic) tables that are related (Datta & Thomas, 1999:290). Basic understanding of an ERD is fundamental (Ballard et al., 1998:2). This implies a firm knowledge and understanding of database functionality, coupled with a programming skill set.

It is imperative for a DW/BI developer to be able to work with the table schema of the database (Alliance, 2015). Structured query language (SQL) is a programming language often used in the database space. This is a specialised language for updating, deleting, and requesting information from databases.

The understanding of different source systems, OLTP systems, ERDs, database functionality, and programming languages such as SQL are key competencies identified under the source systems component. The next section focuses on the ETL systems.

4.4.5 Extract, transform, and load systems

The second component of the DW/BI architecture is the ETL systems. The ETL systems are discussed in detailed in Chapter 3. The ETL systems' building process is considered the first comprehensive task in DW development; it is thus imperative to have a solid understanding of the source, the destination, and the mapping areas (El-Sappagh *et al.*, 2011:92, Breslin, 2004:15, Bergamaschi *et al.*, 2011:717).

According to Kimball (2004:1) many DW/BI professionals cannot break down the ETL processes to justify why it "is often estimated to consume 70 percent of the time and effort of building a data warehouse." Kimball (2004:1) identified 38 ETL subsystems that are required in the DW/BI backroom; this is an attempt to justify the immense effort required by the ETL process. Table 4-2 below lists the subsystems, gives a short description for each, and lists the skills required. Table 4-2 is a good illustration of the fact that an ETL project is list of small manageable iterative tasks. The detailed study of these subsystems is outside the scope of this study.

Table 4-2 Summary: 38 ETL subsystems and required skills as per Kimball (2004:1)

The 38 ETL Subsystems		
Subsystem	Description	Skills required
1. Extract system	Source, filtering, sorting, and data staging.	• ODBC/JDBC driver's knowledge • Data structures knowledge • Database programming • Knowledge of source systems (mainframes, relational databases, XML sources, flat files, Web logs, and enterprise resource planning (ERP) systems)
2. Change data capture system	Capturing changes made at the data source and extracting changed data only.	• Event programming • Database administration (DBA) skills
3. Data profiling system	Data rule analysis, and value rule analysis.	• Database programming • Analytical skills
4. Data cleansing system	Complete parsing of data in correct format.	• Database programming • Database administration (DBA) skills.
5. Data conformer	Enforcement of special conformed dimension attributes and fact table measures.	• Database programming • Dimensional modelling
6. Audit dimension assembler	Assembly of metadata context surrounding each fact table load.	• Database programming • Dimensional modelling
7. Quality screen handler	Checking for data quality issues.	• Database programming • Event programming
8. Error event handler	Reporting and responding to all ETL error events.	• Event programming • Database programming
9. Surrogate key creation system.	Producing stream of surrogate keys, independently for every dimension.	• Database modelling • Database programming
10. Slowly Changing Dimension (SCD) processor	Transformation logic for handling three types of time variance possible for a dimension attribute.	• Database programming • Dimensional modelling

11. Late arriving dimension handler.	Insertion and update logic for dimension changes that have been delayed in arriving at the data warehouse.	• Database programming • Dimensional modelling
12. Fixed hierarchy dimension builder	Data validity checking and maintenance system for all forms of many-to-one hierarchies in a dimension.	• Database programming • Dimensional modelling
13. Variable hierarchy dimension builder	Data validity checking and maintenance system for all forms of ragged hierarchies of indeterminate depth.	• Database programming • Dimensional modelling
14. Multivalued dimension bridge table builder	Creation and maintenance of associative (bridge) table used to describe a many-to-many relationship between dimensions.	• Database programming • Dimensional modelling • Database administration (DBA) skills
15. Junk dimension builder	Dimensions consisting of miscellaneous low cardinality flags and indicators found in most production data sources.	• Database programming • Dimensional modelling • Database administration (DBA) skills
16. Transaction grain fact table loader	Updates transaction grain fact tables including manipulation of indexes and partitions.	• Database programming • Dimensional modelling • Database administration (DBA) skills
17. Periodic snapshot grain fact table loader	System for updating periodic snapshot grain fact tables.	• Database programming • Dimensional modelling • Database administration (DBA) skills
18. Accumulating snapshot grain fact table loader.	System for updating accumulating snapshot grain fact tables including manipulation of indexes and partitions, and updates to both dimension foreign keys and accumulating measures.	• Database programming • Dimensional modelling • Database administration (DBA) skills
19. Surrogate key pipeline	Pipelined, multithreaded process for replacing natural keys of incoming data with data warehouse surrogate keys.	• Database programming • Dimensional modelling • Database administration (DBA) skills.
20. Late arriving fact handler	Insertion and update logic for fact records that have been delayed in arriving at the data warehouse.	• Database programming • Dimensional modelling • DBA skills
21. Aggregate builder.	Creation and maintenance of physical database structures to improve query performance.	• Database programming • Dimensional modelling • Database administration (DBA) skills.

22. Multidimensional cube builder.	Creation and maintenance of star schema foundation for loading multidimensional (OLAP) cubes, including special preparation of dimension hierarchies as dictated by the specific cube technology.	• Database programming • Dimensional modelling • Database administration (DBA) skills. • OLAP (aggregations, partitioning, calculations, calculations, etc.) skills
23. Real-time partition builder	Special logic for each of the three fact table types that maintains a “hot partition” in memory for incremental loads.	• Database administration (DBA) skills
24. Dimension manager system	Administration system for the “dimension manager”.	• Database programming • Dimensional modelling • Database administration (DBA) skills
25. Fact table provider system	Receives conformed dimensions sent by the dimension manager.	• Database programming • Dimensional modelling • Database administration (DBA) skills
26. Job scheduler	System for scheduling and launching all ETL jobs.	• Database administration (DBA) skills. (Production automation)
27. Workflow monitor.	Dashboard and reporting system for all job runs initiated by the Job Scheduler.	• Database administration (DBA) skills. (Production automation)
28. Recovery and restart system	Resuming halted and failed jobs.	• Database administration (DBA) skills. (Production automation)
29. Parallelising / pipelining system	Multiple processing for streaming data flows.	• Database programming • Dimensional modelling • Database administration (DBA) skills
30. Problem escalation system	Automatic plus manual system for raising an error condition.	• Database administration (DBA) skills
31. Version control system.	Check-out and check-in of all ETL modules and jobs.	• Database administration (DBA) skills. (Production automation)
32. Version migration system	Development to test production systems.	• Database administration (DBA) skills. (Production automation)
33. Compliance reporter	Comply with regulatory statutes to prove the lineage of key reported operating results.	• Reporting skills

34. Lineage and dependency analyser	Display the ultimate physical sources and all subsequent transformations of any selected data element.	• Database programming • Dimensional modelling • Database administration (DBA) skills
35. Security system	Administer role-based security on all data and metadata in the ETL pipeline.	• Database administration (DBA) skills
36. Backup system	Backup data and metadata for recovery, restart, and security.	• Basic security and assurance knowledge which are part of database administration (DBA) skills
37. Metadata repository manager	Capturing and maintaining all ETL metadata, transformation logic.	• Database programming • Dimensional modelling • Database administration (DBA) skills
38. Project management system	Comprehensive system for keeping track of all ETL development.	• Moving data can be a completed process that requires project management skills (Balasubramanyam, 2012)

It is evident from Table 4-2 that the 38 subsystems require skill sets that include both general IT skills as well as specialised DW/BI skills. The source systems and data destinations are modelled using standard models (e.g. ERDs or star schemas), whilst the ETL system (mapping area) does not have a standard model (El-Sappagha et al., 2011:91). This is another factor that impacts the immense effort required for the ETL process. Powerful ETL tools are used in this process and DW/BI professionals should have good knowledge of how they function. Sometimes the ETL tool is inadequate in terms of performance; DW/BI professionals must then use their acquired programming skills to enhance them (Balasubramanyam, 2012:1).

There are twelve key skills in the rightmost column of Table 4-2; these are identified from the literature review in Chapter 3. These knowledge bases are presented in Table 4-3 below.

Table 4-3: ETL skills set and knowledge

Skills / knowledge	Description
1. Analytical skills	Analytical skills are an important factor in the ETL systems. This means the ability to develop understanding, generating algorithms and analysis of a sequence of instructions that will automate performing a specific task or solving a given problem. For example, different DW/BI problems situations mandate different solutions that suit each scenario. Analytical skills such as “deductive database solutions, undertaking data mining exercises” can be applied (RMIT, 2015).
2. DW/BI tools knowledge	Application software and other technologies are used by DW/BI professionals to extract, transform, and load data to target destinations such as data stores and data warehouses. InfoSphere DataStage, Integration Services (part of SQL Server), Oracle Warehouse Builder (OWB), and Informatica are some of the leading ETL tools (Thoo et al., 2013:5).
3. Dimensional modelling	Dimensional models are the ultimate ETL deliverables, and thus some familiarity with the basic principles of dimensional modelling is essential (Kimball, 2015). Dimensional modelling is a technique that was invented by Kimball to conceptualise a business as a set of entities such as dimensions and facts tables and the relationships among them (Kimball, 2014, O’Leary, 1999:49). Dimensional models are discussed in details in Chapter 3.
4. Database programming	Solid database programming skills is essential in DW/BI. In Chapter 3 a running example is given where SQL scripts are used to clean up data.
5. Data structures	ETL development requires various data structures which developers must be familiar with.
6. Database drivers connections	ETL developers must be able to create a connection to different databases and others sources using drivers and such as ODBC/JDBC driver’s.
7. Database administration (DBA) skills.	The movement of data from different source systems via the use of the ETL systems to the presentation server, and the DW/BI application tools requires some DBA skills. Important processes such as access to data and server management (security, space and speed) mandate advanced DBA skills.
8. Programming including event programming.	Programming or coding has always been a necessary skill set in DW/BI. Inmon (2008:1) explains the inception of coding in BI/DW that one will have to “write code in order to get data from one source to the appropriate target.” Despite the advancement of technology where writing and maintaining transformation code has been shifted to automation, there is still a need for custom coding in the ETL process (Inmon, 2008, Alliance, 2015).
9. Online analytical processing (OLAP) skills	Aggregations, partitioning, calculations, calculations, etc.
10. Project	Moving data can be a complete process that requires project management

management skills	skills (Balasubramanyam, 2012).
11. Reporting skills	Reporting on ETL processes and stages.
12. Source systems knowledge	Knowledge and understanding of different source systems such as databases and flat files is essential to ETL developers.

The third DW architecture component (the presentation sever is) presented next.

4.4.6 Presentation server

Dimensional modelling is important for ETL; ETL can be described as the ultimate deliverable for a DW. Knowledge of how data are modelled, stored, and retrieved in the DW and online analytical processing (OLAP) cubes is of utmost importance. According to Kimball (2004: it can be safely assumed that DW/BI applications use *“dimensionally-structured, organized by business process”* data; knowledge regarding dimensional modelling is thus important for DW/BI professionals. Data can further be stored in OLAP cubes that have different atomically-grained levels, as well as calculated and aggregated summaries; this is to ensure optimal performance tuning. Data stored in a cube can be queried using MDX language; cube development and manipulation is thus a core DW/BI skill.

The presentation server is a data location area where data are managed by DW/BI teams. Managing the presentation server requires DW/BI professionals to have dimensional modelling and database administration (DBA) skills (Sonderegger & Strategist, 2012). Large amounts of data come in various forms and velocity; DBA skills and knowledge of unstructured data (“Big Data”) is thus a necessity. There is also an element of performance tuning for fast storing, retrieving, and managing of data (Sonderegger & Strategist, 2012).

The researcher concludes from the above that knowledge of dimensional modelling, OLAP, DBA skills, as well as MDX and SQL programming are competencies that are important in the presentation server component.

4.4.7 BI/DW applications

BI/DW applications are tools that provide management and organisational employees with reliable information and insights, in order to enhance decision making and social collaboration, and facilitate better performance. These tools provide means for efficient reporting; reporting is facilitated by analysis of data, statistics and analytics, and dashboards that display relevant indicators. These tools enable a skilled business user to define reports; a DW/BI developer must maintain and customise the reports (Thorntwaite, 2006). The DW/BI team must equip themselves with a wide range of expertise in order to deliver information in a correct and timely fashion. This range includes delivering information options such as user interface design, *ad hoc* reporting and querying, analysis, trending and predictive graphical techniques, as well as information push (email/intranet) and pull (on request).

Knowledge of reporting tools and skills are required to create and deliver data for dashboards and static reporting structures. DW/BI professionals must have a general or basic understanding of the *“underlying theory and application of key reporting software on the market”* (Alliance, 2015). Some of the DW/BI applications have the *“underlying scripting language in reporting software”* that makes basic programming skills a necessity (Alliance, 2015). Users require an easy-to-use front end application (interface) to access data in the DW; it must provide comprehensive analytical capabilities (Collier, 2011). The goal is to tailor these tools to the way users work; this requires skills regarding custom programming or commercial application customisation. Knowledge of SQL and database reporting is essential. These analytical skills also include working with spreadsheets and data mining tools.

User requirements require an organised way to determine customer needs; so, an understanding of the systems development lifecycle (SDLC) is a necessity. The development of DW/BI user interfaces follows the DW/BI lifecycle methodology (Kimball *et al.*, 2008a:3).

Business intelligence leverages other systems and processes that already exist, e.g. OLTP. Accordingly, the dependence of a DW upon a variety of already existing

databases leads thereto that a DW discussion starts with at the OLTP systems (Nilakanta *et al.*, 2008:263, Rifaie *et al.*, 2008:58). There is dependency relationship between OLTP and DW/BI that requires a general understanding of the difference between the two.

Developers that work in DW/BI inevitably become proficient in understanding the business problems and transforming them into solutions (Pitre, 2013). DW/BI tool sets require developers to have a general business understanding to accomplish this. Technological advancements also resulted in mobile DW/BI. The use of wireless devices removed desktop restrictions and increased accessibility to data (Stangarone, 2012, Cowie & Burstein, 2007:372).

In summary, DW/BI professionals must have: general business knowledge, programming skills, database reporting skills, SDLC skills, user interface designing skills, Microsoft Excel skills, mobile DW/BI technology skills, big data understanding and handling skills, data mining skills, and an understanding of the difference between OLTP and DW/BI.

4.4.8 Metadata

Generally metadata is described as data about data. However Kimball *et al.* (2008d:115) argue that this definition is too shallow and does not paint a clear picture in a mind. Kimball *et al.* (2008d:115) define metadata as “all the information that defines and describes the structures, operations, and contents of the DW/BI system”. This includes a clear definition of technical components and business processes that make up the DW/BI system, and how they work together.

The argument that students should have metadata skills also entail an understanding of the business, technical, and process metadata required by all the different architecture components of the DW. Like professionals, students must be able to create and develop procedures to maintain metadata.

4.4.9 Deployment

The final activity in the initial DW iteration of the lifecycle methodology is deployment (Kimball *et al.*, 2008a:115). Deployment refers to the process of preparing to hand over the DW/BI solution to the business users. There are key activities that happen in the deployment process of the DW/BI solution; it includes testing, preparing the hosting environment, documentation, preparing for user training, as well as system and organisational support.

Students should have skills related to, and an understanding of: DW/BI testing, the deployment process, how to document the process, as well as end-user training and support.

4.4.10 Growth and maintenance

The nature of DW/BI solutions is characterised by continuous change and growth; maintenance is thus key so that an organisation can avoid being 'a dinosaur'. There are always other business cases that arise after the solution is deployed; they must be catered for as part of maintenance. A change in business requirements, or in velocity and data volumes, may affect all activities in the cycle methodology; hence, the execution of the methodology is cyclic.

The role of good metadata management is crucial in maintaining and extending the DW/BI solution. The skills required for these phases of the DW/BI solution are encapsulated in the phases and components described in this section.

4.4.11 Skills from literature review

This section presents a summary of the base skill sets identified from Section 4.4.1 to Section 4.4.10. This process involves multiple cross referencing and comparison until a final traceable set is established. This process necessitates a coding scheme to help with traceability of skills and their sources (i.e. where and how they were identified). The following scheme is derived and used going forward in the chapter.

Lit + sequential number + group + skill short name

Lit = Literature (indicates that skill is from literature review).

Sequential number = Count number starting from 1 up to the maximum number of identified skills.

Group = Which group does the skill fall under (For example, Table 4-4 has five groups [1] DW/BI concepts, [2] Source systems, [3] ETL systems, [4] Presentation server, and [5] Business intelligence applications).

Skill short name = Short description of the skill required.

For example the first skill under DW/BI concepts number is represented as follows:

Lit_1_Concepts_ Components

Lit = Literature review

1 = Skill sequential number

Concepts = From DW/BI concepts (Group)

Components = Short name for DW/BI components knowledge

A summary of skill sets and the short descriptions of their applications are presented in Table 4-4 below. DW/BI concepts are discussed first. These skill sets are classified according the DW architecture components.

The discussion thus far demonstrates that DW/BI is a highly multi-skilled discipline. The DW/BI team delivers solutions using different techniques and skill sets that are either DW/BI specific or general IT. Some skills such as programming (Lit_14_Sources_DBProg, Lit_21_ETL_DBProg, and Lit_34_Presentation_DBProg) are repeated in different groups. This gives evidence that their relevance expand across DW/BI groups.

Table 4-4: Literature review skill sets/knowledge

No	Skills / knowledge	Description
BI concepts		
Lit_1_Concepts_Components	BI components knowledge (Data, technology, processes, and people).	BI is an umbrella term that covers data, technology, processes, and human knowledge.
Lit_2_Concepts_Architecture	DW/BI architecture knowledge and understanding.	According to Kimball <i>et al.</i> (2008d:100) architecture answers the question, "how will we do it?" part.
Lit_3_Concepts_Lifecycle	DW/BI lifecycle.	Overall knowledge of the components of the DW/BI lifecycle that supports the architecture.
Lit_4_Concepts_Methodologies	Knowledge of alternative DW/BI Methodologies.	Kimball and Inmon are known to be the two catalysts of DW/BI philosophies that are influencing the DW/BI industry (George, 2012, Anupindi & Coady, 2005). However there are other architecture models.
Project planning and management		
Lit_5_ProjPlan_ProjectM	Project management.	Project management is one of the key activities that Kimball <i>et al.</i> (2008a:111) discuss in the DW/BI lifecycle methodology.
Business requirements and analysis		
Lit_6_BusinessReq_Matrix	Business requirements collection and analysis in order to develop a business requirements matrix.	Programme and project business requirements collection and analysis are required to set up the business requirements matrix (Kimball <i>et al.</i> , 2008d:110).
Technical architecture		
Lit_7_TechArch_TechArchitect	Overall technical architecture knowledge.	According to Kimball <i>et al.</i> (2008d:100) architecture answers the question, "how will we do it?" part.
Lit_8_TechArch_DW_Reporting	DW/BI reporting.	DW/BI reporting, data management, security, band width, and storage requirements are all inclusive in the DW/BI architecture (Inmon, 2000).
Lit_9_TechArch_Data manage	Overall data management.	Component of technical architecture (Inmon, 2000).
Lit_10_TechArch_DataSecurity	Data Security.	Component of technical architecture (Inmon, 2000).
Lit_11_TechArch_Networking	Networking (band width and topologies).	Component of technical architecture (Inmon, 2000).

Lit_12_TechArch_HardwareKnow	Computer hardware knowledge (storage and computing requirements).	Component of technical architecture (Inmon, 2000).
Source systems		
Lit_13_Sources_DatabaseFunc	Database functionality.	Database functionality understanding.
Lit_14_Sources_DBProg	Database programming skills (SQL).	SQL database programming for updating, deleting, and requesting information from databases.
Lit_15_Sources_ERD	ERD knowledge.	A basic understanding of ERD is fundamental (Ballard <i>et al.</i> , 1998:2).
Lit_16_Sources_OLTP	OLTP understanding.	Knowledge of OLTP is important and it cannot be divorced from the ERD (Datta & Thomas, 1999:290).
Lit_17_Sources_SourceProcesses	Source systems and business processes knowledge.	A good understanding of source systems and business processes is a solid foundation for DW/BI (Bergamaschi <i>et al.</i> , 2011:717)
ETL systems (DW lifecycle extract, load and transform (ETL) functions)		
Lit_18_ETL_Analytical	Analytical skills.	Analytical skills such as “deductive database solutions, undertaking data mining exercises” can be applied (RMIT, 2015).
Lit_19_ETL_Tools	BI tools knowledge.	InfoSphere DataStage, Integration Services (part of SQL Server), Oracle Warehouse Builder (OWB), and Informatica are some of the leading ETL tools (Thoo <i>et al.</i> , 2013:3).
Lit_20_ETL_DimModeling	Dimensional modelling.	Dimensional models are the ultimate ETL deliverables, and thus some familiarity with the basic principles of dimensional modelling is essential (Kimball, 2015).
Lit_21_ETL_DBProg	Database programming.	Solid database programming skills is a must have in DW/BI.
Lit_22_ETL_DataStructures	Data structures.	ETL development is open to various data structures which developers must be familiar with.
Lit_23_ETL_Connections	Database driver connections.	ETL developers must be able to connect to different databases and others sources using drivers and such as ODBC/JDBC driver's.
Lit_24_ETL_DBA	Database administration (DBA) skills.	Important processes such as access to data and server management (security, space and speed) mandate serious DBA skills.
Lit_25_ETL_Prog	Programming.	Despite the advancement of technology

		where writing and maintaining transformation code has been shifted to automation, there is still a need for custom coding in the ETL process (Inmon, 2008, Alliance, 2015).
Lit_26_ETL_OLAP	OLAP skills.	aggregations, partitioning, calculations, calculations, etc.
Lit_27_ETL_ProjectM	Project management skills.	Moving data can be a completed process that requires project management skills (Balasubramanyam, 2012).
Lit_28_ETL_Reporting	Reporting skills.	Reporting on ETL processes and stages.
Lit_29_ETL_Source	Source systems knowledge.	Knowledge and understanding of different source systems such as databases and flat files is essential to ETL developers.
Presentation server		
Lit_30_Presentation_DimModeling	Dimensional modelling knowledge (e.g. Start schema).	Presentation server is a stage where data is managed by DW/BI team. Dimensional modelling skills are required. Since dimensional models are the ultimate ETL deliverables, some familiarity with the basic principles of dimensional modelling is necessary.
Lit_31_Presentation_DBA	DBA skills.	Managing data in various forms and velocity, and performance tuning make DBA skills a big necessity (Sonderegger & Strategist, 2012).
Lit_32_Presentation_MDX	Multidimensional dynamics expression (MDX) skills.	Data stored in a CUBE is queried using MDX language.
Lit_33_Presentation_OLAP	Knowledge of online analytical processing (OLAP) cubes.	Data is stored in different atomically-grained levels and pre-calculated and aggregated summaries for performance tuning.
Lit_34_Presentation_DBProg	Database Programming (SQL).	It is imperative for a DW/BI developer to be able to work with the table schema of the database and SQL (Alliance, 2015).
Business intelligence applications		
Lit_35_BIApplication_Business	General business understanding.	BI tool sets requires developers to have a general business understanding to accomplish that work (Pitre, 2013).
Lit_36_BIApplication_Requirements	Project requirements collection and analysis.	Collection and analysis of business requirements for specific business applications.
Lit_37_BIApplication_Prog	Programming skills.	Some of the DW/BI applications have the “ <i>underlying scripting language in</i>

		<i>reporting software</i> ” that makes basic programming a necessity (Alliance, 2015).
Lit_38_BIApplication _ DBReporting	Database reporting.	Knowledge of SQL and database reporting is also essential for creating user reports and dashboards.
Lit_39_BIApplication _ ReportTools	Knowledge of reporting tools.	Skills needed to create and deliver data using a structured reporting tool for creating dashboards and static reporting structures.
Lit_40_BIApplication _ SDLC	Systems development lifecycle (SDLC) understanding.	User requirements require an organised way to determine customer needs and thus SDLC understanding a necessity.
Lit_41_BIApplication _ Interface	User interface design knowledge.	Users require an easy-to-use front end application interface into the data warehouse that provides comprehensive analytical capabilities (Collier, 2011).
Lit_42_BIApplication _ MSEXcel	Microsoft Excel skills.	Microsoft Excel is powerful front end tool for self-service DW/BI. Pivot tables are often linked to OLAP cubes.
Lit_43_BIApplication _ MobileBI	Mobile DW/BI.	Wireless devices have removed desktop restrictions and have increased accessibility of data (Stangarone, 2012, Cowie & Burstein, 2007:372).
Lit_44_BIApplication _ BigData	Big data understanding and handling.	There is still a challenge of establish value from huge volumes of data because it is useless if not used for reporting (Sharmila & Subramani, 2013:5).
Lit_45_BIApplication _ DataMining	Data mining skills and knowledge.	Data analytics is grounded mostly in data mining and statistical analysis (Chen <i>et al.</i> , 2002).
Lit_46_BIApplication _ OLTP/BI	Knowledge of the difference between OLTP and DW/BI.	There is dependency relationship between OLTP and DW/BI that requires a general understanding of the difference between the two.
Metadata		
Lit_47_Metadata	Metadata knowledge including business, technical and process metadata.	Kimball <i>et al.</i> (2008d:115) define metadata as “all the information that defines and describes the structures, operations, and contents of the DW/BI system”.
Deployment		
Lit_48_Deploy_process	Deployment process.	Knowledge of the overall deployment process.
Lit_49_Deploy_Testing	Testing.	Knowledge of the importance and procedures of testing.

Lit_50_Deploy_Doc	Documentation.	Knowledge of the documentation required by testing and deployment.
Lit_51_Deploy_User_training	End-user training and support.	Knowledge and skills in user training and continuous support.

4.4.12 Categorisation of skills from literature review

A distinction between DW/BI specific and other or general IT skill sets must be made in order to compare information from universities that has literature on DW/BI. Universities' programmes, and especially modules in DW/BI, do not provide a list of all the general IT skills required to complete their DW/BI module.

Processes such as data sourcing, transformation, loading, and presentation make DW/BI unique within the IT discipline. There is a temptation to classify skill sets that is used to achieve these processes to be only DW/BI specific. This is not true because some skill sets are core to both DW/BI and other IT disciplines. For example, data extraction can be done by using programming language such as Java with embedded SQL scripts. This does not make Java a core DW/BI skill because it can be used across many IT disciplines. On the other hand, OLAP can be said to be a core DW/BI skill because it is specific to DW/BI. The task to categorise the identified skill sets in general IT and DW/BI specific skills sets require mapping the following tables: Table 4-1 (ACM) and Table 4-4 (skill sets derived from the literature review). The outcome of this is: two new lists containing general IT skills required for DW/BI development, and specific skills required for DW/BI development. The list for general IT skills is presented in Table 4-5 below. The first column ('No') represents the newly allocated skill number generated using the coding scheme described below. The second column '*Skills / knowledge*' represents the skill number reference from Table 4-4. The third column '*ACM knowledge area*' represents the knowledge area number and the item matching the skill. The following scheme is used to number skills:

Lit + Skill category + sequence number + description

Lit = Literature (indicates that skill is from literature)

Skill category = General IT skills (Gen) or DW/BI specific skills (BI)

Sequential number = Allocated number starting from 1 up to the maximum number of identifies skills.

Table 4-5: General IT skills / knowledge

No	Skills / knowledge (code number)	ACM Knowledge area
Lit_1_Gen_ProjectM	Lit_5_ProjPlan_ProjectM Lit_27_ETL_ProjectM	4. System Integration and Architecture
		Project Management
Lit_2_Gen_Requirements	Lit_6_BusinessReq_Matrix Lit_36_BIApplication_Requirements	4. System Integration and Architecture
		Requirements
		Acquisition and Sourcing
Lit_3_Gen_TechArchData	Lit_9_TechArch_DataManag	2. Information Management
		Data Organization Architecture
Lit_4_Gen_DBFunc	Lit_13_Sources_DatabaseFunc	2. Information Management
		Concepts and Fundamentals
Lit_5_Gen_DBProg	Lit_14_Sources_DBProg Lit_21_ETL_DBProg Lit_34_Presentation_DBProg	2. Information Management
		- Database Query Languages - Data Organization Architecture - Data Modelling
Lit_6_Gen_ERD	Lit_15_Sources_ERD	2. Information Management
		Data Modelling
Lit_7_Gen_OLTP	Lit_16_Sources_OLTP	2. Information Management
		- Data Organization Architecture - Data Modelling - Managing the Database Environment
Lit_8_Gen_Processes	Lit_17_Sources_SourceProcesses	4. System Integration and Architecture
		- Requirements - Organizational Context - Architecture
Lit_9_Gen_	Lit_18_ETL_Analytical	12. Math and Statistics for IT

Analytical		• Basic Logic • Discrete Probability • Functions, Relations and Sets • Hypothesis Testing • Sampling and Descriptive Statistics • Graphs and Trees • Application of Math & Statistics to IT
Lit_10_Gen_ETL_Tools	Lit_19_ETL_Tools	7. Platform Technologies
		• Operating Systems • Computing Infrastructures • Software • Firmware
Lit_11_Gen_DataStructures	Lit_22_ETL_DataStructures	3. Programming Fundamentals
		• Fundamental Data Structures
Lit_12_Gen_DBConnections	Lit_23_ETL_Connections	4. System Integration and Architecture
		• Acquisition and Sourcing • Integration and Deployment
Lit_13_Gen_DBA	Lit_24_ETL_DBA Lit_31_Presentation_DBA	2. Information Management
		• Managing the Database Environment
Lit_14_Gen_Prog	Lit_25_ETL_Prog Lit_34_Presentation_Prog	7. Platform Technologies
		• Fundamental Programming Constructs • Object-Oriented Programming • Algorithms and Problem-Solving • Event-Driven Programming
Lit_15_Gen_Reporting	Lit_28_ETL_Reporting	4. System Integration and Architecture
		• Project Management
Lit_16_Gen_Source	Lit_29_ETL_Source	4. System Integration and Architecture
		• Acquisition and Sourcing
Lit_17_Gen_Business	Lit_35_BIApplication_Business	4. System Integration and Architecture
		• Requirements • Organizational Context
Lit_18_Gen_DBReporting	Lit_38_BIApplication_DBReporting	4. System Integration and Architecture
		• Integration and Deployment

Lit_19_Gen_ReportTools	Lit_39_BIApplication_ReportTools	7. Platform Technologies
		• Operating Systems • Computing Infrastructures • Software • Firmware
Lit_20_Gen_SDCL	Lit_40_BIApplication_SDCL	4. System Integration and Architecture
		• Requirements • Acquisition and Sourcing • Integration and Deployment • Project Management • Testing and Quality Assurance • Organizational Context • Architecture
Lit_21_Gen_Interface	Lit_41_BIApplication_Interface	5. Human Computer Interaction
		• Human Factors • Aspects of Application Domains • Human-Centred Evaluation • Developing Effective Interfaces • Accessibility • Emerging Technologies • Human-Centred Computing
Lit_22_Gen_MSExcel	Lit_42_BIApplication_MSExcel	5. Human Computer Interaction
		• Human-Centred Computing
Lit_23_Gen_BigData	Lit_44_BIApplication_BigData	2. Information Management
		• Managing the Database Environment • Special-Purpose Databases
Lit_24_Gen_DataMining	Lit_45_BIApplication_DataMining	12. Math and Statistics for IT
		• Basic Logic • Discrete Probability • Functions, Relations and Sets • Hypothesis Testing • Sampling and Descriptive Statistics • Graphs and Trees • Application of Math & Statistics to IT
Lit_25_Gen_OLTP/BI	Lit_46_BIApplication_OLTP / DW/BI	2. Information Management
		• Concepts and Fundamentals • Database Query Languages • Data Organization Architecture • Data Modelling • Managing the Database Environment

		• Special-Purpose Databases
Lit_26_Gen_Metadata	Lit_47_Metadata	2. Information Management
		• Managing the Database Environment
Lit_27_Gen_Deployment	Lit_48_Deployment process Lit_50_Deploy_Doc	4. System Integration and Architecture
		• Integration and Deployment
Lit_28_Gen_Testing	Lit_49_Deploy_Testing	5. Human Computer Interaction
		• Accessibility
Lit_29_Gen_UserTraining	Lit_51_Deploy_User_training	6. Social and Professional Issues
		• Professional Communications
Lit_30_Gen_DataSecurity	Lit_10_TechArch_DataSecurity	9. Information Assurance and Security
		• Fundamental Aspects • Security Mechanisms (Countermeasures) • Operational Issues • Policy • Attacks • Security Domains • Forensics • Information States • Security Services
Lit_31_Gen_Networking	Lit_11_TechArch_Networking	13. Networking
		• Foundations of Networking • Routing and Switching • Physical Layer • Security • Network Management • Application Areas
Lit_32_Gen_HardwareKnow	Lit_12_TechArch_HardwareKnow	7. Platform Technologies
		• Operating Systems • Architecture and Organization • Computing Infrastructures • Enterprise Deployment Software • Firmware • Hardware

The ACM mapping reveals the knowledge areas that are important (e.g. 2 and 4 are extremely important). Table 4-6 below presents the total count for each knowledge area that is mapped in Table 4-5. A higher count indicates that the knowledge area is more important.

Table 4-6: Knowledge area ranking

Knowledge area No.	Count
4. System Integration and Architecture	10
2. Information Management	10
5. Human Computer Interaction	3
7. Platform Technologies	3
12. Math and Statistics for IT	2
13. Networking	1
3. Programming Fundamentals	1
6. Social and Professional Issues	1
9. Information Assurance and Security	1

A dilemma occurs when one attempts to propose that all these skills are prerequisites for a DW/BI module; that is because many of these skills are advanced IT skills not necessarily covered by general IT programmes at universities.

It was therefore decided to divide the list into general IT skills required on a level representative of a typical general IT degree, and an advanced or more specialised list of general IT skills used on an advanced or specialised level when applied in a DW/BI environment. Table 4-7 below presents some of the advanced general skills that are essential to DW/BI. Kimball *et al.* (2008d:475) refer to these skills in the DW/BI lifecycle. This table was compiled by evaluating the general skills that are discussed in detail by Kimball *et al.* (2008d:475).

Table 4-7: Advance general skills

Skill	Comment
Lit_1_Gen_ProjectM	According to Kimball <i>et al.</i> (2008d:475) project management skills is essential skill to the entire DW lifecycle.
Lit_2_Gen_Requirements	According to Kimball <i>et al.</i> (2008d:64) business analysis skills are key in requirements collection process.
Lit_3_Gen_TechArch_TechAchData	According to Kimball <i>et al.</i> (2008d:110) technical architecture skills is key for BI/DW technical architecture.
Lit_5_Gen_DBProg	According to Kimball <i>et al.</i> (2008d:327) DB programming skills are key to designing the physical database and planning for performance.
Lit_7_Gen_OLTP	According to Kimball <i>et al.</i> (2008d:370) data modelling skills are essential for ETL and DW/BI application development in a DW/BI solution.
Lit_9_Gen_Analytical	According to Kimball <i>et al.</i> (2008d:110) analytical skills are essential for requirements collection process.
Lit_10_Gen_ETL_Tools	According to Kimball <i>et al.</i> (2008d:110) knowledge of ETL tools is essential to especially. ETL tools may be used for data integration outside the field of DW/BI.
Lit_13_Gen_DBA	According to Kimball <i>et al.</i> (2008d:327) DBA skills are key to designing the physical database and planning for performance.
Lit_21_Gen_Interface	According to Kimball <i>et al.</i> (2008d:506) application programming is important for user front development and configuration.
Lit_23_Gen_BigData	The historical amounts of data accumulated thus far by many organisations and its current growth rate led to the coining of “Big data” term (Tankard & Pathways, 2012:5). This growth has given rise in volume (amount of data), velocity (speed of data) and variety (range of data).
Lit_24_Gen_DataMining	According to Kimball <i>et al.</i> (2008d:475) data mining is one of DW/BI application used to analyse data.
Lit_26_Gen_Metadata	According to Kimball <i>et al.</i> (2008d:64) metadata knowledge and handling is key in requirements collection process and the entire DW/BI system.
Lit_28_Gen_Testing	According to Kimball <i>et al.</i> (2008d:475) system testing and quality assurance skills are essential in the deployment phase of a DW/BI solution.
Lit_29_Gen_UserTraining	According to Kimball <i>et al.</i> (2008d:475) training skills are essential in the deployment phase of a DW/BI solution.
Lit_30_Gen_DataSecurity	According to Kimball <i>et al.</i> (2008d:327) information security skills are key to designing the physical database and planning for performance.

The list of skills from literature is now divided into two lists: One of general skills that may be seen as prerequisite skills; these are called “PreGen”. And one of general skills that are further developed through their application in DW/BI; these are called “AdvGen”. These lists are derived and renumbered in Table 4-8. All the advanced skills also appear on the prerequisite list; that is because some of the skills appear on both lists as basic knowledge required. However, advancement also occurs when applying it to DW/BI as indicated in Table 4-8.

Table 4-8: Development of prerequisite and advanced general IT skills

Original general Literature skill	Prerequisite general literature Skill	Advance general literature skill
Lit_1_Gen_ProjectM	Lit_1_PreGen_ProjectM For OLTP systems	Lit_1_AdvGen_ProjectM Developed to include DW/BI systems
Lit_2_Gen_Requirements	Lit_2_PreGen_Requirements For OLTP systems	Lit_2_AdvGen_Requirements Developed to include DW/BI systems
Lit_3_Gen_TechArchData	Lit_3_PreGen_TechArchData For OLTP systems	Lit_3_AdvGen_TechArchData Developed to include DW/BI systems
Lit_4_Gen_DBFunc	Lit_4_PreGen_DBFunc	
Lit_5_Gen_DBProg	Lit_5_PreGen_DBProg For ERD systems	Lit_4_AdvGen_DBProg Developed to include dimensional models
Lit_6_Gen_ERD	Lit_6_PreGen_ERD	
Lit_7_Gen_OLTP	Lit_7_PreGen_OLTP	Lit_5_AdvGen_OLTP To develop more flexible end-user systems
Lit_8_Gen_Processes	Lit_8_PreGen_Processes	
Lit_9_Gen_Analytical	Lit_9_PreGen_Analytical	Lit_6_AdvGen_Analytical To develop more flexible end-user systems
Lit_10_Gen_ETL_Tools	Lit_10_PreGen_ETL_Tools Basic data manipulation	Lit_7_AdvGen_ETL_Tools Advance data manipulation
Lit_11_Gen_DataStructures	Lit_11_PreGen_DataStructures	
Lit_12_Gen_DBConnections	Lit_12_PreGen_DBConnections	
Lit_13_Gen_	Lit_13_PreGen_DBA For ERD systems	Lit_8_AdvGen_DBA Developed to include dimensional

DBA		models
Lit_14_Gen_Prog	Lit_14_PreGen_Prog	
Lit_15_Gen_Reporting	Lit_15_PreGen_Reporting	
Lit_16_Gen_Source	Lit_16_PreGen_Source	
Lit_17_Gen_Business	Lit_17_PreGen_Business	
Lit_18_Gen_DBReporting	Lit_18_PreGen_DBReporting	
Lit_19_Gen_ReportTools	Lit_19_PreGen_ReportTools	
Lit_20_Gen_SDCL	Lit_20_PreGen_SDCL	
Lit_21_Gen_Interface	Lit_21_PreGen_Interface For OLTP systems	Lit_9_AdvGen_Interface To develop more flexible end-user systems
Lit_22_Gen_MSExcel	Lit_22_PreGen_MSExcel Including Pivot tables	
Lit_23_Gen_BigData	Lit_23_PReGen_BigData Only introductory	Lit_10_AdvGen_BigData Implications for DW/BI
Lit_24_Gen_DataMining	Lit_24_PreGen_DataMining Only introductory	Lit_11_AdvGen_DataMining As used from a DW
Lit_25_Gen_OLTP/BI	Lit_25_PreGen_OLTP/BI	
Lit_26_Gen_Metadata	Lit_26_PreGen_Metadata	Lit_12_AdvGen_Metadata Specific DW/BI metadata
Lit_27_Gen_Deployment	Lit_27_PreGen_Deployment	
Lit_28_Gen_Testing	Lit_28_PreGen_Testing	Lit_13_AdvGen_Testing Focussed in a DW/BI environment
Lit_29_Gen_UserTraining	Lit_29_PreGen_UserTraining	Lit_14_AdvGen_UserTraining Focussed in a DW/BI environment
Lit_30_Gen_DataSecurity	Lit_30_PreGen_DataSecurity	Lit_15_AdvGen_DataSecurity Focussed in a DW/BI environment
Lit_31_Gen_Networking	Lit_31_PreGen_Networking	
Lit_32_Gen_HardwareKnow	Lit_32_PeGen_HardwareKnow	

The *first proposed guideline* is that: students enrolling for a DW/BI module must have done modules that cover skills presented in the second column of Table 4-8.

The purposes of mapping the general IT skills to the ACM classification are to show that these skills are general, and to group them into known subject areas so that educators can compare degrees' content to determine entry requirements to DW/BI programmes. It is thus not necessary to map the specific DW/BI skills to the ACM list. Table 4-9 therefore contains only a relabelling of the DW/BI specific skills of Table 4-4.

Table 4-9: DW/BI specific skills/knowledge

No	Skills / knowledge (code number)
Lit_1_BI_Components	Lit_1_Concepts_Components
Lit_2_BI_Achitecture	Lit_2_Concepts_Achitecture
Lit_3_BI_DWlifecycle	Lit_3_Concepts_Lifecycle
Lit_4_BI_Methodologies	Lit_4_Concepts_Methodologies
Lit_5_BI_TechAchitect	Lit_7_TechAchitect
Lit_6_BI_ETL_Tools	Lit_19_ETL_Tools
Lit_7_BI_DimModeling	Lit_20_ETL_DimModeling Lit_30_Presentation_DimModeling
Lit_8_BI_MDX	Lit_32_Presentation_MDX
Lit_9_BI_OLAP	Lit_33_Presentation_OLAP
Lit_10_BI_MobileBI	Lit_43_Presentation_MobileBI

The *second proposed guideline* is that DW/BI modules should teach the skills that are presented in Table 4-9.

The *third proposed guideline* is that DW/BI students should develop advanced general IT skills indicated in the third column of Table 4-8 (named as “AdvGen”); these advanced general IT skills can be covered in modules taken concurrently by DW/BI students as part of their programme.

4.5 Summary

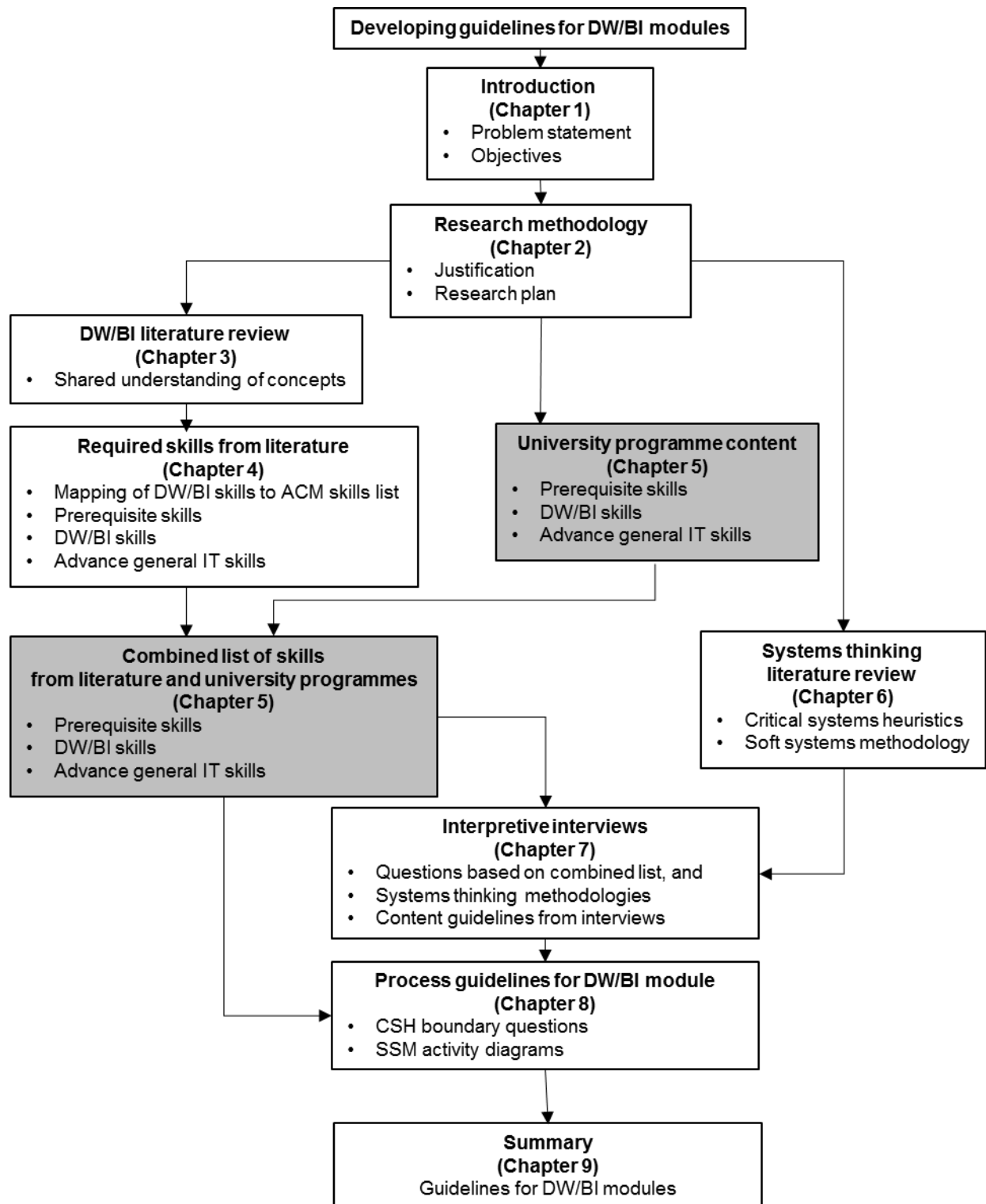
The aim of this chapter is to develop guidelines for DW/BI modules from DW/BI literature. In order to be able to position these guidelines in current IT programmes, the work of the ACM was used as a categorisation aid. This was done to ensure that the guidelines presented for prerequisite knowledge in this study can be linked to existing modules in IT programmes.

It is argued that DW/BI professionals combine their general IT knowledge with their DW/BI specific knowledge in their daily activities. Therefore, the list of skills required, as identified from the literature, was divided into general IT skills and DW/BI skills. It is further argued that the list of general IT skills contain IT skills typically acquired during the first three years of IT education, as well as advanced IT skills not included in the skill set of students who completed first (3 year) IT degrees.

Therefore this chapter provides three initial lists of skills as guidelines for DW/BI programmes:

1. A list of general IT skills students should have before enrolling for the DW/BI module presented in Table 4-8 (second column).
2. A list of DW/BI specific skills for the DW/BI module presented in Table 4-9.
3. A list of advanced general IT skills for the DW/BI module or concurrent modules presented Table 4-8 (third column).

The aim of the next chapter is to investigate current DW/BI programmes and modules at universities in order to enrich these lists.



Chapter 5 Guidelines from institutions of higher education

5.1 Introduction

The objective of this study is to develop guidelines for data warehousing/business intelligence modules in information technology programmes at universities; the guidelines are based on an understanding of the needs of various stakeholders. Two methodologies, i.e. critical systems heuristics and soft systems methodology, are used to understand the various stakeholders' perspectives.

The previous chapter (Chapter 4) presented an initial set of guidelines that was developed from DW/BI literature. Before investigating the views of different stakeholders, current university DW/BI programmes and modules are investigated in this chapter. This is done in order to provide context for the information provided by different stakeholders and to compare current programme/module information with DW/BI literature. The aim of this chapter is therefore to investigate DW/BI programmes in order to identify content guidelines and to consolidate these with the guidelines developed in the previous chapter. This is a twofold process as indicated by the two shaded blocks on the research plan diagram presented on the adjacent page.

The institutions of higher education create curricula as a standard academic “practice to help faculty members build programmes for specific disciplines” (Pyster *et al.*, 2009:94). These programmes are designed to provide students with a broad knowledge foundation (Reeves, 2015b). The purpose of IT programmes is to “provide graduates with the skills” they require to apply in the work place (after graduation) (Lunt *et al.*, 2008:10). The IT skill sets are acquired in general IT disciplines, i.e. Computer Engineering, Computer Science, Information Technology, Information Systems, and Software Engineering (Lunt *et al.*, 2008:16). These programmes equip graduates with the right combination of knowledge and practical hands-on experience.

The chapter consists of three main sections as explained in Section 5.2. Firstly, data collected from other universities' DW/BI programmes are used develop skill set lists for DW/BI modules. The process is reported in Section 5.3 and the final lists are given in Section 5.4. Secondly, these lists are compared with the skill set lists developed in the previous chapter from DW/BI literature (Section 5.5). Finally, a list of guidelines for DW/BI modules, representative of DW/BI literature and university programmes and modules, is presented in Section 5.6. A chapter summary is provided in 5.7.

5.2 Methodology

This section provides the methodology used to achieve the goals of this chapter:

1. Identify skills covered by universities' programmes and modules.
2. Compare the list of skills from DW/BI literature and skills from universities' programmes and modules.
3. Creating a combined list of skills representative of DW/BI literature and skills of universities' programmes and modules.

These steps are discussed in more detail in the following sections.

5.2.1 Identify skills according programmes/modules of institutions on higher learning

In order to identify skills from programmes/modules the following process is followed:

1. Select the top universities in South Africa (SA) and collect data on their applicable programme/modules (Section 5.3.1).
2. Present an overview of their general IT programme requirements (Section 5.3.2).
3. Focus on the DW/BI modules' curriculums to identify skills as guidelines (Section 5.3.3).

4. Identify *general* IT skills required by the DW/BI modules to those identified in literature (Section 5.4.1.)
5. Identify *specific* DW/BI skills from the DW/BI modules to those identified in literature in (Section 5.4.2.)

5.2.2 Compare the lists of skills from DW/BI literature and skills from universities' programmes and modules.

After the lists of general and specific IT skills identified from DW/BI programmes/modules from universities are generated they need to be compared with the lists developed in Chapter 4, compiled from DW/BI literature. There are two steps required in this process:

1. Compare *general* IT skills required by the DW/BI modules to those identified in literature (Section 5.5.1.)
2. Compare *specific* DW/BI skills from the DW/BI modules to those identified in literature in (Section 5.5.2)

5.2.3 Present a combined list of skills

It is argued in Chapter 4, Section 4.4.12 that the lists of general IT skills are on different academic levels. General IT skills are therefore divided into two groups. The first group is identified as skills that are typically covered in the module content of the first three years of IT programmes and are called prerequisite skills. The second group are advanced general IT skills, which consists of general IT skills not included in the skill set of students who completed a first (3 year) IT degree. This leads to the development of three sets of skills of DW/BI modules:

1. A list of prerequisite skills students should have before enrolling for the DW/BI module in Section 5.6.1.
2. A list of DW/BI specific skills for the DW/BI module in Section 5.6.2.
3. A list of advanced general IT skills students should master in Section 5.6.3.

This process outlined in the section starts with an investigation of the programme/module content of the universities.

5.3 Skills from universities' programme/module content

The aim of this section is to identify skills required for DW/BI modules from current university programmes and modules covering DW/BI content as discussed in Section 5.2.1. As this is part of the empirical objectives of the study information is given on data collection and analysis.

5.3.1 Data collection

First, it is necessary to identify the institutions that offer DW/BI modules. A selection of the top 12 universities in South Africa, according to University Ranking by academic performance (URAP), is used (performance, 2015). The list of universities that emerged as the top 12 is presented in Table 5-1.

Table 5-1: Top 12 universities in South Africa

No	Institution
1.	University of Cape town (UCT)
2.	University of the Witwatersrand (WITS)
3.	University of KwaZulu-Natal (UKZN)
4.	Stellenbosch University (SU)
5.	University of Pretoria (UP)
6.	University of Johannesburg (UJ)
7.	North West University (NWU)
8.	Rhodes University (RU)
9.	University of Free state (UFS)
10.	Nelson Mandela Metropolitan University (NMMU)
11.	University of the Western Cape (UWC)
12.	University of South Africa (UNISA)

Out of the top 12 universities only five universities offered a DW/BI module in calendar year 2015. Table 5-1 is expanded to include the '*DW/BI module*' indicator column. The indicator column contained the affirmative 'Yes' for universities that offer DW/BI module and a 'No' otherwise. The new table is presented in below in Table 5-2. This table does not suggest that other institutions without DW/BI modules do not offer DW/BI content in other modules; it means that they do not present a module called "Data Warehousing", or "Business Intelligence", or "Database Systems" where the latter includes data warehousing as a substantial part of the content. The table is ordered according to the '*DW/BI module*' column.

Table 5-2: Top 12 universities in South Africa

Institution	DW/BI module
University of Cape Town (UCT)	Yes
University of Pretoria (UP)	Yes
North West University (NWU)	Yes
University of Free state (UFS)	Yes
Nelson Mandela Metropolitan University (NMMU)	Yes
University of South Africa (UNISA)	No
University of the Witwatersrand (WITS)	No
University of KwaZulu-Natal (UKZN)	No
Rhodes University (RU)	No
University of the Western Cape (UWC)	No
University of Johannesburg (UJ)	No
Stellenbosch university (SU)	No

To collect data, a review of universities' prospectus/brochures, universities' websites, and study guide material was conducted. The researcher focused on two types of data:

1. General programme characteristics such as university name, degree name, programme focus, required modules to complete the degree, and admission requirements.
2. General requirements, DW/BI module content, the expected learning outcomes, and specific competencies taught in each module.

The next section focuses on the general programme characteristics.

5.3.2 General programme characteristics

This section focuses on the general programme characteristics of the identified universities. There are two aspects to be noted that are common among the identified universities' programmes. The first aspect is that DW/BI is not offered as a specialisation programme. Specialisation refers to the fact that DW/BI is not a stand-alone programme; instead, it is offered as a module within a degree programme. The second aspect is that the DW/BI module is only offered at honours level; so, an undergraduate degree with a general IT background is a minimum requirement for student admission. It is therefore important to look at the programme context that contains DW/BI as a module. A summary of the curriculums of the identified universities is presented in Table 5-3 below.

Table 5-3: DW/BI module content

Degree name	Admission requirement	Modules
UCT		
Honours in Computer Science	BSc Computer Science degree and must have 65% average in their final year of Computer Science.	Research methods and New venture planning are compulsory modules carrying 10 credits each. Other modules carry 10 credits each. Students are allowed to choose elective courses worth up to 80 credits from the list below: • Advanced Topics in Information Security • Computational Geometry for 3D Printing • <u>Database Systems</u> • Digital Forensics • Evolutionary Computation • High-Performance Computing • Human Computer Interaction

		• Information Retrieval • Introduction to Image Processing and Computer Vision • Introduction to ICT for Development • Logics for Artificial Intelligence • Mobile Game Development • Network and Internetwork Security • Networks for developing regions • Ontology Engineering • Software engineering • Visual Thinking and Visualization • User Experience in Games and Virtual Environments
UP		
BCom honours (Informatics)	IT degree other than BCom (e.g. BSc (IT), BIS, BSc (IS)),	A student should complete seven modules. At least five courses must be taken from the following modules: • Advance Database Systems (DW) • IS Development • Research Methodology • Managing end-users and projects • E-Commerce • Knowledge Acquisition & Sharing • Enterprise Architecture • Mobile Computing • HCI The remaining two modules may be taken in other departments.
NWU		
Honours in BSc Computer Science and Information Systems	A BSc degree with Information Technology, Computer Science or Business Analytics.	A student should complete a total of four modules each from the list on level I and level II and a research project. Also the database modules must be coupled with DW module within the same calendar year. • Data warehouses I • Databases • Information Systems Engineering I • Computer Security I • Artificial Intelligence I • Decision Support Systems I • Project • Data Warehouses I • Databases II • Information Systems Engineering II • Decision Support Systems II
UFS		
Honours in Computer Science	Three-year B-degree majoring in Computer Science or Computer Information Systems	Students should take the following compulsory modules: • Project • Advance programming 1 • Advance programming 2 • Introduction to research Students should take at least five elective modules from the

		following :Network management • Advance databases • Robotics • Information security • Theory of algorithms • Automata theory and applications • Business intelligence • Data warehousing • Knowledge based systems • Capita Selecta • Advance internet programming • Management information systems • IT project management • Decision support systems • Human computer interaction • Advance computer networks • Object designs • Computer ethics
NMMU		
BSc Honours (Computer Science)	Bachelor's degree (Computer Science)	The Honours course consists of at least eight semester lecture modules with a total credit value of at least 88 and a treatise on an independent project. A two-semester module contributes 22 credits, a one-semester module 11 credits and the treatise on the project 32 credits. A total of at least 120 credits are required for the learning programme. Students must select at least four of the following compulsory modules: • Graphics • Compiler Construction • Usability Engineering • Evolutionary Computing • Virtual Reality • Advanced Programming At most 33 credits selected from the following modules must be offered Students have an option of choosing from the following subjects. • Design in the Digital Domain • Data Warehousing • E-Commerce • Research Frontiers in Computing • Capita Selecta • Mobile Computing Another Honours module which must be approved by the HoD of Computing Science, subject to the condition that it should complement the other modules in the programme. Approval is dependent upon submission of request on appropriate form available in the Department. (Credit vale = 22)

Modules that include DW/BI are shaded in Table 5-3. UTC and UP do not have a module that is explicitly referred to as data warehousing; it is included in the databases module (the database modules of all other universities were also looked at). Table 5-3 illustrates that DW modules are coupled with other general IT modules. Students are given a wide range of modules that they can choose from to couple with data warehousing. Some universities offer more modules than others do. Table 5-4 below presents the source for the module content as per Table 5-3 above.

Table 5-4: Module content sources reference table

No	Institution	Source
1.	UCT	University of Cape Town. 2015. Honours handbook. http://www.cs.uct.ac.za/honours/Honours_Handbook_2015.pdf. Available on line. Accessed 20 September 2015.
2.	UP	- University of Pretoria. 2015. BCOM hon (informatics) handbook. http://www.up.ac.za/media/shared/102/ZP_Files/hons-2015v2_1.zp39704.pdf. Available on line. Accessed 20 September 2015. - Data warehousing fundamentals for it professionals, 2nd Edition. Paulraj Ponniah
3.	NWU	- North West University. 2015. Information sheet: Honours students 2015. http://www.nwu.ac.za/sites/www.nwu.ac.za/files/files/v-fesit-sit/it%20courses/Final%20Information%20Sheet%202015.pdf. Accessed 20 September 2015. - The Microsoft data warehouse toolkit. With SQL Server R2 and the Microsoft Business Intelligence Toolset. 2nd ed. Wiley. Mundy, J., Thornthwaite, W. & Kimball R. 2008. - The data warehouse toolkit. Practical techniques for building data warehouse and Business Intelligence Systems. 2nd ed. Wiley. Kimball, R., Ross, M., Thornthwaite, W., Mundy, J., Becker, B. 2008. - Study guide: ITRI611 - Study guide: ITRI621
4.	UFS	- University of Free State. 2015. Rule book 2015. http://apps.ufs.ac.za/dl/yearbooks/237_yearbook_eng.pdf. Accessed 20 September 2015. - Business Intelligence Practices, Technologies and Management (1st edition). Rajiv Sabherwal & Irma Becerra-Fernandez. (2011). Wiley. - Microsoft. (2012). SQL Server 2012 Tutorials: Analysis Services – Multidimensional Modeling. Step-by-Step - Sharda, R., Delen, D., & Turban, E. (2013). Business Intelligence: A Managerial Perspective on Analytics (3 edition). Prentice Hall. - Study guide: CSID6823 (RIS610) - Study guide: CSID6853 (RIS625)
5.	NMMU	- Computer science. 2015. Post graduate modules. http://cs.nmmu.ac.za/Postgraduate-Information/Postgraduate-Modules. Accessed 20 September 2015. - Study guide: WRDW411 Data Warehousing 411

The next section focuses on the DW module characteristics.

5.3.3 Data warehouse module characteristics

The objective of this section is to investigate the DW/BI modules and analyse their characteristics. The researcher focuses on general DW/BI module prerequisites, content, and expected learning outcomes.

5.3.3.1 Prerequisites

There are requirements that students must satisfy before he/she is admitted to the honours programmes' DW/BI modules. Students must have completed a three year undergraduate degree; students must also have completed a database module or any other introductory module on the design and use of a relational database (refer to Table 5-2). Undergraduate degrees afford students general IT fundamentals.

5.3.3.2 Module content

This section presents the DW/BI module content, the expected core learning outcomes, and specific competencies taught at each university. The expected learning outcomes describe what faculty members expect students to know, and to be able to do, at the end of the module. Table 5-5 below presents a summary of what each university includes in the DW/BI module. The '*Module*' column gives the descriptive name of the DW/BI module. The '*Source*' column gives the list of prescribed books, material, or study guide that is used by the university. The '*Source*' column gives the reference material (prescribed books/materials and study guides) for the '*Module content/Learning outcomes*' column. These sources are used to identify skill sets/knowledge at the same level of detail. The following scheme is used for labelling the *university name* in Table 5-5:

Sequence number + institution + 'Content'

Sequential number = Count number (starting from 1 up to the maximum number of universities).

Institution = University name (abbreviation or short name)

For example, the first content displayed in Table 5-5 is for UCT, and its number is 1_UCT_Content. The table also contains a code assigned to the module content. Codes are repeated as they reoccur. The following coding scheme is used:

Uni + sequential number + group + skill short name

Uni = Indicates that skill is from university curriculum
Sequential number = Count number starting from 1 up to the maximum number of identified skills.
Group = Which group does the skill fall under (General or DW/BI specific)
Skill short name = Skill / knowledge

Table 5-5: DW/BI module content

Module	Module content/Learning outcomes	Code
1 _UCT_ Content		
Database Systems	Object-relational systems	Uni_1_Gen_Analytical
	NoSQL	Uni_4_Gen_DBProg
	Distributed databases	Uni_4_Gen_DBProg
	Data warehousing	Uni_7_Gen_Requirements Uni_9_Gen_DBReporting Uni_10_Gen_OLTP_BI Uni_1_BI_DWPrinciples Uni_2_BI_Architecture Uni_3_BI_DWLifecycle
	OLAP and data mining	Uni_3_Gen_DataMining Uni_7_BI_Tools
	Spatial data management	Uni_6_Gen_SpatialData
	Big Data	Uni_2_Gen_BigData
	Research topics in database systems	
2 __UP_ Content		
Advanced database systems	Data warehousing theory	Uni_10_Gen_OLTP_BI Uni_8_BI_KimballMethod
	Data warehousing principles	Uni_7_Gen_Requirements Uni_1_BI_DWPrinciples Uni_3_BI_DWLifecycle

	Data warehouse architecture	Uni_2_BI_Architecture
	Dimensional modelling	Uni_4_BI_DimModelling
	Extract, load and transform (ETL) functions	Uni_4_Gen_DBProg Uni_7_BI_Tools
	Practical DW application	Uni_1_Gen_Analytical
	OLAP	Uni_6_BI_OLAP Uni_11_BI_MDX
	DW reporting	Uni_9_Gen_DBReporting Uni_5_BI_DWReporting
3 _NWU_ Content		
Data warehousing	General data warehousing concepts	Uni_10_Gen_OLTP_BI Uni_1_BI_DWPrinciples
	Data warehouse methodology	Uni_7_Gen_Requirements Uni_2_BI_Architecture
	Programming (graphic user interface design)	Uni_5_Gen_InterProg
	Data warehouse lifecycle	Uni_3_BI_DWLifecycle Uni_8_BI_KimballMethod
	Alternative data warehousing methodologies	Uni_10_BI_Methodologies
	Dimensional modelling	Uni_4_BI_DimModelling
	Requirements collection	
	Extract, load and transform (ETL) functions	Uni_4_Gen_DBProg Uni_7_BI_Tools
	Create a data warehouse (star schema) from given data and requirements.	Uni_1_Gen_Analytical Uni_4_Gen_DBProg
	Create a data warehouse browser, reports and dashboard from an OLAP cube.	Uni_4_Gen_DBProg Uni_9_Gen_DBReporting Uni_5_BI_DWReporting Uni_6_BI_OLAP Uni_11_BI_MDX
	Suitable documentation should be developed for the practical work.	Uni_8_Gen_ProjDocumentaion
4 _UFS_ Content		
Data Warehousing	Understand of data warehouse development methodology.	Uni_7_Gen_Requirements Uni_10_Gen_OLTP_BI Uni_1_BI_DWPrinciples Uni_10_BI_Methodologies
	Understand and apply Ralph Kimball's methodology.	Uni_2_BI_Architecture Uni_3_BI_DWLifecycle Uni_8_BI_KimballMethod
	Learn the fundamentals of Big Data.	Uni_2_Gen_BigData

	Use ETL concepts including slow changing dimensions for transforming and loading of data into a data warehouse.	Uni_4_Gen_DBProg Uni_4_BI_DimModelling
	Develop all the necessary ETL routines using SQL Server SSIS.	Uni_1_Gen_Analytical Uni_4_Gen_DBProg Uni_7_BI_Tools
	Analyse data from a data warehouse.	
	Create an OLAP database.	Uni_4_Gen_DBProg Uni_6_BI_OLAP
	Create data mining models.	Uni_3_Gen_DataMining
	Development of key performance indicators (KPIs), dashboards and balanced scorecards.	Uni_9_Gen_DBReporting Uni_5_BI_DWReporting
	Understanding machine learning algorithms namely; clustering (k-means), decision trees and artificial neural networks.	Uni_3_Gen_DataMining
	Perform hands-on exercises with a Business Intelligence tool (Microsoft Analysis Services), Data Visualization tool (Tableau) and Data Mining tool (SAS Enterprise Miner).	Uni_1_Gen_Analytical Uni_3_Gen_DataMining Uni_5_BI_DWReporting Uni_7_BI_Tools
	Deliver a Business Intelligence solution. This will entail the following: • Compile business requirements from a business. • Design the dimensional model from the business requirement specification. • Load the dimensional model with data using ETL routines. • Develop and deploy OLAP cubes. • Include Key Performance Indicators (KPI) in the OLAP cubes.	Uni_1_Gen_Analytical Uni_4_Gen_DBProg Uni_7_Gen_Requirements Uni_9_Gen_DBReporting Uni_4_BI_DimModelling Uni_5_BI_DWReporting Uni_6_BI_OLAP Uni_7_BI_Tools Uni_11_BI_MDX
	Delivery the information to end-users in a dashboard and scorecard.	Uni_9_Gen_DBReporting Uni_5_BI_DWReporting Uni_11_BI_MDX
5_NMMU_Content		
Data Warehousing	Data warehousing link to data mining.	Uni_1_Gen_Analytical Uni_3_Gen_DataMining Uni_1_BI_DWPrinciples
	The process and importance of data pre-processing.	Uni_3_Gen_DataMining
	Data warehousing principles	Uni_7_Gen_Requirements Uni_10_Gen_OLTP_BI Uni_2_BI_Architecture Uni_8_BI_KimballMethod
	Dimensional modelling	Uni_4_Gen_DBProg Uni_4_BI_DimModelling

	OLAP Technology	Uni_5_BI_DWReporting Uni_6_BI_OLAP Uni_11_BI_MDX
	Data cube computation and usage and data generalization.	Uni_4_Gen_DBProg Uni_7_BI_Tools Uni_11_BI_MDX
	Data mining of frequent patterns, associations and correlations.	Uni_3_Gen_Datamining
	Classification and prediction	Uni_3_Gen_DataMining
	Cluster analysis	Uni_3_Gen_DataMining
	Data visualization	Uni_3_Gen_DataMining Uni_9_Gen_DBReporting

Table 5-5 presents a general curriculum of what is included in the DW/BI module. The source material for the module content is not on the same level of detail and scope per university. Therefore, there might be some topics that are covered in the module content that are not included in Table 5-5. The extent (scope and depth) to which these skill sets is covered by the universities is not clear from the general curriculum; a further investigation of the detailed content is outside the scope of this study. It is assumed that the terminology used to identify the broad topics, relates to popular DW/BI literature as presented in the previous chapter.

These skills however can be seen as a mixture of general IT skills and DW/BI specific skills as presented in the next section.

5.4 Findings: Skills from universities' programmes and modules

This section presents the skills from universities' programmes and modules in terms of general IT skills from universities' programmes and modules (Section 5.4.1) and specific DW/BI skills from universities' programmes and modules (Section 5.4.2).

5.4.1 General IT skills

General IT skills extracted from module contents are presented in this section. These skills are mapped to the body of knowledge to qualify them as general IT skills. Table 5-6 below has four columns. The first column 'No' is for skills numbering, and a

coding numbering is also implemented in order to identify the source of the skill later in the chapter. The second column 'Skills/knowledge' gives the list of skills. The third column 'Institution' is a list of institutions that offers the skills. The fourth column 'ACM knowledge area' represents the knowledge area number and the item matching the skill. If the skill has a mapping number (knowledge area) it thus means that it falls under a general skill category. This scheme is applied to both the general (Gen) and business intelligence (BI) skills tables.

Table 5-6: General skills from module contents

No	Skills/knowledge	Institution	ACM knowledge area
Uni_1_Gen_Analytical	Analytical skills	1_UCT_Content 2_UP_Content 3_NWU_Content 4_UFS_Content 5_NMMU_Content	12. Math and Statistics for IT
			- Basic Logic - Discrete Probability - Functions, Relations and Sets - Hypothesis Testing - Sampling and Descriptive Statistics - Graphs and Trees - Application of Math & Statistics to IT
Uni_2_Gen_DW/BigData	Big Data	- Uni_1_Content_UCT - Uni_4_Content_UFS	2. Information Management
			Managing the Database Environment Special-Purpose Databases
Uni_3_Gen_DataMining	Data mining skills	- Uni_1_Content_UCT - Uni_4_Content_UFS - Uni_5_Content_NMMU	2. Information Management
			- Managing the Database Environment - Special-Purpose Databases
Uni_4_Gen_DBProg	Database programming (e.g. SQL, objects, NoSQL)	1_UCT_Content 2_UP_Content 3_NWU_Content 4_UFS_Content 5_NMMU_Content	2. Information Management
			- Database Query Languages - Data Organization Architecture - Data Modelling
Uni_5_Gen_InterProg	Interface programming skills	3_NWU_Content	3 Programming Fundamentals
			- Fundamental Programming Constructs - Object-Oriented Programming

			- Algorithms and Problem-Solving - Event-Driven Programming
Uni_6_Gen_SpatialData	Spatial data management	1_UCT_Content	2 Information Management
			- Database Query Languages - Data Organization Architecture - Data Modelling
Uni_7_Gen_Requirements	Requirements collection	1_UCT_Content 2_UP_Content 3_NWU_Content 4_UFS_Content 5_NMMU_Content	4 System Integration and Architecture
			- Requirements - Organizational Context - Architecture
Uni_8_Gen_ProjDocumentation	Project Documentation	3_NWU_Content	4 System Integration and Architecture
			- Requirements - Organizational Context - Architecture
Uni_9_Gen_DBReporting	Database reporting	1_UCT_Content 2_UP_Content 3_NWU_Content 4_UFS_Content 5_NMMU_Content	2 Information Management
			Database Query Languages
Uni_10_Gen_OLTP_BI	Knowledge of the difference between OLTP and DW/BI.	1_UCT_Content 2_UP_Content 3_NWU_Content 4_UFS_Content 5_NMMU_Content	2 Information Management
			- Managing the Database Environment - Special-Purpose Databases

Even though Table 5-6 presents the list of skills as general, there are some skills that are not necessarily general; rather, these are representative of very specific disciplines. For the purpose of this study these are listed as general because they are not specific to DW/BI. For example, data mining, spatial data management, and big data are focus areas on their own. Some universities combine DW/BI with data mining; however, data mining skills and knowledge is not necessarily core to DW/BI.

The general IT skills presented in Table 5-5 highlights the argument presented in Section 4.4.12 that DW/BI knowledge requires an advance level of general IT knowledge.

5.4.2 DW/BI specific skills from module contents

BI specific skills from module contents are presented in Table 5-7 below. The prefixes “Uni_No_BI_” in table Table 5-7 are used later in the chapter to identify these skills as DW/BI specific skills identified from university programmes.

Table 5-7: DW/BI specific skills/knowledge from module contents

No	Skills/knowledge	Institution
Uni_1_BI_DWPrinciples	DW principles	1_UCT_Content 2_UP_Content 3_NWU_Content 4_UFS_Content 5_NMM_Content
Uni_2_BI_Architecture	DW architecture	1_UCT_Content 2_UP_Content 3_NWU_Content 4_UFS_Content 5_NMM_Content
Uni_3_BI_DWLifecycle	DW lifecycle extract, load and transform (ETL) functions	1_UCT_Content 2_UP_Content 3_NWU_Content 4_UFS_Content 5_NMM_Content
Uni_4_BI_DimModelling	Dimensional modelling	2_UP_Content 3_NWU_Content 4_UFS_Content 5_NMM_Content
Uni_5_BI_DWReporting	DW reporting (e.g. standard reports, development of Key performance indicators (KPIs), dashboards and balanced scorecards)	2_UP_Content 3_NWU_Content 4_UFS_Content 5_NMM_Content
Uni_6_BI_OLAP	Online analytical processing (OLAP)	2_UP_Content 3_NWU_Content 4_UFS_Content 5_NMM_Content

Uni_7_BI_Tools	BI tools knowledge	1_UCT_Content 2_UP_Content 3_NWU_Content 4_UFS_Content 5_NMMU_Content
Uni_8_BI_KimballMethod	DW Kimball methodology knowledge	2_UP_Content 3_NWU_Content 4_UFS_Content 5_NMM_Content
Uni_10_BI_Methodologies	Knowledge of alternative DW Methodologies	3_NWU_Content 4_UFS_Content
Uni_11_BI_MDX	Multidimensional dynamics expression (MDX) skills	2_UP_Content 3_NWU_Content 4_UFS_Content 5_NMM_Content

The set of skills presented in Table 5-7 is compared with the DW/BI skills from literature in the next section.

5.5 Comparison of university skills and skills from literature

The aim of this section is to provide a traceable comparison between the skills for DW/BI modules developed in Chapter 4 to the skills identified from university programmes in the previous section. This comparison is done in terms of general IT skills (Section 5.5.1), and specific DW/BI skills (Section 5.5.2).

5.5.1 Comparison of advance general IT skills from literature and skills from universities

The researcher compares advance general IT skills identified from literature with general IT skills identified universities' programmes. The purpose of this process is to determine skills that are not included in universities' curriculums, but are otherwise mentioned in literature and *vice versa*. As it was highlighted previously, these skills might be covered by other programmes. Also, the extent (scope and depth) to which the skill sets/knowledge is included by the universities is not clear from the general

curriculums. What is also evident is that the literature guidelines are more detailed than the universities' guidelines. The literature's advanced general IT skills mapped to universities' skills are presented in Table 5-8 below.

Table 5-8: Advanced general IT skills mapped to universities' skills

Literature skills	University skills
Lit_1_AdvGen_ProjectM	Uni_8_Gen_ProjDocumentaion
Lit_2_AdvGen_Requirements	Uni_7_Gen_Requirements
Lit_3_AdvGen_TechArchData	
Lit_4_AdvGen_DBProg	Uni_4_Gen_DBProg
Lit_5_AdvGen_OLTP	
Lit_6_AdvGen_Analytical	Uni_1_Gen_Analytical
Lit_7_AdvGen_ETL_Tools	
Lit_8_AdvGen_DBA	
Lit_9_AdvGen_Interface	Uni_5_Gen_InterProg
Lit_10_AdvGen_BigData	Uni_2_Gen_BigData
Lit_11_AdvGen_DataMining	Uni_3_Gen_DataMining
Lit_12_AdvGen_Metadata	
Lit_13_AdvGen_Testing	
Lit_14_AdvGen_UserTraining	
Lit_15_AdvGen_DataSecurity	
	Uni_6_Gen_SpatialData
NEW: Lit_16_AdvGen_DBReporting	Uni_9_Gen_DBReporting
NEW: Lit_17_OLTP_BI	Uni_10_Gen_OLTP_BI

One should be careful of making conclusions about the empty cells in the second column. These skills are not explicitly named in universities' modules' content; however, they may be presented as part of other items listed in the module's documentation. In Table 5-8 there are three universities' skills that are not mapped to literature skills. The knowledge represented by these skills is presented in general DW/BI literature, with exception of spatial data (only from Uni_1_UCT), which is present in applied DW/BI literature. It is therefore appropriate to extend the list of advanced general skills to include the two general skills that are prominent in universities' programmes as indicated in the last two rows of Table 5-8.

5.5.2 Comparison of DW/BI specific skills from literature and skills from universities

Table 5-9 compares the second guidelines according to literature and universities' curriculums.

Table 5-9: DW/BI specific skills mapped to universities' skills

Skills from Literature	Skills from universities' programmes
Lit_1_BI_Components	Uni_1_BI_DWPrinciples
Lit_2_BI_Achitecture	Uni_2_BI_Architecture
Lit_3_BI_DWlifecycle	Uni_3_BI_DWLifecycle
Lit_4_BI_Methodologies	Uni_10_BI_Methodologies Uni_8_BI_KimballMethod
Lit_5_BI_TechAchitect	Uni_2_BI_Architecture
Lit_6_BI_ETL_Tools	Uni_7_BI_Tools
Lit_7_BI_DimModeling	Uni_4_BI_DimModelling
Lit_8_BI_MDX	Uni_11_BI_MDX
Lit_9_BI_OLAP	Uni_6_BI_OLAP
Lit_10_BI_MobileBI	

Mobile DW/BI (Lit_10_BI_Mobile DW/BI) is not mapped and therefore not explicitly covered by universities. Mobile DW/BI is a new trend where DW/BI solutions are delivered to handset devices. All the identified DW/BI specific skills from universities are covered in the list from literature.

5.6 Guidelines for DW/BI programmes from literature and universities' programmes

This section presents a combined summary of skills/knowledge from the literature review and universities' programmes.

5.6.1 Prerequisite skills from literature and universities' programmes

The *first proposed guideline* is that: students enrolling for a DW/BI module must have done modules covering the skills presented in Table 5-10 below; these are prerequisite skills (this table is an excerpt of Table 4-8 in Section 4.4.11).

Table 5-10: Development of prerequisite and advanced general IT skills

Prerequisite general literature skill	Description
Lit_1_PreGen_ProjectM	Project management of OLTP systems including project management of technical systems such as ETL.
Lit_2_PreGen_Requirements	Requirements collection and analysis for OLTP systems and end-user applications focussed on reporting.
Lit_3_PreGen_TechArchData	Data management and technical architecture for OLTP systems based on ERD data models.
Lit_4_PreGen_DBFunc	Database concepts and functionality of relational databases.
Lit_5_PreGen_DBProg	Database programming of relational databases.
Lit_6_PreGen_ERD	Database modelling of relational data in ERD models.
Lit_7_PreGen_OLTP	Overall modelling and management of data in OLTP systems.
Lit_8_PreGen_Processes	Data flow and other operational processes in OLTP systems.
Lit_9_PreGen_Analytical	Analytical skills in terms of all components of OLTP systems including requirements analysis and technical analysis.
Lit_10_PreGen_ETL_Tools	Programming skills to do basic data manipulation.
Lit_11_PreGen_DataStructures	Knowledge and manipulation skills of general data structures such as objects, link lists and arrays.
Lit_12_PreGen_DBConnections	Creation of connections between different types of databases.
Lit_13_PreGen_DBA	DBA skills for relational databases.
Lit_14_PreGen_Prog	General programming skills.
Lit_15_PreGen_Reporting	General reporting skills.
Lit_16_PreGen_Source	Knowledge of different source systems and platforms including ERP systems.
Lit_17_PreGen_Business	Knowledge of typical business processes.
Lit_18_PreGen_DBReporting	Reporting linked to database systems.
Lit_19_PreGen_ReportTools	Use of reporting tools.
Lit_20_PreGen_SDLC	Good knowledge of SDLC of OLTP systems.
Lit_21_PreGen_Interface	Development of user interfaces for OLTP systems.

Lit_22_PreGen_MSEExcel	Microsoft Excel knowledge including pivot tables.
Lit_23_PReGen_BigData	Introductory knowledge of Big Data and unstructured data.
Lit_24_PreGen_DataMining	Introductory knowledge of Data Mining.
Lit_25_PreGen_OLTP/BI	Knowledge of the comparison of OLTP and DW/BI systems.
Lit_26_PreGen_Metadata	Knowledge of metadata.
Lit_27_PreGen_Deployment	Knowledge of the deployment of new systems.
Lit_28_PreGen_Testing	Good testing knowledge.
Lit_29_PreGen_UserTraining	Knowledge of user training and support.
Lit_30_PreGen_DataSecurity	Awareness of security risks and procedures.
Lit_31_PreGen_Networking	Knowledge and skills of network topologies and their implementation.
Lit_32_PeGen_HardwareKnow	Knowledge of hardware systems such as distributed computing.

5.6.2 BI specific skills from literature and universities' programmes

The *second proposed guideline* is that DW/BI modules should cover the DW/BI specific skills as presented in Table 5-11 below.

Table 5-11: Second guideline DW/BI specific skills

Skill	Description
Lit_1_BI_Components	BI components knowledge (Data, technology, processes, and people).
Lit_2_BI_Achitecture	DW architecture knowledge and understanding.
Lit_3_BI_DWlifecycle	DW lifecycle extract, load and transform.
Lit_4_BI_Methodologies	Knowledge of alternative DW Methodologies.
Lit_5_BI_TechAchitect	Knowledge of the technical architecture of a DW.
Lit_6_BI_ETL_Tools	Vendor specific tools that enable better ETL processing.
Lit_7_BI_DimModeling	Dimensional modelling.
Lit_8_BI_MDX	Multidimensional dynamics expression (MDX) skills.
Lit_9_BI_OLAP	Knowledge of online analytical processing (OLAP) cubes.
Lit_10_BI_MobileBI	Mobile DW/BI.

5.6.3 Advanced general IT skills from literature and universities' programmes

The DW/BI skills incorporate specific advanced general IT skills presented in Table 5-12. These skills should be addressed as part of a DW/BI module.

Table 5-12: Advanced general skills

Literature Skills	Description
Lit_1_AdvGen_ProjectM	According to Kimball <i>et al.</i> (2008d:339) project management skills is essential skill to the entire DW lifecycle.
Lit_2_AdvGen_Requirements	According to Kimball <i>et al.</i> (2008d:64) business analysis skills are key in requirements collection process. This should include skills to develop a requirements bus matrix.
Lit_3_AdvGen_TechArchData	According to Kimball <i>et al.</i> (2008d:110) technical architecture skills is key for BI/DW technical architecture.
Lit_4_AdvGen_DBProg	According to Kimball <i>et al.</i> (2008d:327) DB programming skills are key to designing the physical database and planning for performance.
Lit_5_AdvGen_OLTP	According to Kimball <i>et al.</i> (2008d:370) data modelling skills are essential for ETL and DW/BI application development in a DW/BI solution.
Lit_6_AdvGen_Analytical	According to Kimball <i>et al.</i> (2008d:348) analytical skills are essential for requirements collection process.
Lit_7_AdvGen_ETL_Tools	According to Kimball <i>et al.</i> (2008d:110) knowledge of DW/BI tools is essential to especially.
Lit_8_AdvGen_DBA	According to Kimball <i>et al.</i> (2008d:327) DBA skills are key to designing the physical database and planning for performance.
Lit_9_AdvGen_Interface	According to Kimball <i>et al.</i> (2008d:506) application programming is important for user front development and configuration.
Lit_10_AdvGen_BigData	The historical amounts of data accumulated thus far by many organisations and its current growth rate led to the coining of "Big data" term (Tankard & Pathways, 2012:5). This growth has given rise in volume (amount of data), velocity (speed of data) and variety (range of data).
Lit_11_AdvGen_DataMining	According to Kimball <i>et al.</i> (2008d:475) data mining is one of DW/BI application used to analyse data.
Lit_12_AdvGen_Metadata	According to Kimball <i>et al.</i> (2008d:64) metadata knowledge and handling is key in requirements collection process and the entire DW/BI system.
Lit_13_AdvGen_Testing	According to Kimball <i>et al.</i> (2008d:364) system testing and quality assurance skills are essential in the deployment phase of a DW/BI solution.
Lit_14_AdvGen_UserTraining	According to Kimball <i>et al.</i> (2008d:15) training skills are essential in the deployment phase of a DW/BI solution.

Lit_15_AdvGen_DataSecurity	According to Kimball <i>et al.</i> (2008d:327) information security skills are key to designing the physical database and planning for performance.
Lit_16_AdvGen_DBReporting	Advance reporting skills are required to enable the development of flexible reports required by DW/BI applications.
Lit_17_OLTP_BI	Advance understanding of differences between OLTP and DW/BI systems.

The chapter is summarised in the next section.

5.7 Summary

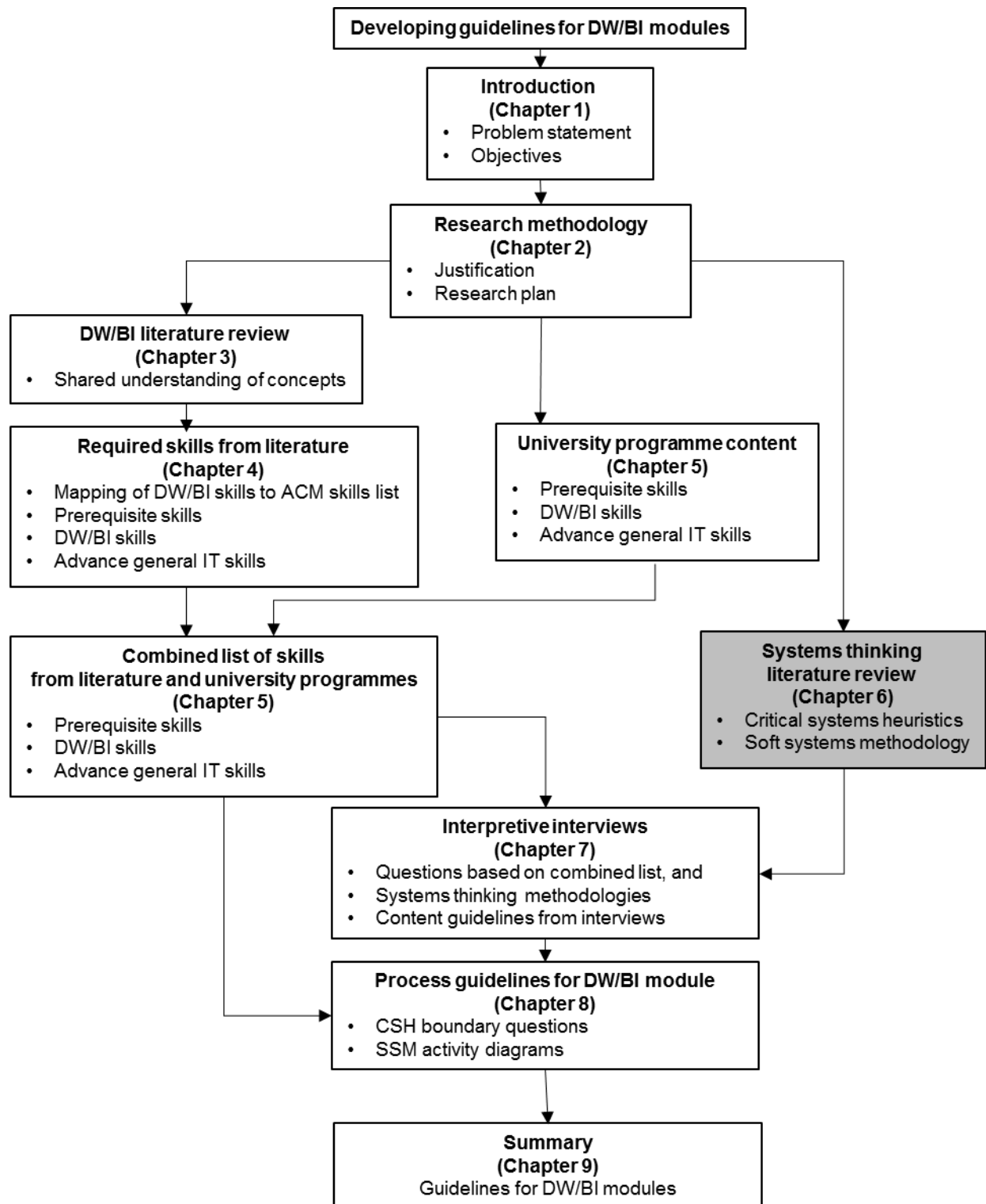
The purpose of this chapter is to focus on what expertise is (or ought to be) possessed by DW/BI professionals, or what counts (should count) as relevant knowledge in this field, as investigated in current DW/BI programmes/modules at South African universities.

This chapter has three objectives:

1. Conduct a study to understand what are currently being presented in universities' DW/BI programmes. A document investigation of DW/BI programmes at identified institutions of higher learning was performed. The researcher focused on three types of data: general programme characteristics, general module content, and specific competencies taught in each module.
2. Compare the skills identified from DW/BI modules with the skills identified in literature in Chapter 4.
3. Create guidelines for DW/BI modules representative of both DW/BI literature and university programme/module content. Three sets of skills contained in three guidelines are presented in this chapter. The first guideline identifies skills that students enrolling for DW/BI modules must have as prerequisite – listed in Table 5-10. The second guideline is a list of specific DW/BI skills that was extracted from the literature, and supported by universities' programmes

– presented in Table 5-11. The third guideline is a list of advanced general IT skills that students must achieve in order to be successful in DW/BI – provided in Table 5-12.

These guidelines are extended in the next chapter to include requirements according to former students and data warehousing professionals. Since systems thinking methodologies are used to guide the process of stakeholder involvement, a literature review is first provided in order to create a shared understanding of these concepts.



Chapter 6 Systems thinking methodologies

6.1 Introduction

The objective of this study is to develop guidelines for data warehousing/business intelligence modules in information technology programmes at universities; the guidelines are based on an understanding of the needs of various stakeholders. Two methodologies, i.e. critical systems heuristics and soft systems methodology, are used to understand the various stakeholders' perspectives. A systems approach is recommended to determine how, in this changing environment of business intelligence education practice, we can improve our understanding of it.

Methods such as linear thinking and analysis, which are processes of deconstructing a topic or problem to study the smaller pieces, were mostly used in the past (Bartlett, 2001:2, Ulrich, 1996). These traditional analysis methods focus on separating the subject matter (problem) into individual parts in order to study it (Jackson, 2001:233). However, traditional analysis methods do not work so well on complex problems. Checkland (2011:497) and Ulrich (2003:1) are systems thinkers – Checkland developed soft systems methodology to accommodate subjectivity; and Ulrich developed critical systems heuristics to help the disadvantaged in situations involving conflict and imbalances. Rather than isolating smaller parts of the system being studied, systems thinking aims to expand the view by taking into account larger numbers of interactions of the problem that is being studied. These complex realities of life, where nothing stays the same, necessitated better approaches to structure and address them. A great number of interactions, and the absence or ineffectiveness of immediately observable solutions, requires a systems approach. Systems thinking provides a means to understand the whole because it is a framework for seeing interrelationships rather than things (Senge, 1990:10).

Soft systems methodology and critical systems heuristics are the two systems methodologies studied in detail in this chapter. This chapter is organised as follows: Section 6.2 discusses systems and systems thinking. An illustrative business intelligence education problem situation is presented in Section 6.3. Section 6.4

discusses what soft systems methodology is as well as how and why it is used. Critical systems thinking methodology is presented in in Section 6.5. Section 6.6 discusses what critical systems heuristics is and how and why it is used. Lastly, Section 6.7 is a chapter summary.

6.2 Systems and systems thinking

An understanding of what constitutes a system forms the basis of systems thinking. Checkland (1997a:667) describes the system as a whole that must be viewed as having properties that makes it more than the sum of its parts. Ackoff (1971:662) defines a system as “a set of interrelated elements”. Isolating or reducing complexity into small manageable components has always been an accepted approach (Dettmer, 2006:1). This meant considering each component in isolation from the others (Pan *et al.*, 2013:1028). However, in reality, there are always some degree of interdependencies between components that forms a whole (Chan, 2001c).

The behaviour of individual components is important to the whole system; so, determining and mapping the relationship between components is very important in a system (Ackoff, 1971:661). Here, behaviour refers to how individual processes accomplish their functions. When many elements and computations are present, there are to some extent interrelationships, interaction, and interconnectivity between the elements of the system (Chan, 2001b:1). There is a direct or indirect connection between components of a system Ackoff (1971:662). Small or big changes can have a significant impact on the overall behaviour of the interrelated parts. Chan (2001a:1) states that “with co-evolution, elements in a system can change based on their interactions with one another and with the environment”. Parts of the problem cannot be understood in isolation. It is important to have a balanced focus on both the whole and its parts, and to take multiple perspectives into account (Checkland & Poulter, 2006:7).

Figure 6-1 shows the adaptive nature of a system; this system has subsystems and an environment in which it exists. According to Checkland (2011:128) the operation of any system depends on four requirements:

The boundary of a system is the scope of interest or concern, and can change as the scope of interest changes. Any external variable that does not affect the systems variables are not part of its environment. A system has structure that defines its parts (elements) and their relationships. Systems' parts are intercalated elements.

Systems are generally open, i.e. they interact with the environment. Ackoff (1971:663) argues that systems' environments "consist of all variables which can affect its state". A system takes inputs from its environment, and transforms these inputs into desired outputs. This means that a system reacts to the events of its components. A reaction event is a change that is caused by another event.

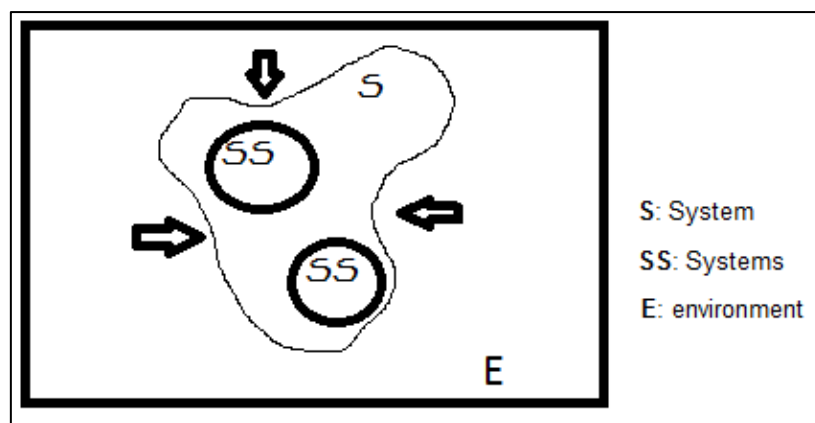


Figure 6-1: Adaptive system (Adapted from Checkland and Poulter (2006:203))

Figure 6-1 above shows that systems can also be part of larger systems. For example, a car's security system can be a part of a car.

The term system thinking refers to a process of understanding of thinking in terms of 'wholes' (Parker, 2000:44, Checkland, 1997a:667, Ackoff, 1971:661). Holistic thinking is a feature of systems thinking because it allows one to see the bigger picture rather than examining each part; i.e. the whole system is examined. Holistic improvement of the system is the aim of systems thinking, rather than focusing on snapshots of isolated parts of the system (Ackoff, 1971:661). Systems thinking allows the researcher to understand the dynamics of the relationships (one affects the other) and the complexities that influence the system (Ackoff, 1971:661). This is done through the understanding of emergent systems properties. These properties

are gained by the system as a result of being a whole system; the separate parts do not have these properties when considered in isolation (Checkland, 1997a:668).

The issue of data warehousing/ business intelligence (DW/BI) education must not be underestimated. This field of study is becoming increasingly competitive, more complex, and more difficult to manage. Two systems methodologies, soft systems methodology (SSM) and critical systems heuristics (CSH) are used in this study to better understand perspectives of different stakeholders in the DW/BI education environment. These systems thinking methodologies are presented later in the chapter by means of an illustrative example. The next section provides the context of the illustrative example.

6.3 Illustrative problem situation

A problem situation is used as an example to demonstrate and explain systems concepts in this chapter. Please note that this is an illustration to demonstrate and explain concepts in this chapter. Data used in this chapter is intuitively deduced and thus it is not collected as part of the empirical study. Chapter 7 presents with the empirical study.

Problem situation: The researcher is currently working in industry; recently, he became aware of a problem regarding higher education graduates. He intuitively assumed the problem to be that students are not being taught satisfactorily to prepare them for the DW/BI industry. So, he decided to contribute towards the solution of the problem by enrolling for an MSc degree to research the problem. It became clear that different stakeholders are involved in the situation; they have different views of DW/BI education. The investors (industry stakeholders) and faculty members (institutions of higher education) may view DW/BI education as a business investment venture in information technology (IT) graduates; lecturers may view it as a means to enable students to acquire employment; whilst parents and students may view it as an opportunity to improve knowledge and education. The success or failure of DW/BI education means different things to the stakeholders with these different viewpoints.

The systems thinking methodologies discussed in the next sections (SSM and CSH) are demonstrated using this problem situation, namely the researcher's perspective of the BI/DW education problem while pursuing a Master's degree.

6.4 Soft systems methodology

Formulating and defining the need that must be met is always part of the problem (Checkland & Poulter, 2006:xii). This is because in real life different stakeholders approach the same problematic issue or process differently, based on their different perspectives. Traditional problem solving strategies often fail to work on complex and messy realities of life, and these are thus difficult to manage. The term 'messy' here refers to people's views and culture that are not well defined. The prescriptive answers about what is the "right" answer to these messy problems may not be appropriate. An approach/methodology, such as SSM that can be used to enquire and manage knowledge of these messy problems, is necessary. SSM is a well-established academic and practiced method that focus on identifying the needs of all the stakeholders involved (Bjerke, 2008:7). It is a suitable method to systematically find out what the current situation is; it also provides areas of improvement that are accommodative to all the current needs.

Soft system was developed by Checkland (2011:497) in the 1960s. Checkland (2011:497) defines SSM as a methodology for 'wicked' problems that resist sharp definitions or structured approaches; it is used to formulate conceptual models for structural thinking in complex human situations and the understanding thereof. SSM uses a collection of models; it is viewed as an organised learning system. These models are devices or tools that are used to explore messy (soft system) problem areas in an organised (hard systems) way which then leads to questions of the real world situation (Patel, 1995:13). Checkland (1997a:669) categorises the individual's perception of a system as either hard (a structured systemic approach) or soft (ill-structured real world problem areas representing multiple worldviews). Hard systems can be predicted to some extent, while soft systems are not predictable – they are messy and involve human activity (Checkland, 2013:1, Patel, 1995:13). "Soft systems is used for understanding motivations, viewpoints, and interactions and

addressing both the qualitative and quantitative dimensions of the situation” (Bjerke, 2008:15).

The wholeness of a system is a fundamental concept that is incorporated in SSM (Patel, 1995:14). There are a number of elements and their interactions that cause complexity in systems; intricate inter-twinings or inter-connectivity of many individual elements result in complexity. SSM is used in an enquiry process that is applied in problem analysis and solving of complex and unstructured (messy) situations (Bjerke, 2008:14). This is achieved applying the principles of “systems thinking” in a cyclic action research approach, with the intention of learning, reflecting and understanding of various people’s perceptions that are involved in a problem situation. Figure 6-2 below illustrates the difference between hard and soft systems.

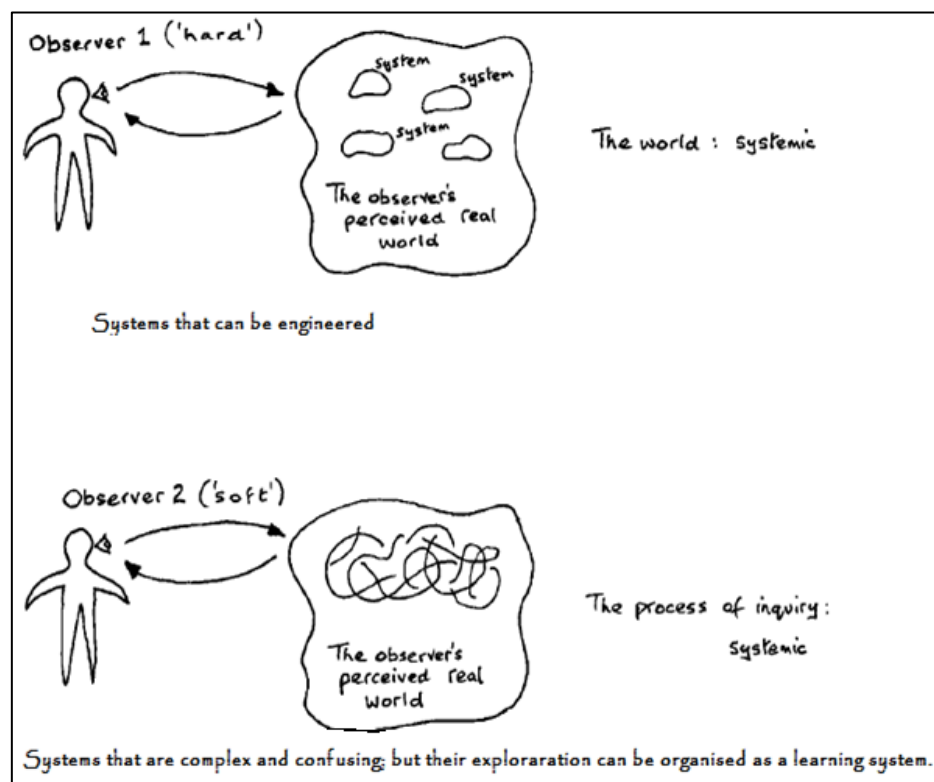


Figure 6-2: The hard and soft systems stances (Checkland, 2000:19) ¹

¹ Checkland promotes the uses of hand drawn figures

Soft systems methodology helps the researcher to gain knowledge by using an action research approach; it also explores the notion that humans act purposefully (Checkland & Poulter, 2006:2). Purposeful action implies that “something needs to be done about this”. The ultimate goal of this process is to improve the problematic area. Checkland and Poulter (2006:11) oppose using the words “problem” and “solutions”; they say that “problem” suggests that there is a clearly defined problem at hand and that “solution” implies that SSM will fix it. They also argue that the goal of SSM is to seek accommodation rather than consensus; this accommodation must be desirable and culturally feasible to the all parties that are involved.

6.4.1 Soft systems methodologies main activities

Checkland and Poulter (2006:13) explain that the SSM learning cycle is a process consisting of four main activities – these are discussed in the subsequent sections. Figure 6-3 below summarises the SSM learning cycle.

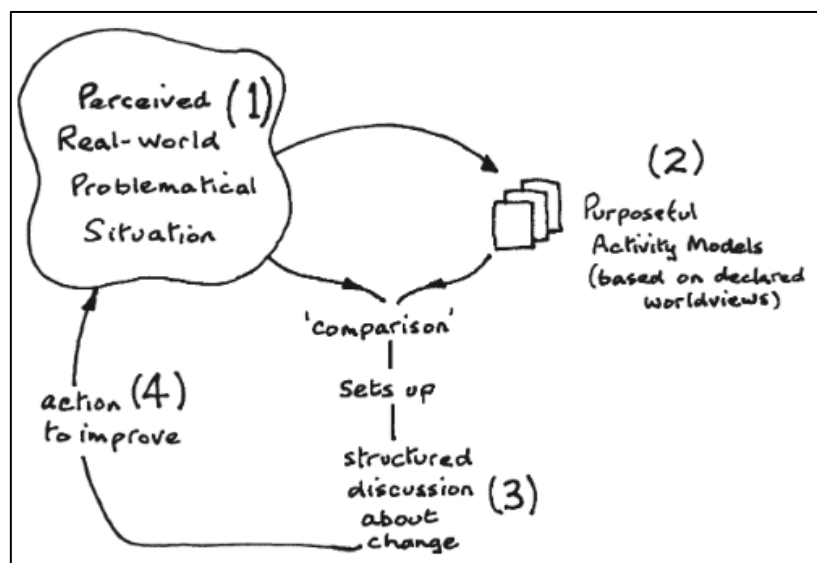


Figure 6-3: SSM learning cycle (Adopted from Checkland and Poulter (2006:207))

6.4.1.1 Perceived real-world problem situation

The first activity entails exploring the initial problematical situation. Checkland and Poulter (2006:209) suggest unpacking the problem situation in an unstructured format. The problem situation may also be experienced by the researcher, which leads to presumptions about the nature of the situation. The first step is to formulate the problem situation in terms of input, transformation, and output. In the DW/BI education problem students who are unprepared for industry must be transformed into students that are better prepared through proper education. This transformation is done in such a way to allow the researcher to obtain an MSc degree. Figure 6-4 illustrates the transformation process involved with this problem situation.

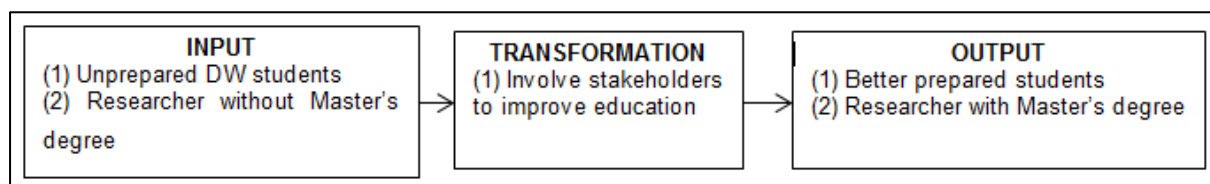


Figure 6-4: DW/BI Education problem transformation element

The second step is to express the problem in terms of a detailed rich picture, within which the problem occurs. The main goal of this rich picture is to capture logic, culture, and relationships. Checkland and Poulter (2006:209) provide structures, processes, climate and people, issues expressed by people, and conflicts as guidelines to be included in the rich picture. A rich picture is a holistic drawing or an impression of a problem area of interest that is formulated using interviews and existing documents (Patel, 1995:15, Checkland & Poulter, 2006:209). This picture is hand drawn and takes the form of real world activities and words in an attempt to convey a clear understanding. A rich picture drawing facilitates a process of unstructured learning in the real world. Consider the illustrative problem situation in Section 6.3 as an example. A rich picture of the perceived proposed problem situation is depicted in Figure 6-5.

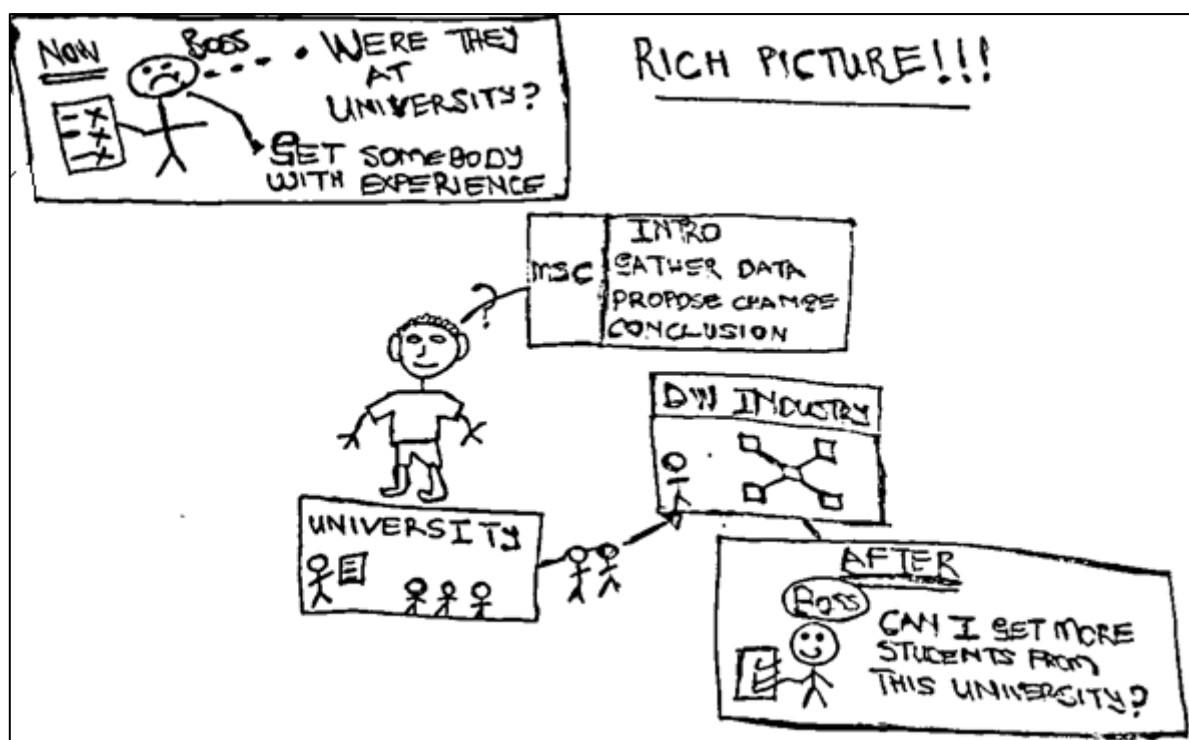


Figure 6-5: DW/BI education problem rich picture

After completion of the rich picture, the second activity is to create purposeful activity models as was indicated earlier in Figure 6-3.

6.4.1.2 Making relevant purposeful activities

The second activity is to make relevant purposeful activities based on worldviews. SSM aims to unpack the problem situation into a set of relevant perspectives. Each perspective provides a particular model base; a model thus can never be a definitive description of the real world (Checkland & Poulter, 2006:218). Different stakeholders can, for example, have different perspectives about the problem situation that was illustrated in Section 6.3. These perspectives will inform the purposeful action to be taken in order to improve the problem situation. The models are ideal situations. They model one way of visually examining involute authenticity of the world.

This process of constructing purposeful activities starts by a statement that expresses a purposeful activity from a particular worldview. A detailed process of constructing purposeful activities is explained in Section 6.4.2 later in the chapter.

The purpose of the current discussion is to provide a holistic view of the overall process depicted in Figure 6-3.

6.4.1.3 Comparison and structured discussions

The result of the second activity is a set of conceptual models representative of the worldviews of the different stakeholders. The third activity uses the models to question the real situation; its purpose is to structure a discussion about the situation and changes that will be 'arguably' desirable (Checkland & Poulter, 2006:218).

The conceptual models are used in comparison (and contrast) with reality and with each another. The difference between the models and the real world are used as the basis for a discussion that intends to find the changes that are both arguably desirable and also culturally feasible (Checkland & Poulter, 2006:193). These are often misunderstood; the aim is accommodation of different worldviews. The main goal is to find the to-be-changed situation. Once that has been identified the next step is then to seek a settlement or a compromise regarding people's conflicting worldviews; they should be able to live with this compromise.

6.4.1.4 Define or take the action to improve the problem situation

The last activity is to define or take the action to improve the problem situation. The previous activity identifies some potential changes. These changes may be technical (or not); these changes must be appropriate for the future. Now, the most desirable and feasible changes identified are implemented. Figure 6-3 (earlier in the chapter) shows the cyclical SSM learning process that extends from exploring a problematical situation, to defining it or taking action to improve it. The end results of the fourth activity point back to the first activity, i.e. the perceived real world problem situation. This is a repetitive process. The next section details the development of purposeful activity diagrams.

6.4.2 The SSM learning cycle (making purposeful activity models)

The purpose of this section is to demonstrate how learning can be acquired in an organised way, by making use of models of purposeful activities (refer also to the

second activity (2) in Figure 6-3 earlier in this chapter). Such models are constructed through the declared worldview; they form the basis “for asking questions about real situation” (Checkland & Poulter, 2006:218). Figure 6-6 below illustrates the guidelines that help to build the models of purpose activity.

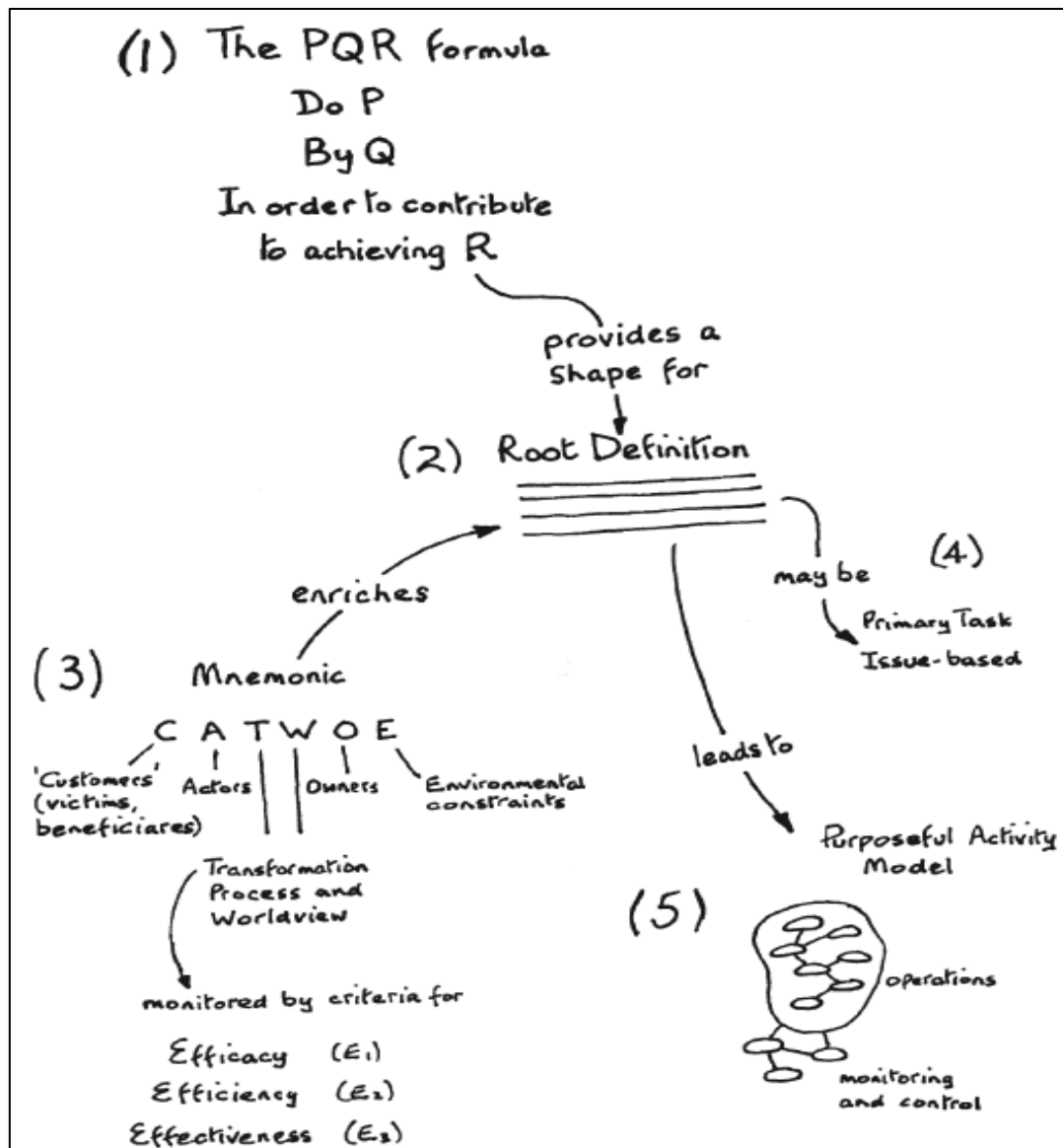


Figure 6-6: Purposeful activity building guidelines (Adopted from Checkland and Poulter (2006:220))

To explain the above guidelines, consider the illustrative problem situation in Section 6.3 of this chapter.

6.4.2.1 The PQR formula

A sentence of the form: 'Do **P** by **Q** in order to contribute to achieving **R**', must firstly be constructed by a SSM practitioner. The PQR formula is used as the *enriching process* for the root definition (RD). The RD enrichment is achieved by answering questions such as: what? how? and why? Checkland and Poulter (2006:219) give the transformation formula PQR meaning: Do P (what?) by Q (how?) in order to help achieve R (why?). The key element in this formula is Q, which represents the transformation process. The other two (P and R) are the supporting elements that define Q in the equation. Transformation is the process of converting or achieving the output given a specific input.

In this DW/BI education problem, the researcher wants to *improve the data warehouse syllabus (P)* by *involving stakeholders* guided by research methodology (Q) in order to *better prepare students*, and meet the MSc degree requirements (R).

6.4.2.2 CATWOE

The second step is the application of the CATWOE model. One of the known SSM tools, and central to the process of deriving a root definition, is CATWOE. The CATWOE mnemonic is used to define the customer (C), actors (A), transformation process (T), Weltanschauung (the worldview) (W), ownership (O), and environmental constraints (E) (Checkland & Poulter, 2006:220). The CATWOE analysis is used to articulate the scope and components (Checkland & Poulter, 2006:222). This means that it is used to check that the root definition is expressed properly (Patel, 1995:16). It can be seen as a thinking stimulant about problems and solutions. Also, it identifies the elements and context of a purposeful and involves finding out about the following elements:

C- Customer (beneficiaries or victims) – affected by the transformation.

A- Actors – people that will do the activities which make up the transformation.

T- Transformation – a purposeful activity carried out by actors.

W- Worldview – derived from the rich picture.

O- Owner – who can stop or change the purposeful activity.

E- Environment – constraints such as regulations, politics, budget etc.

Table 6-1 illustrates CATWOE applied to the DW/BI education problem situation. The industry and students are considered to be customers in the problem situation; the actors (the doers of action enabling that transformation) are viewed as the lecturers and students.

Table 6-1: CATWOE analysis of DW/BI education problem situation

Element	Improvement of students	Obtain MSc degree
C	Industry and students.	The researcher.
A	Lecturer teaching and students who are learning.	The supervisor and researcher.
T	Data warehouse education not aimed at industry needs – to – data warehouse education aimed at industry needs.	A well planned MSc study.
W	Universities can improve data warehouse education to meet the needs of industry.	Improve the situation while doing Master's degree.
O	Faculty management but also the lecturer.	The supervisor and faculty director of research.
E	Constraints of degree programmes such as student credits; Types of evaluation; Time constraints of industry.	The requirements of the MSc degree as stated by the university.

The evaluation criteria elements, i.e. efficacy, efficiency, and effectiveness, are used in regard to the problem situation. This evaluation process is to assess if the transformation process is: (1) producing the intended outcome, (2) achieved with minimal resources, and (3) achieves a long term goal (Checkland & Poulter, 2006:222: 222). The evaluation criteria related to the stated DW/BI education problem are applied in Table 6-2.

Table 6-2: Evaluation criteria for DW/BI education problem

Evaluation criteria	Improvement of students	Obtain Master's degree
Efficacy	Proposed guidelines are achievable within the constraints of the university and students	The project is achievable in terms of the requirements for Master's study
Efficiency	A module designed according to the guidelines is realistic in terms of the time required of students and lecturers and the input required by industry.	The effort required to perform the study is realistic according to credit –hour requirements of university and input required from the supervisor.
Effective	Students are better prepared to contribute to industry after completion of a data warehousing module designed according the proposed principles.	A Master's degree has been obtained in accordance to the regulations of the university.

6.4.2.3 Root definition

The RD is a rich description of the activity formulated in order to improve the efficiency of the analysis (Checkland & Poulter, 2006:219). The following is an example of a RD of the illustrative example:

Root definition - a well planned and executed MSc study that satisfies the criteria for good research and that will lead to the development of module guidelines that are desirable and feasible for universities to implement that will better address the needs of industry for DW/BI professionals. The DW/BI professionals should participate in the study within the constraints of their availability. The study should adhere to the requirements without taking unreasonable amounts of time from the student and supervisor.

This root definition describes an agreeable future view of the DW/BI education problem situation.

6.4.2.4 Primary vs issue based model

In the fourth step, Checkland and Poulter (2006:223) advocate that a choice must be made between: (1) a primary task, and (2) an issue based RDs. Primary task RDs are based on the notion that they accept problem situation area boundaries, such as

limiting the DW/BI education problem situation to the computer science department. The issue based RDs are those that require a broader consideration in a problem situation. For example, consider again the illustrative problem situation in Section 6.3. The RDs and models are said to be issue based if their boundaries do not coincide with departmental boundaries.

6.4.2.5 Purposeful activity models

The fifth step is the development of activity models based on the RDs. At this stage activities required to achieve the transformation outcome are bound together. Checkland and Poulter (2006:225) advice on the following logical process for building the SSM activity models:

1. Assemble guidelines using PQR, CATWOE, and the RD.
2. Start from T and W in a CATWOE to name the purposeful action as a transformation. This means identifying activities related to:
 - (1) entity to be transformed,
 - (2) transformation process, and
 - (3) entity in a transformed state.
3. Structure activities according to their dependencies.
4. Add the monitoring control activities.
5. Check the mutual dependency of guidelines and models.

As an example, Figure 6-7 is drawn based on the DW/BI education problem; it illustrates the flow of relevant activities.

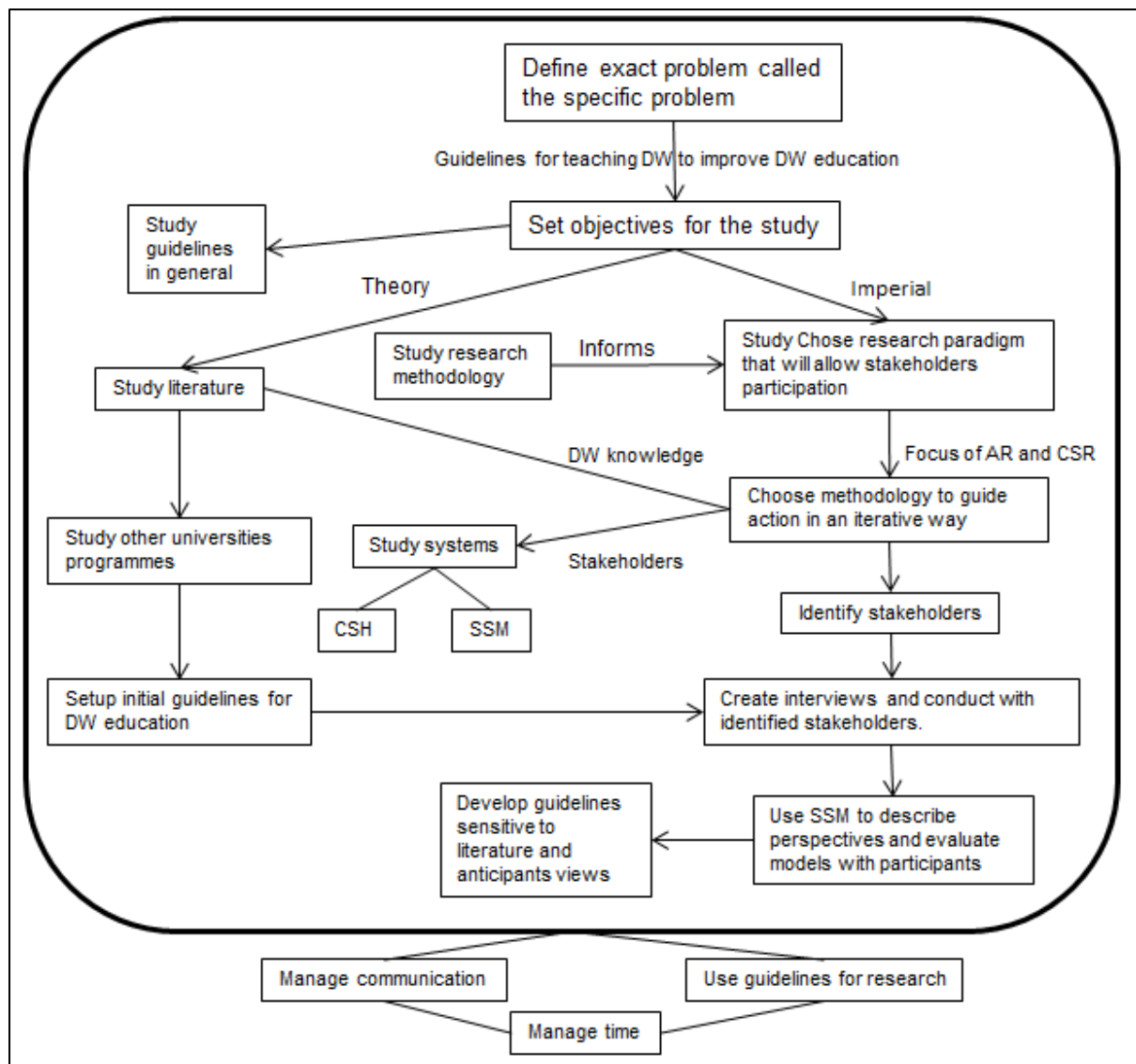


Figure 6-7: Illustration: DW/BI education purpose activity model

The next section discusses the application of SSM and action research in the empirical portion of this study.

6.4.3 Position of the study

This section explains how SSM and action research are applied to help solve knowledge management problems (acquiring of knowledge based on different perspectives), within the context of DW/BI education in institutions of higher learning. SSM is used in this study to build reflective stakeholders models that are purposeful to the problem situation. The purpose is to make sense of this situation by understanding the DW/BI education problem situation, what the stakeholders

perceive it to be, and what they think it ought to be. Figure 6-8 below illustrates the application of SSM to the problem situation.

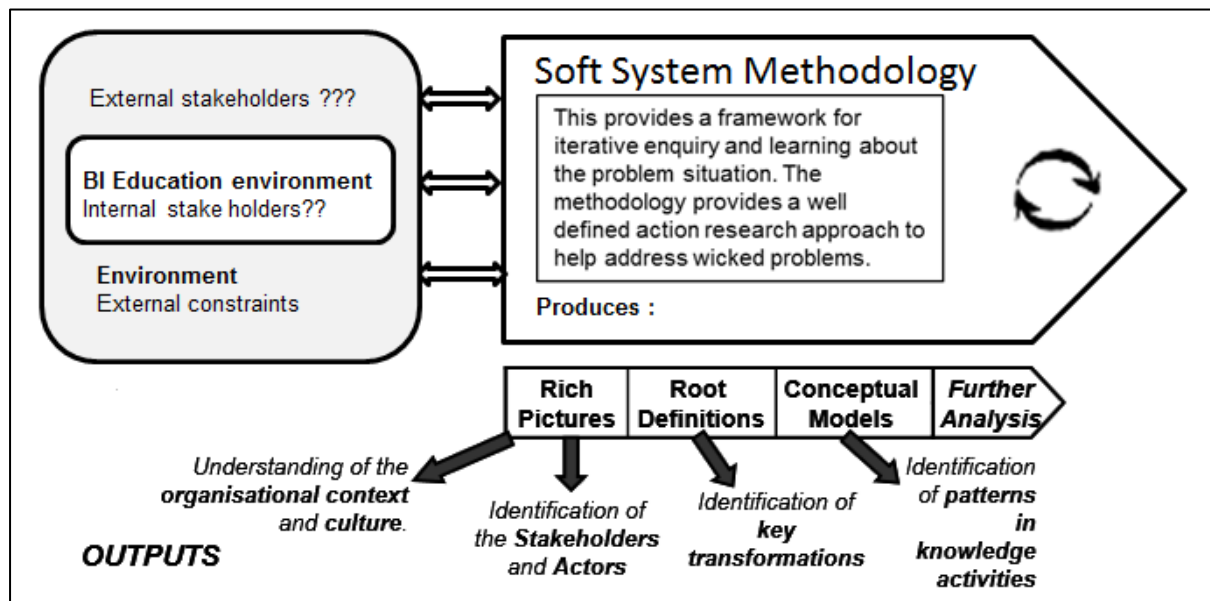


Figure 6-8: Applying SSM to knowledge management in the DW/BI education problem

Interviews are developed and presented in Chapter 7 aiming at understanding the perceptions of different stakeholders on the problem environment of this study. Analysis of the interview data not only provides information on module content, but also on the process to improve DW/BI education at universities. Interview questions were guided by the results of Chapter 5 (combined skill sets from universities' programme/ modules and DW/BI literature) as well as the components of the systems thinking methodologies discussed in this chapter. SSM diagrams are used to present the different perspectives of different groups of stakeholders who participated in the empirical research and are presented in Chapter 8.

6.5 Critical systems thinking

The systems movement evolved with time, and there are strands that developed within it. Within systems the critical systems stance was added. Ulrich (1983:20) defines the term "critical" as the process of discerning or judging carefully while taking care to safeguarding oneself against error. Parker (2000:47) simplifies this concept by stating that in a traditional context criticality may mean "existing standards or norms against which something is judged". However, Ulrich advocates

that critique must be used in a 'higher', reflective sense, to such an extent that the standards that govern actions are also questioned. "To be 'critical' then, above all else, means to become self-reflective in respect to the presuppositions flowing into one's own judgements." (Ulrich, 1983:20).

Critical systems thinking (CST) is one approach that emerged as a reaction to hard systems approaches shortcomings with regard to their applications in a given social and political context (Simon, 2003:1). Early approaches that made use of systems ideas (such as systems analysis, systems engineering, and operational research) were appropriate for tackling certain well-defined problems (Jackson, 2001:223). Broader perspectives proved hard systems approaches' inadequacies (Bawden, 1991 2362). According to Jackson (1985:135) hard systems have three points of critique: (1) the existence of real world is taken for granted, (2) a search for regularities and casual relationships between systems interaction, and (3) the deterministic stance of human behaviour. Hard systems approaches are not suitable for use in social systems because "social systems are based on the interactions of living systems" (Fuchs & Hofkirchner, 2009:111).

Critical systems thinking was developed "to allow analysis of complex societal problems and intervention to resolve such problems" (Jackson, 2001:233). This means taking a critical approach to systems' problem solving. This is an attempt to reduce human-unverified claims and thoughts, and to find the truth by questioning and criticising individuals' cognitive developments (Yenice, 2011:497).

6.6 Critical systems heuristics

Complex situations may bring about opposing perspectives that requires a constructive approach (Ulrich & Reynolds, 2010a:243). This includes making sense of situations, understanding assumptions, and seeing the bigger picture. Multiple perspectives are problematic and need a critical approach; the fact that that there may be more than one right way to solve an issue, requires critical intervention (Ulrich, 2001:1). This complexity necessitates a framework within critical systems thinking, to organise and articulate our thinking in order to make the process clear and traceable. This is a "process of using systems concepts to organise evaluation,

critical reflection, and reflexion that enhances the possibilities for order and rigour in action research” (Hart & Caceres, 2012:6). Ulrich (1983:24) developed critical system heuristics (CSH) to be applied in situations that require such social planning.

Critical system heuristics “is a framework for reflective practice based on practical philosophy and systems thinking” (Ulrich, 2003:1). He uses three terms: (1) *critical*, (2) *systems*, and (3) *heuristics*. Firstly, Ulrich (2005b:1) argues that a complex problem situation requires a *critical* approach, because there is no single right way to resolve the issue. Problem situations require the *critical* element, which is the process of expressing or involving an analysis of the merits and faults. And thus, there can be more than one right answer with alternative assumptions. Secondly, social settings require a method/approach such as *heuristics* to discover the meanings attached to them. *Heuristics* means to discover the meanings attached by identifying and exploring problem aspects, such as assumptions, questions, or solution strategies, as opposed to deductive (algorithmic) procedures (Ulrich, 2005b:1). For example, to apply “common sense” is a heuristic that is applied to a problem based on an individual’s experience and observation of a situation. Third, Ulrich (2005b:1) uses the term *systems* to refer to totality or wholeness of relevant problems that are based on the relevant theoretical or practical judgements they depend on.

6.6.1 Boundary judgements

The application of the systems concept to a segment of the real world involves making assumptions about what is to belong to the system in question (Parker, 2000:46). Ulrich (2005b:2) refers to these assumptions as boundary judgements. From his perspective ‘boundary judgment’ means defining what is relevant knowledge for a particular inquiry, who is relevant in generating it, and having stake in it. CSH is guided by boundary judgements that determine which empirical observations and value considerations are relevant to the phenomenon (Ulrich, 2005b:2). These empirical observations and value considerations are formulated as a claim or position, e.g. using processes of opinion formation, problem solving, decision-making, action, or conflict resolution. The assertions or suggestions that are proposed to be relevant in the reference system are called claims (Ulrich, 2005b:2).

The merits of these claims are then assessed (filtered) by checking the impact that they potentially have in the reference system. Ulrich (2005b:2) proposes that merits of a claim are to be examined by the following questions:

1. What difference does it make in practice?
2. Who will or not benefit?
3. What improvement or emancipation does it bring to?

“The sum-total of these considerations of fact and of value make up the reference system that gives meaning to a particular claim and conditions its validity” (Ulrich, 2005b:2).

Boundary judgements help to avoid an infinite regress of grounding the underlying validity claims in systems thinking, and in the elimination of impartiality claims (Ulrich, 1994: 14). Systems need a focused approach in order to eliminate impartiality, or to deliver emancipation or a reflective view.

Critical systems heuristics achieves this focus by making use of boundary judgments. This helps to determine which empirical observations and values count and which ones do not count by exploration of a wide range of perspectives. Ulrich (2005b:1) advocates a critical (critique) approach towards boundary judgement. A boundary-critical attitude takes into consideration three important factors: the (reference) system, facts, and values. As alluded to earlier, a system means a perceived situation or circumstances of underpinning boundary judgments. Facts refer to all observations that are taken into account under the reference system. Values refer to all the evaluations done in reference to the system. Ulrich (2005b:1) emphasises “judgments of fact as well as of value depend on boundary judgments and are connected through them”. Figure 6-9 illustrates the eternal triangle of boundary judgments, facts and values. For example, the expansion/decrease of a reference system’s boundary may add/decrease facts and values. The impact of such a change must be evaluated.

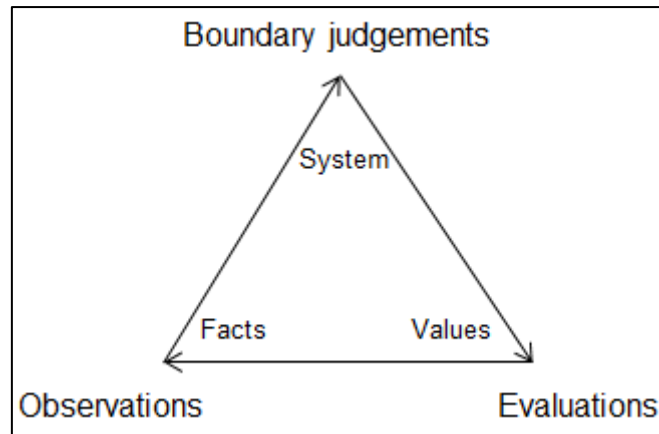


Figure 6-9: The 'eternal triangle' of boundary judgments, facts, and values (Adopted from (Ulrich, 2005b:1))

According to Ulrich (2005b:1), boundary judgements can either be used reflective (applying them self critically) or emancipatory (for critical purposes against those who may not handle them so self-critically). Table 6-3 below illustrates four perspectives for examining the selectivity of any claim. The four issues help to make claims clear and valid because they reflect on the reference system. The empirical selectivity implies what is current; the normative selectivity implies the ideal situation.

Table 6-3: Four perspectives for examining selectivity (Adopted Ulrich (2005b:1))

Perspective	Empirical selectivity ('Is mode')	Normative selectivity ('Ought mode')
Facts	Actual mapping What 'facts' are considered relevant and which ones are left out?	Ideal mappings What 'facts' ought to be considered relevant and which ones should be left out?
Values	Actual mapping What 'values' are considered relevant and which ones are left out?	Ideal mapping What 'values' ought to be considered relevant and which ones should be left out?

According to Ulrich (2005b:2) there are four basic boundary issues that are important in CSH; these are the basis of:

1. Motivation - sense of purpose and value (why?).
2. Power - authority to decide the fate of what can and cannot happen?

3. Knowledge - What experience and expertise support the claim?

4. Basis of legitimacy - Where does legitimacy lie?

Ulrich (2005b:3) deepened this by defining three categories for each boundary issue. Figure 6-10 gives an overview of boundary categories, issues, those involved, and those affected within referenced systems.

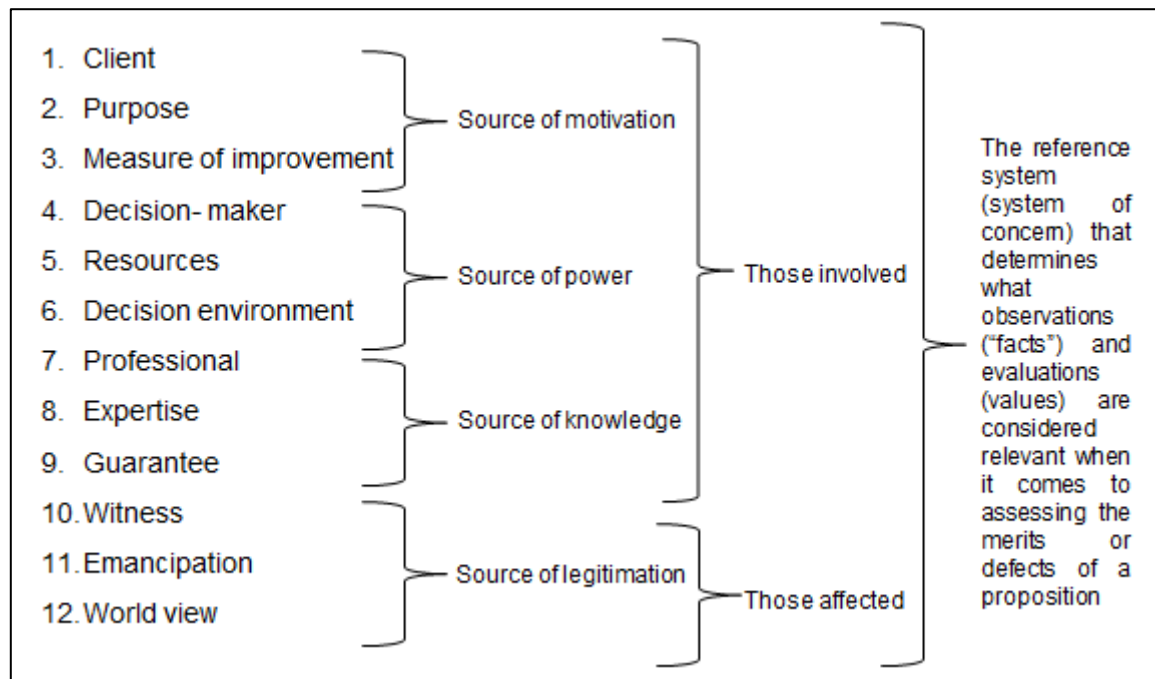


Figure 6-10: Table of boundary categories (Adopted from Ulrich (2005b:1))

Ulrich (1996:33) developed a set of twelve questions to guide systemic inquiry into a current situation, and also to consider potential alternatives. He then categorically classified these questions into four groups, i.e. sources of: motivation; power/control; knowledge, and legitimisation – refer to column 2 in Table 6-4 below. Those that are involved in a situation are then identified using sources of motivation, power, and knowledge groups. The affected parties in the situation are identified through the source of legitimation group. Each of these four groups of sources has three questions coupled to it; these questions critique the first order boundary judgments being made in the conceptual modelling process. In each group the first question is concerned with the role of relevant stakeholder (refer to No. 1, 4, 7, and 10 in Table 6-4 below). The second question of each group relates to the relevant concerns of the role identified in the first question. The third question of each group relates to

issues important in qualifying the first two questions. The questions are asked in two forms: first, the current/presents ('is', 'are', 'have'), and second, the ideal ('ought to be') states.

Table 6-4: Table of boundary categories (Adopted from Ulrich (1983:33))

No.	Questions: First order enquiry	
1	Sources of motivation	Who is / ought to be the client (beneficiary) of the system S to be designed or improved?
2		What is / ought to be the purpose of S such that it is serving the interests of the client?
3		What are / ought to be S's measures of success (or improvement)?
4	Sources of power	Who is / ought to be the decision maker, who has the power to change S's measures of improvement?
5		What are / ought to be the components (activities, resources, constraints) of S within the control of the decision maker?
6		What resources and conditions are / ought to be part of S's environment? i.e. not controlled by the decision maker?
7	Sources of knowledge	Who is / ought to be involved as the designer of S?
8		What kind of expertise is / ought to be informing the design of S? i.e. concerns who fulfils role of "designer" – experience, knowledge, skills etc.
9		What are / ought to be the designer's assumptions underpinning their design?
10	Sources of legitimization	Who is / ought to be the relevant stakeholders (witnesses) affected by S and what is / ought to be their role?
11		To what extent and how do / ought these stakeholders to have control over their own interests in S?
12		Whose "worldviews" or perspectives are / ought to be underpinning the design of S?

6.6.2 Demonstration of critical systems heuristics

This section presents a demonstration of the application of CSH to the illustrative problem that was discussed in Section 6.3. Table 6-5 below illustrate the boundary judgement for the DW/BI education problem. Answers are provided to the questions in the "ought to" mode.

Table 6-5: Boundary judgements: DW/BI education problem (MSc study perspective)

Question	Improvement of students	Obtain MSc degree
Who is / ought to be the client (beneficiary) of the system S to be designed or improved?	Industry, university and students.	The researcher and the supervisor to a lesser degree.
What is / ought to be the purpose of S such that it is serving the interests of the client?	(a) The industry gets to benefit by getting graduates that are more ready for work (b) The university gets better recognition of their quality from industry (c) The students get to be more marketable.	To obtain the degree.
What are / ought to be S's measures of success (or improvement)?	Module content must represent the needs of both the industry and students, but all these must be within the university constraints of operation.	The requirements for a Master's degree.
Who is / ought to be the decision maker, who has the power to change S's measures of improvement?	The faculty management and the lecturer can decide to stop the development of the new module.	The supervisor and the director of research of the faculty.
What are / ought to be the components (activities, resources, constraints) of S within the control of the decision maker?	Experienced and expert knowledge of DW/BI development.	Participants of the study; data collection and analysis methods.
What resources and conditions are / ought to be part of S's environment? I.e. not controlled by the decision maker?	The other modules in the programme of the students that has an impact on their availability and the requirements of the faculty in term of credit-hours.	The regulations of the university governing Master's degrees.
Who is / ought to be involved as the designer of S?	The industry experts and the lecturer.	Student and supervisor.
What kind of expertise is / ought to be informing the design of S? i.e. concerns who fulfils role of „designer“ – experience, knowledge, skills etc.	The lecturer and the industry professionals have knowledge of DW/BI content and the lecturer has knowledge of education matters.	The supervisor has knowledge of research methodology and systems thinking which needs to be conveyed to the student in order for him to take ownership of the study.
What are / ought to be the designer's assumptions underpinning their design?	The universities as producers of graduates and the industry as the receivers of students from universities are able to develop a module that is to the real benefit of the students.	It is possible to improve the students and meet the requirements of Master's degree.
Who is / ought to be the relevant stakeholders	The industry, students, and the lecturers are directly affected.	The interests of the r students' family should be taken into

(witnesses) affected by S and what is / ought to be their role?		account.
To what extent and how do / ought these stakeholders to have control over their own interests in S?	Each stakeholder must find an accommodative ground that is acceptable to address their needs with regards to DW/BI education. The lecturer should represent the needs of the students.	The supervisor is sensitive to their needs.
Whose „worldviews“ or perspectives are / ought to be underpinning the design of S?	Industry, students and the academic professionals (lecturers, dean) should be addressed by the proposed guidelines.	The perspectives of the student and the supervisor sensitive to the requirements of the faculty.

The next section explains the application of CSH in this study.

6.6.3 Position of the study

As discussed in the previous section, CSH presents a way of making explicit the assumptions underlying systems design through boundary judgements. CSH boundary judgements are thus used in this study as an approach to better understand the different perspective of the stakeholders in the problem situation. The purpose is to attempt to make the boundary judgments explicit and defensible. CSH, with the set of 12 questions, can guide (and improve) the systemic enquiry into a current situation and enable consideration of potential alternatives.

The CSH boundary judgements are used together with SMM components to better understand and represent the views of different stakeholders in the problem environment. The representation and accommodation of the different perspectives are presented in Chapter 8.

6.7 Summary

The purpose of this chapter is to present SSM and CSH methodologies, and to create a shared understanding of these underlying concepts. In Chapter 1 the two methodologies were chosen as instruments to be used in the research problem because of their systems approach nature (Patel, 1995:14). The discussion of the two methodologies in this chapter is to explain these methodologies prior to their

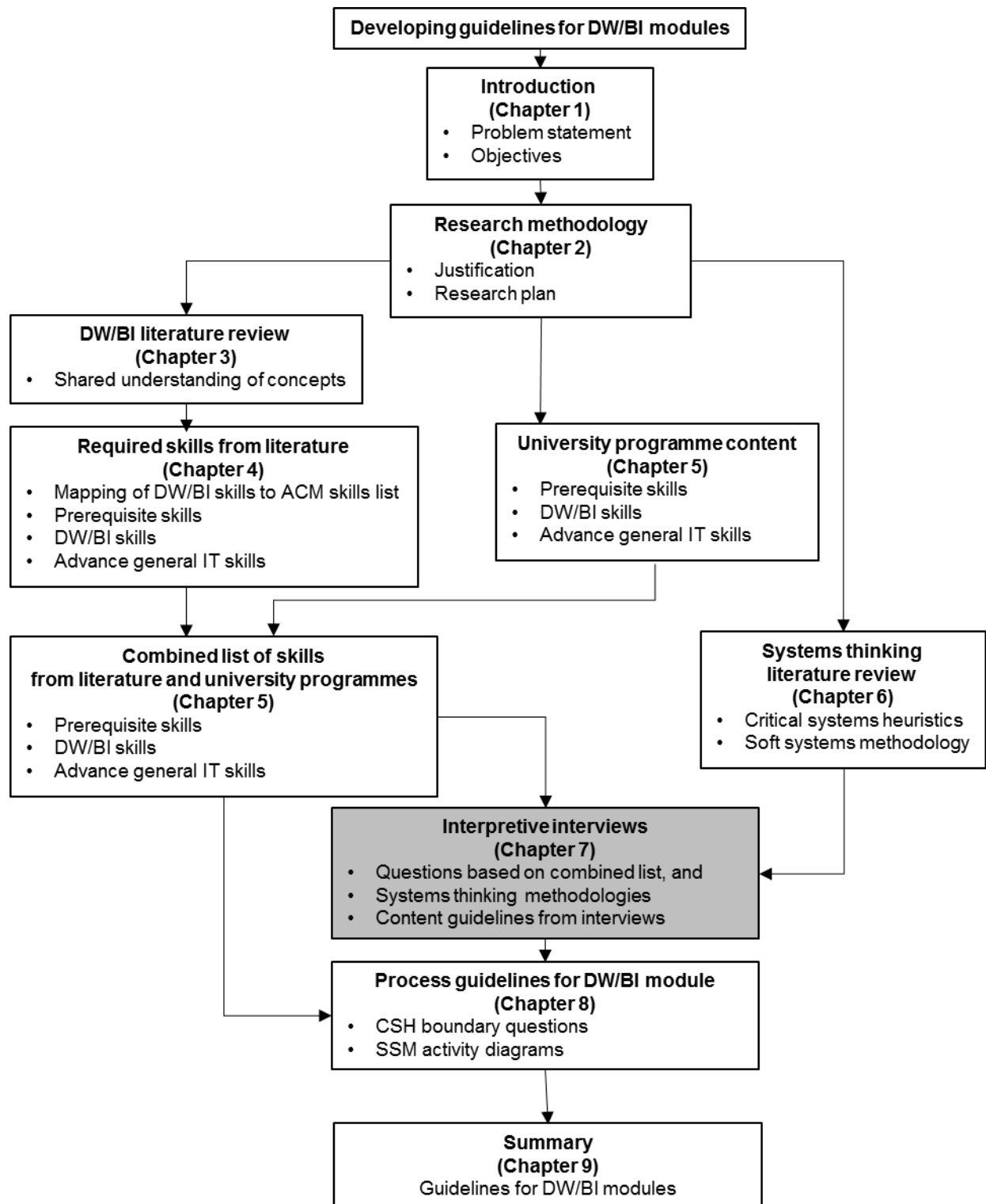
application. An illustrative DW/BI problem situation is used to demonstrate the application of both SSM and CSH.

Soft systems methodology provides a structured and systematic approach for analysing human activity systems. All defined human activities of a system under investigation are viewed as more than just the sum of the parts, hence a holistic approach.

Critical systems heuristics seeks to be holistic; it aims to ensure that theory fortifies practice and is tested in practice. It is proposed that CSH, through its attention to managing the assumptions underlying systems design, may be of use in attempting to deal with the stated problem situation. CSH is used as a second order boundary critique to the problematic situation.

The components of these methodologies together with the list of combined guidelines from universities' programmes/modules are used to guide the empirical study to understand the perspectives of different stakeholders in DW/BI education, presented in the following chapter.

This page is left open intentionally.



Chapter 7 Stakeholders' perspectives on DW/BI education

7.1 Introduction

The objective of this study is to develop guidelines for data warehousing/business intelligence modules in information technology programmes at universities; the guidelines are based on an understanding of the needs of various stakeholders. Two methodologies, i.e. critical systems heuristics and soft systems methodology, are used to understand the various stakeholders' perspectives. Business intelligence was presented in Chapter 3, with the purpose of creating a shared understanding of the underlying concepts. Data collected from universities' business intelligence programmes/modules were used to understand what they currently teach; the knowledge presented in Chapter 3, as well as the universities' programmes, was used to create the initial guidelines in Chapter 5. The purpose of this chapter is to gain an understanding of other stakeholders' (i.e. former students, lecturers and industry) perspectives regarding skills required, as well as the process to be followed, to develop a module in business intelligence/data warehousing on post graduate level. The interview content is also guided by concepts from systems thinking methodologies – refer to Chapter 6 for a discussion.

This chapter focuses on presenting findings from interviews conducted and is organised according to the research process. Section 7.2 presents the interpretive interview development. Section 7.3 presents the data collection process. Section 7.4 presents data analysis. Section 7.5 presents conclusions drawn from coded data. Section 7.6 presents findings that to enrich the initial guidelines, presented in Chapter 5, according to stakeholders' (i.e. former students, industry and faculty members) perspectives. Section 7.7 provides a summary for the chapter.

7.2 Interpretive interview design

The objective of an interview design is to collect exploratory information. For the purpose of this study, the required information is the understanding of the former

students, lecturers, and industry experts with regard to the skills graduates should possess when they enter the industry, and the process of developing a business intelligence/data warehouse (DW/BI) module. Data collection is guided by the initial guidelines developed in Chapter 5, as well as the components of soft systems methodology (SSM) and critical systems heuristics (CSH) as discussed in Chapter 6.

7.2.1 Participants

The targeted participants are experienced business intelligence (BI) professionals in the industry, former students, and faculty members of universities. More detail of the participants is given in Section 7.3.1.

7.2.2 Interview questions

The interpretive interview is made up of four sections, i.e. biographical information, stakeholder identification, actions to be taken to improve DW/BI education, and module content.

7.2.2.1 Section 1: Biographical information

The first section of the interview contains questions on biographical information of the participants as indicated in Table 7-1 below. Table 7-1 presents Section 1 of the interview.

Table 7-1: Biographical information of participants

Section 1: Biographical information
Question
What is your current role in DW/BI? What other roles have you played in DW/BI and how many years of experience associated with each role?
Did you receive formal education to study the DW/BI concepts? If not would you prefer formal education in DW/BI?

7.2.2.2 Section 2: Stakeholder Identification

This section is to elicit information for the questions as raised by CSH.

Section 2 of the interview is presented in Table 7-2 below. The first (left) column presents the twelve CSH boundary questions in the current/present ('is', 'are', 'have') state and the ideal ('ought to be') state applied to DW/BI education. The second column presents the motivation for asking the question. The motivation is linked to the two methodologies: CSH and SSM.

Table 7-2: Section 2: Stakeholders

Section 2: Stakeholder	
Question	Motivation for asking
Who do you think is/ought to be benefiting from a module in data warehousing in a 4 th year IT programme?	• CSH – to know who the beneficiary is. • SSM – to know the client.
According to you, what is/out to be the purpose of DW/BI module such that it is serving the interests of those who benefit from it?	• CSH – to know the purpose of DW/BI module. • SSM - to know the purpose of DW/BI module.
What do think are/ought to be the measures of success (or improvement) for DW/BI module?	• CSH – to know the measure of improvement. • SSM – to know impact of the transformation.
Who do you think has/ought to have decision powers to make a change to improve BI modules?	• CSH - to know the decision maker. • SSM – to know the owner.
What components (activities, resources, constraints) of DW/BI module do you think are/ought to be within the control of the decision maker?	• CSH – to know the resources. • SSM – to know transformation's inputs.
What resources and conditions do you think are/ought to be part of DW/BI education environment but are outside the control of decision maker?	• CSH - to know the environment. • SSM – to know transformation's inputs and the environment.
Who do think is/ought to be involved in the development of DW/BI module?	• CSH- to know stakeholders. • SSM – to know the actors.
What kinds of expertise do think is/ought to be consulted or involved in the development of DW/BI	• CSH – to know needed expertise. • SSM – to know actors expertise.

module?	
Who or what do think is/ought to be assuring successful outcome or conditions of DW/BI module?	• CSH – to know the guarantor. • SSM – to assess if the transformation process is [1] producing the intended outcome (efficacy), [2] achieved with minimal resources (efficiency), and [3] achieves a long term goal (effective).
Who do you think are/ought to be the relevant stakeholders affected by DW/BI module and what is their role?	• CSH – to know the involved and affected stakeholders. • SSM – to know client (Who are the winners and losers?).
What control do you think the affected stakeholders have/ought to have in DW/BI module development or education?	• CSH – to know how the interests of stakeholders are protected. • SSM – to know clients' interests.
According to you, DW/BI module development or education is/ought to representing the worldviews or perspectives of which stakeholders?	• CSH – to know whose views are represented. • SSM – to identify the Weltanschauung.
What do you think should be the process to develop an improved DW/BI module? Please give detailed steps - It would be appreciated if you could draw some kind of activity diagram that shows dependencies.	

The next section presents the stakeholders' views on desired content of a DW/BI module.

7.2.2.3 Section 3: Module content

This section presents questions relevant to DW/BI module content. Table 7-3 below presents Section 4 of the interview. The first (left) column presents questions in the ideal ("ought to be") state to improve DW/BI module education. The second column presents the motivation for asking the question. The aim of these questions are to support or extend the guidelines presented in Chapter 5 (Section 5.6), as well as to support elements of the two systems thinking methodologies, i.e. CSH and SSM.

Table 7-3: Module content

Section: Module content	
Question	Motivation for question
What is your overall perspective of what universities teaches in DW/BI modules? Do you think that the current DW/BI curriculum is preparing students to meet industry skills requirements?	• CHS – to know “is” mode and to get overall impression. • SSM – to know what the context of the situation. • SSM – to gain an understanding of the real/actual problem. • SSM – what is the wider impact of any solution? • To verify the guidelines of Chapter 5.
What is your own experience of DW/BI concepts?	• CHS – to know the “is” mode and to get overall impression. • SSM – to know the context of the situation. • SSM – to gain an understanding of the real/actual problem.
Please tell us what you think universities should teach DW/BI students?	• SSM – to know what are the transformation’s (improvement) outputs. • To verify the guidelines of Chapter 5.
What should be viewed as prerequisite knowledge/skills to enrolling for DW/BI module?	• SSM – to know what are the transformations’ (improvement) input. Identify the participant’s perspective. • To verify the guidelines of Chapter 5.
What general IT knowledge/skills should be taught in parallel with the DW/BI module?	• SSM – to know what are the transformations’ (improvement) outputs. Identify the participant’s perspective. • To verify the guidelines of Chapter 5.
A typical data warehousing project utilizes a host of software solutions including data modelling tools, ETL tools, OLAP tools, DBMS (in many cases specialized data warehousing DBMS). What is your overall view of universities modules in terms in terms of providing students with exposure to these software tools?	• SSM – to gain an understanding of the real/actual problem. • SSM – to know what is the wider impact of any solution. • To verify the guidelines of Chapter 5.
DW/BI module is a discipline that includes many subjects. There are theoretical teaching and practical teaching in this course. What is your overall view of the teaching approach?	• SSM – to gain an understanding of the real/actual problem. • SSM – to know what is the wider impact of any solution. • To verify the guidelines of Chapter 5.
What is your overall view of the assessment component for student knowledge within the DW/BI module?	• SSM – to know what is the wider impact of any solution. • To verify the guidelines of Chapter 5.

7.3 Data collection process

After finalising the questions, the next step was to invite participants in the interview process. The next section presents a summary of the interview process.

7.3.1 Participants

There were approximately 25 potential participants contacted and asked to participate in this research. They are known to the researcher and involved in the DW/BI industry. Only 16 agreed or were able to participate in the research. Of these, 11 were face to face interviews, and other participants responded in writing.

The 16 participants are listed in Table 7-4; it also includes other data such as experience, roles played, and whether they have formal DW/BI education or not. This information represents the biographical information of participants. Respondents are numbered P1 – P16. Each responded are referred to by their number throughout this chapter. The order of this table corresponds to the order in which the interviews were conducted.

Table 7-4: Participants

Participant no	Response type	Years in DW/BI	Roles played in DW/BI	Formal DW/BI education	Would you have preferred formal education in DW/BI?	Learning environment
P1	Telephone interview (Skype)	11	• Lecturer • BI developer	Yes	NA	University
P2	Interview	7	• BI developer	No	Yes	On the job
P3	Interview	20	• BI architect • BI developer • Domain specialist	No	Yes	On the job
P4	Interview	9	• BI developer	No	Yes	On the job

			• BI analyst			
P5	Interview	0		Yes	NA	University
P6	Interview	7	• BI developer • Microsoft developer architect	Yes	NA	University
P7	Interview	0		Yes	NA	University
P8	Interview	3	• ETL developer	Yes	NA	University
P9	Interview	15	• BI manager • BI consultant • Lecturer	No	Yes	On the job
P10	Interview	10	• Developer, analyst • Team lead • Architect • Quality a	Yes	NA	University
P11	Interview	8	• BI developer	No	Yes	On the job
P12	Written	6	• BI developer • BI analyst	No	Yes	On the job
P13	Written	5	• Lecturer	No	Yes	On the job
P14	Written	14	• ETL developer • BI developer • Data modelling • Architect	No	Yes	On the job
P15	Written	5	• BI developer	No	Yes	On the job
P16	Written	6	• BI developer • Team lead	Yes	Yes	University

Table 7-5 below presents a summary of Table 7-4. Out of the 16 participants, 7 (44%) have formal DW/BI education. The remainder of the participants (9 or 56%) do not have formal DW/BI education. Participants that do not have formal DW/BI education unanimously indicated that they would have preferred to have formal DW/BI education before they started a DW/BI career. Participants P5 and P7 have no DW/BI role experience (0 years' experience), even though they have formal DW/BI education. The two participants are included to learn about the perceptions of people not actively involved in DW/BI, but in the general IT industry.

Table 7-5: Participants' summary

Response	Formal DW/BI education	%	Learning environment
Yes	7	44	University
No	9	56	On the job
Total	16	100	

The interview process is presented in the next section.

7.3.2 Interview process

Face to face and telephonic interviews presented the researcher with the opportunity to ask follow-up questions where necessary; written responses did not allow for this. Effort was made to get the participant's perception of the subject matter related to the question. For example, the researcher anticipated that participants that do not have formal DW/BI education may be unable to give an answer about the current state of DW/BI education, since they do not have experience of it. Each question in Section 2 of the interview is qualified by asking the participant to give his/her perception. For example, the code excerpt below shows the response to Question 5 of participant 15 (P15). Even though P15 did not receive any formal DW/BI education, a response was given. Harris and Brown (2010:1) advice researchers to "create tightly aligned and structured instruments" to be used for data collection method.

Question 5

T - Who do you think currently has the decision powers to make a change to improve DW/BI modules? What is your perception?

P15: I think universities have the decision powers to improve DW/BI modules.

T - Who do you think ought to have decision powers to make a change to improve DW/BI modules?

P15: I think universities should still have the decision to improve DW/BI modules without ignoring the input from students and the demands of stakeholders in the DW/BI job market.

7.4 Data analysis

Different methods may be used to analyse textual data, e.g. content analysis, ethnography, grounded theory, and phenomenology (Hsieh & Shannon, 2005:1278). The researcher explored research methodologies in Chapter 2; interpretive coding and cross case analysis were selected as applicable methods for this study. Content analysis is a method that is widely used in a qualitative research (Hsieh & Shannon, 2005:1277).

According to Zhang and Wildemuth (2009a:3) qualitative data analysis requires a “set of systematic and transparent procedures”; thus, they recommend the following 8 step process:

7.4.1 Preparing the data

This step involves transcribing data into text, reading the transcripts, and choosing relevant content. According to Zhang and Wildemuth (2009a:3) researchers must decide if they want to use all the questions of the interview, whether they want to apply literal verbal translation or only summarise, and whether observations during the interview should be transcribed or not.

7.4.2 Define units of analysis

In this step the sections are identified as units and subunits.

7.4.3 Develop categories of scheme

This step involves developing coding schemes for inductive and deductive content analysis. Deductive reasoning, which generates concepts or variables from previous

studies or theory, is largely used. Inductive coding is also used where no theories are available and they only emerge through data examination and constant comparison (Zhang & Wildemuth, 2009a:3). This process is achieved by labelling relevant pieces of your data such as words, phrases, sentences, or sections. Labels can be about actions, activities, concepts, differences, opinions, processes, etc. Categories are created by bringing several codes together.

7.4.4 Test your coding schemes on a sample of text

This step involves coding of text especially based on the units that are defined in Step 2. Zhang and Wildemuth (2009a:3) encourage researchers to develop and validate the coding scheme early in the process by checking for consistency using sample data.

7.4.5 Code all the text

After coding rules have been tested and sufficient consistency is achieved, only then can coding be applied to all text.

7.4.6 Assess the coding consistency

To eliminate or minimise errors, coding consistency should be checked on the entire data set.

7.4.7 Draw Conclusions from the coded data

The ultimate goal of data analysis is to make sense of all themes or categories and their properties identified in the study. This is achieved by making inferences and reconstructions of the meanings derived from the data.

7.4.8 Report your findings

According to Zhang and Wildemuth (2009a:5) “qualitative content analysis uncovers patterns, themes, and categories important to a social reality”. The main objective is to uncover categories and how they are connected. Table 7-6 below presents the application of these steps in this study.

Table 7-6: 8 step process to perform content analysis (Zhang & Wildemuth, 2009a:308)

Step	Application to this study
1. Preparing the data	Data is prepared by transcribing the interviews and organising written responses into one uniform standard and format to be use for Atlas.ti software application. This step is explained in Section 7.4.1.
2. Define units of analysis	Each section of the interview is treated as a unit, as explained in Sections 7.4.2 & 7.4.3.
3. Develop categories of scheme	Code categories are developed based on the questions objectives (e.g. beneficiaries' code category is based on 'Who are the current DW/BI education beneficiaries'). Code development is presented in Section 7.4.3.
4. Test your coding schemes on a sample of text	Sample data is selected and coded and codes are tested in Section 7.4.4.
5. Code all the text	This step involves coding of text especially based on the units that are defined in Step 2.
6. Assess the coding consistency	Code consistency is done and presented in Section 7.4.6.
7. Draw conclusions on the coded data	Data analysis is conclusions are drawn in Section 7.5.
8. Report your findings	Guidelines are presented in Section 7.6.

7.4.9 Step 1: Preparing the data

The process of preparing data involves transcribing interviews. All written responses were also edited to remove any information that may identify participants or companies. Both interviews and written responses are organised or formatted to one uniform standard to be used in Atlas.ti software application. Participants are given an ID starting with letter P, and sequential numbers from 1 to 16. Each participant's response document is associated with the ID. Refer to Figure 7-1 as an example.

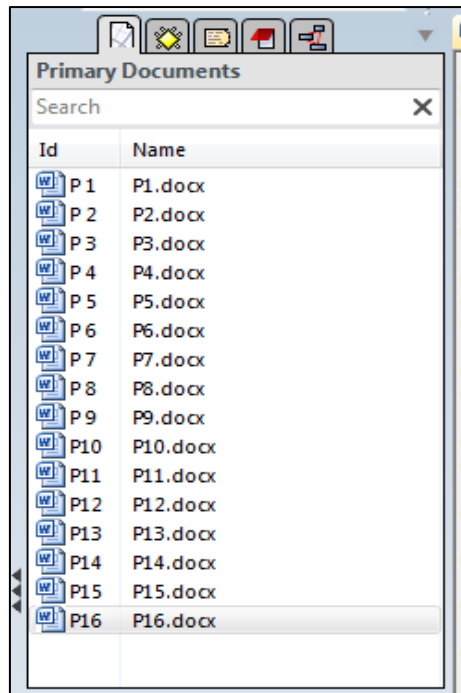


Figure 7-1: Participants' IDs from Atlas.ti

7.4.10 Step 2: Define the unit of analysis

The research interview contains three sections, as indicated in Section 7.2.2 of this chapter. Each section is treated as a unit, and subunits are also defined based on the question and the responses of participants.

7.4.11 Step 3: Develop categories and a coding scheme

This process involves labelling relevant pieces of data such as words, phrases, sentences, or sections. These labels are informed by actions, activities, concepts, differences, opinions, processes, etc. revealed in data. Categories are created by bringing several codes together after the establishment of relationships or close relationships. For example, Figure 7-2 below presents the streamlined codes-to-theory model for qualitative inquiry.

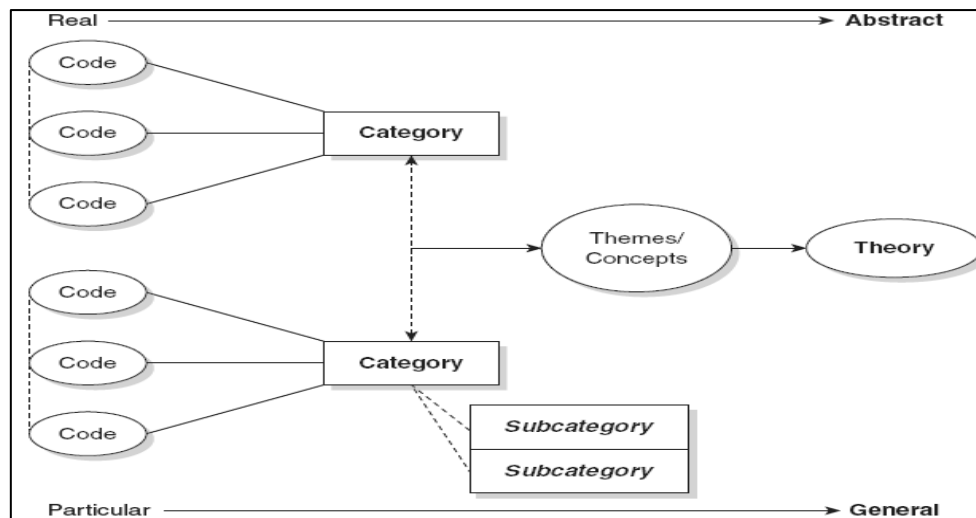


Figure 7-2: Streamlined codes-to-theory model for qualitative inquiry (Saldana, 2008:12)

The first step is to create codes in the text. The second step is to group codes into categories and subcategories. The last step is to formulate themes of concepts that will lead to theory. This process starts from left to right as depicted in the Figure 7-2.

7.4.12 Step 4: Testing the coding scheme on a sample of text

Question 3 in Section 2 of the interview is selected for the purpose of conducting unit testing. The responses of two participants (P10 and P11) are selected for this initial coding. Undirected coding is conducted on the two respective participants' responses to this question. Table 7-7 presents the participants' responses to Question 3 in the first column; the second column shows the relevant codes that are assigned to each response. P10's response is assigned four codes: [a] *Ideal beneficiary – Business*, [b] *Ideal affected – Business*, [c] *Current success measure - technical abilities*, and [d] *Ideal DW/BI purpose*. The response of P11 is assigned three codes: [a] *Ideal beneficiary – Business*, [b] *Ideal affected – Industry*, and [c] *Ideal affected - Students*.

For this study, distinction between business and industry is made. Business refers to end users of DW/BI in firms or organisations; industry refers to DW/BI developers of DW/BI solutions.

Table 7-7: Code scheme unit test on sample of participants (P10 and P11)

Participants' responses	Codes
Question 3 T: Who or what other stakeholders that ought to benefit from DW/BI education programmes? <i>P10: The business [a][b] that recruiting the DW/BI graduate they should be benefiting. Because right now what we are looking for is a real BI developer, but what we are getting is more towards a strong technical resource; meaning that they know the tools sets, they know the methodology reasonably well, but I think they lack business sense [c]. The real value is knowing where you fit in the business and using these tools to actually promote business [d]. And that's why I say business is not really benefiting from a guy like that. They understand the benefit of BI but they don't realise it.</i>	<i>[a] Ideal beneficiary – Business</i> <i>[b] Ideal affected - Business</i> <i>[c] Current success measure - Technical abilities (run queries, database connection, etc.)</i> <i>[d] Ideal DW/BI purpose - Value of an enveloped solution that pushes business forward.</i>
Question 3 T: Who do you think ought to be benefiting from a module in data warehousing in a 4th year IT programme? <i>P11:</i> <i>(1) The organisation [a] that employs the students.</i> <i>(2) Students [b] benefit because they become more marketable or more employable.</i> <i>(3) The customers that are being served by BI consulting houses. (Business) [c]</i>	<i>[a] Ideal affected – Industry</i> <i>[b] Ideal affected – Students</i> <i>[c] Ideal beneficiary – Business</i>

It is clear from this unit testing that participants may have provided rich information that addresses other aspects of the research. It is also evident that comparing and analysing each participant's response per question as a unit will narrow down codes. A participant's holistic view may not be well represented by unit testing. Using the respondent's entire response will lead to much wider code groups and a much better holistic worldview.

7.4.13 Step 5: Code all text

Each participant's response is coded and all codes created previously are reused. New codes are also created and added to the list until all participants' responses are analysed. This is consistent with cross case analysis as discussed by Seaman (1999:559). All 16 participants' responses are coded using textual analytic software

(Atlas.ti). There are 290 codes created. These are presented according to different themes in Section 7.5. Table 7-8 below presents an extract of undirected codes of all participants' responses.

Table 7-8: Alphabetical extract of undirected codes and counts for all participants' responses

Codes	Participants (P)																Totals
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Assessment component - No experience of it	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
Assessment component - Practical projects	1	0	0	1	1	1	1	0	1	1	1	0	1	0	1	1	11
Assessment component - Practical should be lined to industry	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2
Assessment component - A continuous assessment method	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
Assessment component - Tests	1	0	0	1	1	1	1	0	1	1	1	0	1	0	1	1	11
BI Process - Define success measures	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	2
BI Process - Choose technology to teach	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	2
BI Process - Conduct course comparison analysis	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
BI process - Define industry requirements	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	3
BI process - Define the module prerequisites and get better students	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
BI process - Draw insight from independent solutions in industry	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
BI process - Identify and involve relevant stakeholders	1	1	0	0	0	1	1	1	0	0	1	1	1	1	0	1	10

BI Process - Identify core skills needed	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
BI process - Incorporate big business models	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
BI process - Past students to review	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
BI process - Review periodically	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
BI process - Start with what we have	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	2

Undirected coding is applied to the participant's entire response and codes are created. Responses to questions that have the same codes are compared. As an example, Table 7-9 presents participants' responses to different questions in the first column, and the second column shows the relevant codes that are assigned to each response. The first response to Question 4 in Table 7-9 is from participant 10 (P10). Question 4 is explicitly asking about the purpose of DW/BI; however, the response of P10 covers other aspects of the study such as what to teach DW/BI students and the perspective of what the universities teach. This response has three distinct codes assigned: *[a] Current DW/BI purpose: To teach Kimball methodology*, *[b] What to teach: Business understanding?* and *[c] Perspective of what universities teaches: No interpretation or business case function in DW/BI solution*.

The response of P10 to Question 4 can be compared to a response of P4 to Question 18. Question 18 is an open-ended question about what the universities should teach the DW/BI students. The response of P4 to the same question is also presented and one unique code *[a] What to teach: Business understanding* is assigned.

The third row in Table 7-9 presents Question 16 and a response from P11 and the codes assigned to it. This is an open-ended question about the participant's perspective of what the universities are currently teaching. There is only one distinct code assigned to this response and that is *[a] Perspective of what universities teach: Transition from OLTP to DW/BI solution*.

Table 7-9: Undirected code scheme testing on participant entire response

Participant response	Codes
Question 4 T: According to you, what is currently the purpose of DW/BI module such that it is serving the interests of those who benefit from it? <i>P10: I can only speak from experience, because I am still involved with one university in terms of helping them with the input. But the core focus of that course initially was to teach the Kimball methodology [a]. This gradually evolved also to include Microsoft toolsets. It has also evolved to include third party tools sets to see what trends and analysis can be made. So that's what it does currently. The tools facilitate what you need to go through but it does not analyse the data for us. So I guess where I am getting at is that there is a missing leg of business understanding [b]. The problem you end up with a graduate that can create reports with numbers and graphics but when you ask them what does this report mean they tell you what is on the axis and what the measure is and that is nothing. That's just the data, that's just the information. There is no interpretive, not business case function. There is simply this is the data. So there is no value add [c]. The other thing that I think is missing is cross subjects area analysis where we don't stick to one box, we try to say we have promotion data on one hand on the other we have sales data on the other and how do I make it tie up? And that again it goes with the analytical concept of saying I need to explain this analytical phenomenal that I saw in the data and where did I see it. I can't just say these are numbers. [a]</i>	<i>[a] Current DW/BI purpose: To teach Kimball methodology</i> <i>[b] What to teach: Business understanding</i> <i>[c] Perspective of what universities teach: No interpretation or business case function in DW/BI solution.</i>
Question 18 T: Please tell us what you think universities should teach DW/BI students? <i>P4:</i> <i>a) Deep understanding of tools and how to use them</i> <i>b) Good understanding of data warehousing</i> <i>c) Business understanding [a]</i> <i>d) Documentation</i> <i>e) Source to target mapping</i> <i>f) Dimensions matrixes</i>	<i>[a] What to teach: Business understanding</i>
Question 16 T - What is your overall perspective of what universities teaches in DW/BI modules? <i>P11: Universities teach a transition from database management into what can also be done once you have data, such as building a data warehouse. This is a transition from transactional database into BI. They don't teach BI or DW extensively. [a]</i>	<i>[a] Perspective of what universities teach: Transition from OLTP to DW/BI solution.</i>

Different parts of the text can be compared and similarities can be drawn when using the participant's entire response as the unit of analysis.

7.4.14 Step 6: Assess the coding consistency

It is imperative that coding should be tested for consistency; more so, it has already been argued that the participant's entire response should be used instead of his/her individual answers. Codes consistency check includes the following:

1. The first codes are created, beginning with participant 1 (P1). Codes created for P1 are used for analysis of the answers of participant P2, and so forth. This incremental process that ensured that unique codes are created and reused is necessary for consistency purposes.
2. The incremental code process is reversed to ensure that all codes that are created later are used in responses that were coded at an earlier stage.
3. The process of creating codes is guided by the objective of the interview. Each of the 23 questions has a purpose, and that purpose is used to create codes. For example, the purpose of Question 3 (Who do you think is currently benefiting from a module in data warehousing in a 4th year IT programme?) is to know who the beneficiaries of DW/BI education are. A category called *Current beneficiary* is created. This process ensures that a focus on the question objectives is not lost even if the participants gave rich responses that may address other aspects of the research. The second benefit of this process is that codes can be used across all participants, thereby ensuring consistency.
4. Codes that refer to similar ideas are grouped as one category. For example, the excerpt presented below is the response of P3 to Question 3 about who the current DW/BI beneficiaries are. Three codes: [a] *Current beneficiary – Universities*, [b] *Current beneficiary – Students*, and [c] *Current beneficiary -*

Industry are assigned. The three codes are grouped into one category Current beneficiaries.

Question 3 T - Who do you think is currently benefiting from a module in data warehousing in a 4th year IT programme? <i>P11: I think <u>universities</u> [a] are more benefiting that the students should. And why I am saying that because some BI courses a chapter a chapter or not covered extensively. Most relevant topics are not covered at all. <u>Students</u> [b] don't have a clear picture of what actual is a BI even though basics they will have. Universities are getting paid for the courses. <u>Organisations</u> [c] they somehow benefit because students have the basics</i>	<i>[a] Current beneficiary - Universities</i> <i>[b] Current beneficiary - Students</i> <i>[c] Current beneficiary - Industry</i>

5. Codes are created for other aspects that were not so explicitly asked for in the interviews, but are relevant to the study. A theme called *Other concerns* is created.
6. All text is read through over and over again to ensure that no relevant information is overlooked. Text that is not relevant to the study is not coded.

7.5 Draw conclusions from the coded data

This section presents the data analysis, and conclusions drawn from data, as prescribed in Step 7 of Zhang and Wildemuth (2009a:3). Code families are discussed and conclusions are drawn for each family; these are presented after the discussion. Figure 7-3 below presents code families created in Atlas.ti software. Two types of hermeneutic interpretation are used throughout the chapter. The first type is the interpretation of faith. This means that data is interpreted from the perspective of stories as told by participants (Ruthellen, 2004:193). The second interpretation type is the interpretation of suspicion that will only be applied in this chapter when data raises the researcher's suspicions. This approach is used for deeper or hidden meaning exploration (Ruthellen, 2004:210).

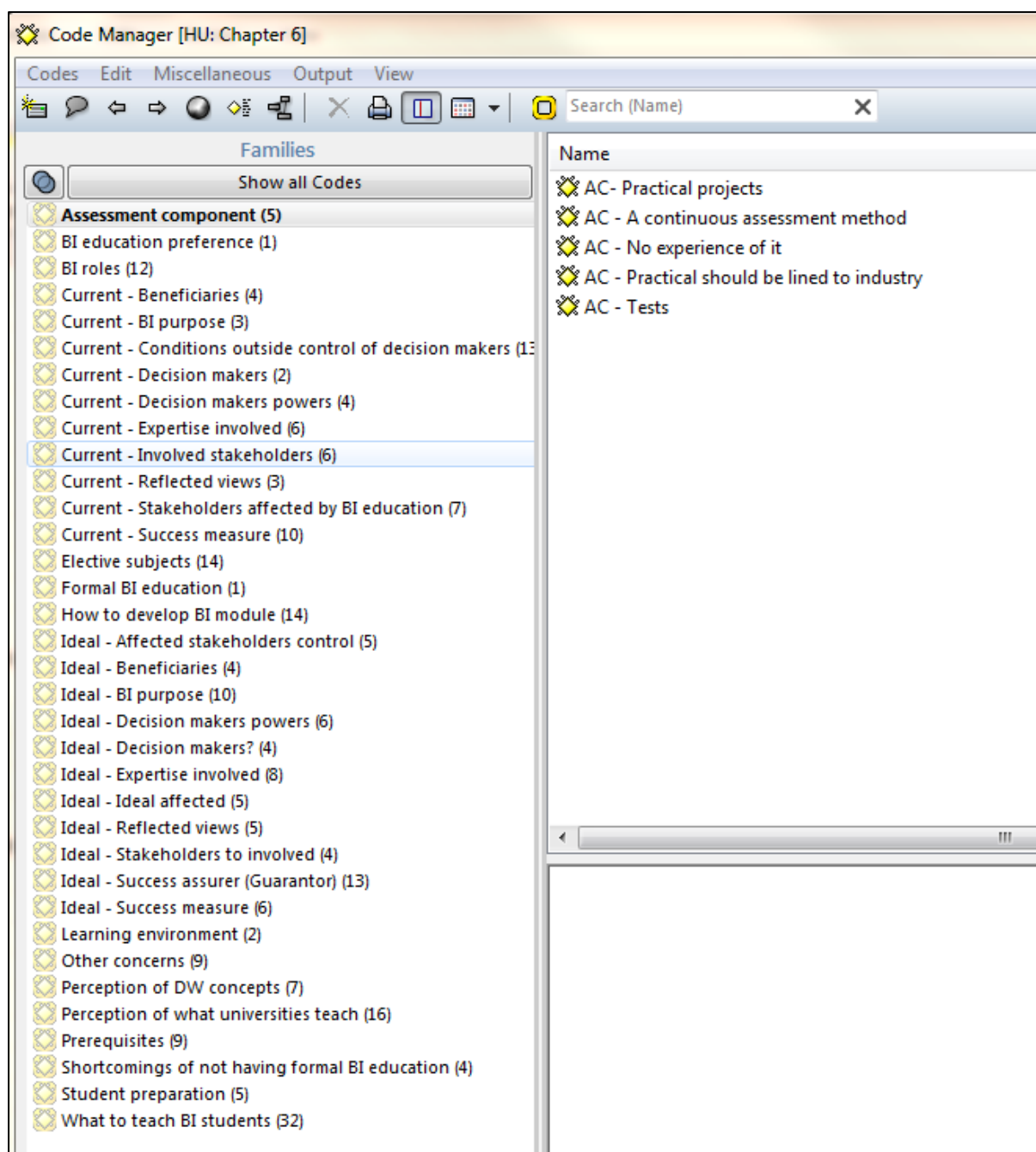


Figure 7-3: Code families

The next section presents themes created from codes and code families. Throughout the chapter, participants with a 0 in the count column indicate that no response was given that addresses that particular theme.

7.5.1.1 Theme 1: DW/BI roles

Question 1 of the interview inquires which roles participants played in their DW/BI careers. Various roles are identified from participant's responses and codes are

created for each. A *DW/BI roles* theme is formed as an outcome of the coding process. Table 7-10 below presents codes per document for the *DW/BI roles* code family. The sum of the occurrences of each of these codes is presented in the 'Totals' column.

Table 7-10: Count of codes per document for code family: *DW/BI Roles*

DW/BI roles	Participant (P)																Totals
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
BI developer	1	1	1	1	0	1	0	0	0	1	0	1	0	1	0	1	9
BI architect	0	0	1	0	0	1	0	0	0	1	0	0	0	1	0	0	4
BI analyst	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	3
ETL developer	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	3
Lecturer	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	3
BI consultant	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	2
None	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	2
Team lead	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	2
BI manager	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
Data modelling	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
Domain specialist	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Quality assurance	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
Totals	2	1	3	2	1	2	1	1	3	5	1	2	1	4	1	2	32

It is evident that the *DW/BI developer* role is played mostly by participants, because it has the highest total sum number of 9. This is because the *DW/BI developer* role is all inclusive of most sub-roles in the realm of DW/BI. For example, below is a code excerpts from P2 and P3 about DW/BI roles. P2 has played an end-to-end role in DW/BI. End-to-end (general) means involvement in the entire DW/BI development cycle (data extraction, staging, transformation, loading, and presentation). Similarly so, participant P3's response to question 10 says that a DW/BI developer in South Africa should know the 'breadth' of DW/BI.

Question 1

T: What is your current role in DW/BI education/industry? What other roles have you played in DW/BI education and how many years of experience associated with each role?

P2 - I've always been a BI developer that is involved in end to end solution. If there is anything that needs to be done from administration, etc. I've done it all.

Question 10

T: What kinds of expertise do you think ought to be consulted or involved in the development of DW/BI module?

P3 - In the SA context BI is understood to know the whole breath of it. In Europe and USA you have specific roles like ETL developer, report developer, etc. You need to get each specific BI role like ETL, DBA. Etc. and academic expertise is important. These expertise are not all involved currently. Companies must push their internship programme knowledge structure back into universities to complement what the universities are offering. Start the internship early on at the university level.

Professionals in the DW/BI field are not always involved in a general role on projects. Sometimes projects are segmented; thus, specific roles are required to be played. For example, data transformation will require an ETL developer role. As depicted in Table 7-10, DW/BI architecture, analyst, and ETL developers are some of the roles played by DW/BI professionals.

7.5.1.1.1 Conclusions drawn on DW/BI roles played.

Most professionals play a *DW/BI developer* (general) role in the industry (i.e. their skills cut across the DW/BI life cycle). This evidence supports the guidelines that are presented in Section 5.6 of Chapter 5. The guidelines say that students should be taught the entire DW/BI life cycle to have the knowledge and skills (basic and advanced). Even though students are generally prepared for a DW/BI developer role, they can later choose to specialise in specific areas of the DW/BI life cycle. The DW/BI module should prepare students with basic and advanced knowledge and skills; they should be sufficiently prepared so that they are able to fulfil any function on a DW/BI project (DW life cycle).

7.5.1.2 Theme 2: DW/BI education preference?

The motivation for Question 2 of the interview is to establish whether participants had any formal DW/BI education from universities or not. Eight out of sixteen

responded with a ‘Yes’. One participant did not give a response. Table 7-11 below presents participants and their responses. Participants without formal DW/BI education were asked if they would have preferred to have had it. A code family for *Would you have preferred to have DW/BI education?* is created.

Table 7-11: Count of codes per document for code family: *DW/BI education preference*

Code family	Participants (P)							
	2	3	4	9	11	12	14	15
Would you have preferred to have DW/BI education	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Participants stated various reasons why they would have preferred to have had formal DW/BI education. Codes are created for the *Why I would have preferred DW/BI education?* code family. Table 7-12 below presents codes per document for *Why participants prefer DW/BI education*. The sum of the occurrences of each code is presented in the ‘Totals’ column. From the codes presented, *Professionals learn DW/BI the hard way on the job* is the most stated reason with total sum of 3. For example, the following code excerpt from the response of P4 is an example of the code *Professionals learn DW/BI the hard way on the job*.

Question 2

T: Reasons why DW/BI education would have been good or bad for you?

P4: It would have a put a lot of things into perspective, and a lot sooner. I would have progressed a lot quicker. So starting without any background or any knowledge and not knowing how business works, everything had to be learned on the job at the client and under pressure. It would have been very nice to have a little bit of background as to what is the whole purpose of doing this.

Other codes such as *Help progress fast in DW/BI career*, *Standardised common basic knowledge*, and *To get a better perspective* complete the list of the most stated reasons with total sum of 2 each.

Table 7-12: Count of codes per document for code family: *Why participants prefer DW/BI education*

Why I would have preferred DW/BI education?	Participants (P)																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Totals
Professionals learn DW/BI the hard way on the job	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	3
Help progress fast in DW/BI career	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	2
Standardised common basic knowledge	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2
To get a better perspective	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	2
It would have prepared me better	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
Totals	0	1	2	3	0	0	0	0	1	0	1	0	0	1	1	0	10

The data presented in Table 7-12 is consistent with that of Table 7-4. It represents the views of participants (P2, P3, P4, P9, P11, P14, and P15) with no DW/BI education. It is expected that participants P5 and P7 will have zeros since they have played no role in DW/BI. Other participants (P1, P6, P8, P10, and P16), with the exception of P12 and P13 who did not give a response to Question 2, have zeros because they have formal DW/BI education.

7.5.1.2.1 Conclusions drawn from DW/BI education preference

The motivation behind Question 2 is to better understand the perspectives of the participants; however it addresses a very important aspect. It is evident that DW/BI professionals who never had formal DW/BI education wish they had it. Their experience proved that it is easier and better to learn DW/BI at university than on the job because it will help them (students) to progress faster in their career. Formal DW/BI education also helps standardise general basic DW/BI knowledge thus

leading to best practice and a better perspective of the overall DW/BI picture. At high level this evidence will inform the universities on whether to offer DW/BI modules as part of their programmes or not; and shows that there is a demand for DW/BI education. The guidelines presented in Chapter 5, Section 5.6, are supported by this evidence by answering: *What to teach in a DW/BI module?* It is advisable that DW/BI professionals have formal DW/BI education.

7.5.1.3 Theme 3: Beneficiaries

Question 3 asked participants about beneficiaries of DW/BI education. Various beneficiaries are identified and coded from responses. Tables 7-13 and 7-14 below present codes for *Current/Ideal Beneficiaries'* code families. The sum of the occurrences of each of these codes is presented in the 'Totals' column of each table. *Students* are the most stated *Current Beneficiaries* of DW/BI education as reflected by the total sum of 13. *Industry* (BI vendors and other companies that are offering DW/BI services) are the second most stated beneficiaries. *Universities* are the third most stated *Current Beneficiaries*.

Table 7-13: Count of codes per document for code family: *Current DW/BI beneficiaries*

Current Beneficiaries	Participants (P)																Totals
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Students	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	13
Industry	1	0	1	1	0	0	0	1	0	0	1	0	1	0	0	0	6
Universities	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	3
Not many are exposed to DW/BI education	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
Totals	2	1	1	2	0	1	1	2	2	1	3	2	2	1	1	1	23

Table 7-14 below presents five codes for the *Ideal Beneficiary* code family. *Students* and *Industry* are the most stated stakeholders with total sums of 8 and 7 respectively. *Industry* refers to thought leaders of DW/BI industry such as vendors,

consulting companies and other in-house DW/BI experts (developers of DW/BI solutions). For example, the following code excerpt from participant P3 indicates that there is a clear distinction between *Business* and *Industry*. Thought leaders here are referred to as industry (experts)

Question 6

T: Who do you think ought to have decision powers to make a change to improve DW/BI modules?

P3: It is the responsibility of both business that uses BI and the universities (faculty members). And those thought leaders (experts) in BI field. Ask professors to invite us.

Business (end-users of the DW/BI solutions) is the second *Ideal Beneficiary* with a total sum of 6. *Business* refers to owners or executives that take business decisions; they know and define the requirements. The following code excerpt from participants P6 and P10 gives evidence of who is coded as *Business*. Business executives, such as a chief information officers (CIOs), managers, etc. are viewed as representatives of *Business* in the study.

Question 6

T: Who do you think ought to have decision powers to make a change to improve DW/BI modules?

P9: The CIO, executives, and senior managers of big corporates because they are the major key beneficiaries and they know what the expectations are. Nobody else know no the expectations except business.

Question 4

T: who or what other stakeholders that ought to benefit from DW/BI education programmes?

P10: The business that recruiting the BI graduate they should be benefiting. Because right now what we are looking for is a real BI developer, but what we are getting is more towards a strong technical resource; meaning that they know the tools sets, they know the methodology reasonably well, but I think they lack business sense. The real value is knowing where you fit in the business and using these tools to actually promote business. And that's why I say business is not really benefiting from a guy like that. They understand the benefit of BI but they don't realise it.

Table 7-14: Count of codes per document for code family: *Ideal DW/BI education beneficiaries*

Ideal Beneficiary	Participants (P)																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Totals
Students	1	0	1	1	0	0	1	0	0	1	1	0	1	1	0	0	8
Industry	1	0	1	1	0	0	0	0	1	0	0	1	1	1	0	0	7
Business	0	0	0	1	0	1	1	0	1	1	1	0	0	0	0	0	6
Business analyst	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
BI professionals	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Totals	2	1	2	3	0	2	2	0	2	2	2	1	2	2	0	0	23

7.5.1.3.1 Conclusions drawn on DW/BI beneficiaries

This data show a component of CSH and SSM respectively. CATWOE is a SSM tool that a researcher can use to define the scope, components, identify, and categorise all stakeholders of the system being analysed (Checkland & Poulter, 2006:220). The first component of CATWOE is customers (C). *Students* and *Industry* should be the customers of DW/BI education. SSM is only concerned with the ideal world; hence, current DW/BI education customers are ignored.

Identifying a client of a system under study is a purpose of CSH boundary questions. *Students*, *Industry*, and *Universities* are the *Current Beneficiaries* (clients) of DW/BI education. *Students* and *Industry* are the *Ideal Beneficiaries* (clients) of current DW/BI education. The researcher suspects that participants P1, P9, and P11 identified *Universities* as an *Ideal Beneficiary*, because they play(ed) the role of lecturer.

7.5.1.4 Theme 4: DW/BI module purpose

Question 4 asked participants about the purpose of DW/BI education. Codes are created for various answers that identify the DW/BI *purpose*. Table 7-15 and Table 7-16 below present codes for *Current/Ideal DW/BI module purpose* code families.

The sum of the occurrences of each of these codes is presented in the 'Totals' column in each table. *Teaching DW/BI technology (tools)* and *Training and equipping students with DW/BI knowledge and skills* are the most stated *Current DW/BI purpose* with total sum of 4 occurrences each. *Teach students accurate reporting* and *Teach students Kimball methodology* are second most stated with a sum of 2 each.

Table 7-15: Count of codes per document for code family: *Current DW/BI education purpose*

Current DW/BI purpose	Participants (P)																Totals
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Teach DW/BI technology (tools)	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	4
Train and equip students with DW/BI knowledge and skills	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	4
Teach students accurate reporting	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	2
To teach Kimball methodology	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	2
The value is not there because students still needs to be trained	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
To add value to business	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
To answer business questions	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
Value is not seen because only a few have formal DW/BI education	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Totals	1	1	1	1	0	1	1	1	1	2	1	1	1	1	1	1	16

Table 7-16 presents codes for *Ideal DW/BI purpose code* family. Teaching value of an enveloped solution that pushes business forward emerged as the most stated *Ideal DW/BI purpose* with an occurrence of 8. For example, the following code excerpt for participants P7 and P9 for Questions 4 and 5 supports this stand.

Question 4

T: According to you, what ought to be the purpose of DW/BI module such that it is serving the interests of those who benefit from it?

P9: “..... The value is in the thinking that goes into it and that is what universities should be teaching.....”

Question 5

T: What do think ought to be the measures of success (or improvement) for DW/BI module

P7: “.... The outcome of BI is to drive business to be more profitable/reduce costs where possible.....”

Table 7-16: Count of codes per document for code family: *Ideal DW/BI education purpose*

Ideal DW/BI purpose	Participants (P)																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Totals
To teach value of an enveloped solution that pushes business forward	0	1	1	1	0	1	1	1	1	1	1	0	0	0	1	0	10
Make students DW/BI ready for industry	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	3
BI is to drive business to be more profitable/reduce costs where possible	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
Knowing your purpose in DW/BI industry	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Should be to teach more than Kimball books (Industry trends)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
To teach business models (business)	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
To teach the value of data, the ethics around using the data	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
Ideal DW/BI purpose - To answer business questions	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
Totals	1	1	2	1	0	1	2	1	1	3	1	1	1	1	1	1	19

7.5.1.4.1 Conclusions drawn on DW/BI purpose

Teaching DW/BI tools, training, and equipping students with DW/BI knowledge and skills are the current main purpose of DW/BI education, according to the participants. Ideally the main purpose should shift towards teaching value of an enveloped

solution that pushes business forward. Further, this ideal shift will help to make DW/BI students ready for industry. The two modes, *Current* and *Ideal*, pair the second component (purpose) of CSH boundary questions. This data also show the third component of CATWOE tool under SSM, i.e. the transformation (T). The transformation (purpose) of DW/BI education ideally is to transform students from students who do not know how DW/BI can propel the business, to students who do know how DW/BI can be used to propel the business.

7.5.1.5 Theme 5: Decision makers

The purpose of Question 5 is to identify the decision makers in DW/BI education. Codes are created for various stakeholders that are identified from participants' responses. Table 7-17 and Table 7-18 below present the *Current* and *Ideal Decision maker* code families respectively. The total sum of the occurrences of each of these codes is presented in the 'Totals' column in each table. Currently *Universities (faculty members)* are the most reported decision powers in DW/BI education, with a total of 12 references. *Industry* also has the decision powers to some extent, with a total sum of 4. Data show that there is a strong perception that *Universities* are the most involved decision maker, and industry is involved to some extent. For example, the following code excerpt for Question 6 of participant P6 shows this perception. This code excerpt also shows the perception that some industries will only get involved with the agenda of recruiting talent, instead of contributing towards the growth of DW/BI education.

Question 6

T: Who do you think currently has the decision powers to make a change to improve DW/BI modules?

P3: I don't know because I never studied it at university level. I think academics. The current BI consulting companies they get involved only to recruit instead of building this educational space. They lessen their recruiting and that should not be the point, the point is to build the industry and professions.

Table 7-17: Count of codes per document for code family: *Current decision makers*

Current Decision makers	Participants (P)																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Totals
Universities (faculty members)	1	1	1	1	0	1	1	1	0	1	1	0	1	1	1	0	12
Industry (to some extent)	1	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	4
Totals	2	1	1	1	0	2	1	2	1	1	1	0	1	1	1	0	16

Table 7-18 below presents the ideal codes for decision makers. The most reported *Ideal Decision makers* in DW/BI are *Industry (experts)*, *Universities (faculty members)*, and *Business*, with total sums of 9, 8, and 3 respectively. There is strong support of *Industry* and *Universities*. For example, this is shown in code excerpt below from Participant P15.

Question 5

Who do you think ought to have decision powers to make a change to improve DW/BI modules?

P15: I think universities should still have the decision to improve BI modules without ignoring the input from students and the demands of stakeholders in the BI job market.

Table 7-18: Count of codes per document for code family: *Ideal decision makers*

Ideal Decision makers	Participants (P)																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Totals
Industry (experts)	1	1	1	1	0	1	0	1	0	0	1	0	1	0	0	1	9
University (faculty members)	0	0	0	0	0	1	0	1	0	1	1	0	1	1	1	1	8
Business	0	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0	3
Students	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
Totals	1	1	2	1	0	2	0	2	1	1	2	0	2	1	3	2	21

7.5.1.5.1 Conclusions drawn on decision makers

Universities and *Industry* are the *Current Decision makers* to improve DW/BI education. Ideally the third decision maker is *Business* because they provide business requirements. The identified stakeholders for both modes (*Current* and *Ideal*) are the sources of power in the DW/BI education under the CSH methodology. Under CATWOE, *Universities*, *Industry*, and *Business* are the ideal owners (O) of DW/BI education. These are the stakeholders that can change or stop DW/BI education.

7.5.1.6 Theme 6: Conditions outside control of decision makers

From the previous theme on decision makers we know that industry, universities, and business are key decision makers in DW/BI education. Codes are created for *Conditions that are outside the control of decision makers*. Table 7-19 below presents codes for *Current Conditions outside control*. The sum of the occurrences of each of these codes is presented in the 'Totals' column in each table. The most common conditions that universities cannot control relates to DW/BI technology and its trends (hardware and software licensing costs) – with 5 occurrences. For example, the following code excerpt for participant P1, who has experience in the role of a DW/BI lecturer, shows that universities have challenges with software licensing, data, and industry input.

Question 8

T: What resources and conditions do you think currently is part of DW/BI education environment but are outside the control of decision maker?

P1: First it will be the software licensing issues for academic institutions; second will be data that we can use for students projects. We need real data from industry to use. And third but lastly the industry input in terms of expertise.

The second most stated condition relates to industry input of real life problems, with a total count of 4. Other conditions, such as *Universities can't control students' employability*, and *Industry can't control programme requirements*, are the third – both with total count of 2.

Table 7-19: Count of codes per document for code family: *Current conditions outside control*

Current conditions outside control	Participants (P)																Totals
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Universities can't control DW/BI technology and its trends (hardware and software licensing costs)	1	0	0	0	0	0	1	0	0	1	1	0	0	0	1	0	5
Universities can't control industry input of real life problems	0	1	0	0	0	1	0	0	0	1	0	0	1	0	0	0	4
Industry can't control programme requirements	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2
Universities can't control students employability	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	2
Industry has no control how academic programmes should run	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
Lack the practical component	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
Universities can't control DW/BI knowledge on Internet	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
Industry can't control modules to complement DW/BI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
Universities can't control packaging of DW/BI programmes such that it adds value	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Universities can't control real data	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Universities can't control technical expertise	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
Universities can't control cost of hardware and software	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
What DW/BI environment students will be exposed to in the industry	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
Totals	3	1	1	1	0	2	2	0	0	3	3	0	1	2	2	1	22

7.5.1.6.1 Conclusions drawn from conditions outside the control of decision makers

The three major conditions that universities cannot control are: technology, industry input, and students' employability. DW/BI technology trends, and the cost of acquiring it, are outside the control of universities. Also, universities cannot control industry input of real life problems to better prepare students. Universities may request such an input, but it is a condition that cannot be enforced. Similarly so, as much as industry can advise universities, they cannot control programme requirements. Universities have the ultimate say in what the programme requirements should be.

The DW/BI educations' sources of power (BI decision makers identified in Theme 5 for CSH) operate under the decision environment that is described in both modes (*Current* and *Ideal*). Universities cannot control technology, industry input, and students' employability. The three conditions are also identified as the Environment (E), which is component of the CATWOE for SSM.

7.5.1.7 Theme 7: Decision makers' powers

The aim of Question 7 is to identify the powers of decision makers identified in Theme 5. Codes are created for *Decision makers' powers*. Table 7-20 and Table 7-21 below present the *Current* and *Ideal* *Decision makers' powers* respectively. There are 5 codes that reflect the *Current* *Decision makers' powers*. Universities is the most stated to have control powers over academic components, with 9 occurrences.

Table 7-20: Count of codes per document for code family: *Current decision maker's powers*

Current Decision makers' powers	Participants (P)																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Totals
Universities control the academic component (standard and course content	1	0	1	0	0	1	1	1	0	1	1	0	1	1	0	0	9

requirements)																	
Universities control the teaching aspect	0	1	0	1	0	0	0	0	0	0	0	0	0	0	1	0	3
Universities have control over textbooks and other materials	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	2
Business must control course content requirements	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
Choice of technology	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
Totals	1	1	1	1	0	2	1	1	1	1	1	0	1	1	2	1	16

Table 7-21 presents 7 *Ideal decision makers' powers* codes. *Industry should determine DW/BI standards* is the most stated reasons with highest total occurrence of 4. Business and Universities ought to control data and requirements, as well as course content, with 2 occurrences each.

Table 7-21: Count of codes per document for code family: *Ideal decision makers' powers*

Ideal Decision makers' powers	Participants (P)																Totals
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Industry should determine DW/BI standards	0	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	4
Business must control data and requirements	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	3
Universities should control course content	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	2
We don't have a body that recognises what DW/BI professional is	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2
The influx intake	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

of better students																	
Universities control choice of technology	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
Industry should say what subjects that must complement DW/BI	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Totals	1	2	2	2	0	2	0	1	0	1	1	0	0	1	0	0	14

7.5.1.7.1 Conclusions drawn from Decision makers' powers

Data reveal that universities currently have more control over the academic component of DW/BI education, teaching aspects, as well as textbooks and other material used. However, ideally, industry should determine the DW/BI standards, and business must control the requirements going into the DW/BI modules. In the process of identifying decision makers' roles, the academic process should be preserved. For example, this is shown the following code excerpt from participant P9.

Question 8

T – What components (activities, resources, constraints) of DW/BI module do you think ought to be within the control of the decision maker?

P11: Industry does not have control of how the academic programmes should run. And in a way it is good because we would lose the picture (purpose) of academic process. The can only tell the institutions what they need but they never change the fundamentals of academic education.

Industry and Universities cannot change the pace at which technology changes. Its either you adapt or you done. You can't control that.

The CSH methodology seeks to identify the resources or conditions that decision makers have control over. The resources or conditions that sources of power (decision makers) currently have are such that universities have more control of over the academic component of DW/BI education, teaching aspects, as well as textbooks and other material used. However, industry and business should have control over DW/BI standards and requirements for DW/BI modules.

7.5.1.8 Theme 8: Affected stakeholders

Participants were asked about stakeholders that are affected by DW/BI education in Question 12. Affected stakeholders are identified and codes are created for each. Table 7-22 and

Table 7-23 below present the *Current and Ideal stakeholders* that are affected by DW/BI education. Table 7-22 presents four stakeholders that are currently affected by DW/BI education. *Business* is the most stated affected stakeholder by DW/BI education, with 9 occurrences. For example the following code excerpt for participants P7, P9, and P10 shows that *Business* is affected. Most participants (9) identify *Business* as affected stakeholders; *Industry* and *Students* are the second most, with 7 occurrences each.

Question 12

P7: What do you think currently are the relevant stakeholders affected by DW/BI module and what is their role?

P7: From a business point of view it is affecting the decision makers. The executives are affected.

P9: Business owners, decision makers are affected by BI education.

P10: students, business, consulting houses, and the teams they have to work in.

Table 7-22: Count of codes per document for code family: *Current affected stakeholders*

Current affected stakeholders	Participants (P)																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Totals
Business	1	0	1	1	0	1	1	1	1	1	0	0	0	0	1	0	9
Industry	1	1	0	0	0	0	0	1	0	1	1	0	1	1	0	0	7
Students	0	1	0	0	0	1	0	0	0	1	1	0	1	1	0	1	7
BI professionals	0	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	4
Totals	2	3	2	2	0	2	1	2	1	3	2	1	2	2	1	1	27

Table 7-23 presents four Ideal stakeholders that ought to be affected by DW/BI education. *Business* is the most stated affected stakeholder with 10 occurrences.

Students are the second most stated affected stakeholders with total 9 occurrences. And *Industry* is the third most stated stakeholder with 8 total counts. As an example, the following code excerpt for participant P2, P3, and P13 show that Business, faculty members (e.g. lecturer), Students, Industry, and DW/BI professionals should ideally be affected.

Question 12

T: What do you think ought to be the relevant stakeholders affected by DW/BI module and what is their role?

P2: It's the students, the lecturer, its person hiring the students, and it's the professional. BI affects all.

P3: Universities, business people and BI leaders.

P13: Students - developing to enter job market, Industry – employing students ready to enter job market, and University and lecturer – delivering students who are sought after in industry

Table 7-23: Count of codes per document for code family: *Ideal affected stakeholders*

Ideal affected stakeholders	Participants (P)																Totals
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Business	1	1	1	0	0	1	1	1	1	1	0	1	0	0	0	1	10
Students	0	1	0	0	0	1	0	0	0	1	1	1	1	1	1	1	9
Industry	1	1	0	1	0	0	0	1	0	0	1	0	1	1	1	0	8
BI professionals	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	2
Totals	2	4	1	2	0	2	1	2	1	2	2	2	2	2	2	2	29

7.5.1.8.1 Conclusions drawn from affected stakeholders

Stakeholders that are currently affected by DW/BI education are business, industry, students, and DW/BI professionals. Business is affected because it is the ultimate customer that will use a DW/BI solution. Industry, such as consulting firms, is affected because they absorb the students with the aim of delivering service to businesses. And lastly, students are affected because they are the main products in the DW/BI education put through to industry. Ideally business, students, industry, and DW/BI professionals should be the affected stakeholders of DW/BI education.

From the perspective of CSH, Ulrich and Reynolds (2010b:243) define the affected stakeholders as those affected but not involved. However, in this study participants were not asked to differentiate between stakeholders that are involved and affected, and those that are not involved but affected. The assumption is that, in the inception of the study, participants would not know who are involved and who are not.

Business, students, industry, and DW/BI professionals are customers (C of CATWOE) that are affected by DW/BI education transformation under the SSM methodology. These stakeholders receive the output of the transformation (improvement or change). Industry, students, and business were already identified as beneficiaries, and thus showed as the affected.

7.5.1.9 Theme 9: Affected stakeholders' control

Affected stakeholders were identified in the previous theme. Each stakeholder identified has, or ought to have, some form of control over or input into DW/BI education. Table 7-24 below shows that universities currently have more control over DW/BI education. For example, the following code excerpt for participant P2 is a view of an outsider without formal DW/BI education. Participant P2 says that universities control the academic aspect of DW/BI education.

Question 13

T: What control do you think the affected stakeholders currently have in DW/BI module development or education?

P2: Currently I feel that the academics are control the BI education.

Table 7-24: Count of codes per document for code family: *Current affected stakeholders' control*

Current affected stakeholders' control	Participants (P)																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Totals
Universities - control academic aspect	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0	1	4

Business - must have input into course content	0	0	0	0	0	1	0	0	1	0	1	0	0	1	0	0	4
Industry - have no control	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	3
Students - must have input into course content	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
Totals	1	1	1	1	0	2	1	0	1	0	1	0	1	1	0	1	12

Table 7-25 below presents codes per document for the family *Ideal Affected stakeholders' control*. Ideally business and industry ought to have more control of DW/BI education by having input in terms requirements. Universities should have control on the teaching aspect.

Table 7-25: Count of codes per document for code family: *Ideal affected stakeholders' control*

Ideal affected stakeholders' control	Participants (P)																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Totals
Business must define business requirements that goes into DW/BI	0	1	1	1	0	1	0	0	0	1	0	0	0	0	0	1	6
Industry should define business requirements	1	0	1	0	0	0	0	1	0	0	1	0	1	0	1	0	6
Universities control teaching and academic aspect	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	3
BI	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	3

professional control practical projects																	
Student must give input on course content	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
Totals	1	3	2	1	0	1	0	2	0	1	1	0	2	1	1	3	19

7.5.1.9.1 Conclusions drawn from affected stakeholder's control

Stakeholders should be identified first, and their roles should be clearly defined, in order to identify how they are affected. Business and industry are the affected; they have to employ students after completion of the module. Universities are affected because they control and deliver the teaching aspects of DW/BI education. This data are suspicious to the researcher because participants P2, P3, P13, and P16 did not identify universities as a current affected stakeholder in Theme 8 (*Affected Stakeholder*) in Table 7-22. However, in Table 7-24, universities are identified as having the powers to control the academic aspect, and thus making the universities one of the affected stakeholders. Lastly, students are affected because they are the receivers of DW/BI education; they move through the DW/BI education pipeline and into the industry to become DW/BI professionals.

The paragraph above describes the powers associated with the decision makers. These powers define component 5 (resources are or ought to be within the control of decision makers) in CSH. Also in SSM, the above mentioned powers form part of the CATWOE model that is the environment (E) component. These are constraints, such as regulations, politics, budget, etc. that outline environment.

7.5.1.10 Theme 10: Involved stakeholders

Participants were asked to identify stakeholders that must be involved in the development of DW/BI modules. Codes are created for various stakeholders that are identified. Table 7-26 and Table 7-27 below present current and ideal codes for the

Current Stakeholders involved family. *Universities* are the current most stated stakeholders, with a total occurrence of 12.

Table 7-26: Count of codes per document for code family: *Current stakeholders involved*

Current Stakeholders' involved	Participants (P)																Totals
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Universities	1	1	1	1	0	1	1	1	0	0	1	1	1	1	1	0	12
Industry (to some extent)	0	0	0	1	0	1	0	0	0	1	1	0	1	0	0	0	5
BI consulting companies	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Internet and other DW/BI bodies	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
Past students	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
Totals	2	1	1	2	0	2	1	1	0	1	2	1	3	2	1	0	20

Industry is the second most stated involved stakeholder, with a total occurrence of 5. Some participants report that the involvement of industry is not satisfactory. This could be caused by various reasons. For example, below is code excerpt of Question 6 from participants P3 and P9. Participant P3 suggests that the involvement of industry goes as far as recruiting students, instead of adding to and growing the DW/BI educational space.

Question 6

T: Who do you think currently has the decision powers to make a change to improve DW/BI modules?

P3: I don't know because I never studied it at university level. I think academics. The current BI consulting companies they get involved only to recruit instead of building this educational space. They lessen their recruiting and that should not be the point, the point is to build the industry and professions.

Participant P9 was asked a follow up question in support of Question 6's response. Participant P9 argues that the reasons business executives are not involved is

because they don't know what the benefits of DW/BI actually are. Participant P3 then proposed that DW/BI education for executives will be good to resolve this problem.

Question 6

T: Who do you think ought to have decision powers to make a change to improve DW/BI modules?

P9: The CIO, executives, and senior managers of big corporates because they are the major key beneficiaries and they know what the expectations are. Nobody else know no the expectations except business.

(Follow up question)

T: How are going to achieve that? Do you think currently the same executives are doing that?

P9: No; executives don't have the decision powers because they don't know either what are the benefits of BI. We may have to teach them too. They are the ones who are supposed to be driving BI education. As an add-on we should also look at business intelligence education for executives.

Table 7-27 below presents the ideal stakeholders that ought to be involved in DW/BI education. Data reveal that there is a high need for DW/BI *professionals*; *Universities*, through its faculty members, and *Business* should be involved. These three stakeholders have a total sum of 11 occurrences each. *Students* should be involved through the feedback process.

Table 7-27: Count of codes per document for code family: *Ideal stakeholders involved*

Ideal stakeholders' involved	Participants (P)																Totals
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
BI professionals	0	1	1	1	0	0	1	1	1	0	1	0	1	1	1	1	11
Business	0	1	1	1	0	1	1	0	1	1	1	1	0	1	1	0	11
Universities (academic professionals)	1	1	1	1	0	1	1	1	0	0	1	0	1	0	1	1	11
Past students	0	0	0	0	0	1	0	1	0	0	1	0	1	0	0	0	4
Totals	1	3	3	3	0	3	3	3	2	1	4	1	3	2	3	2	37

7.5.1.10.1 Conclusions drawn from involved stakeholders

Currently, universities are the most involved stakeholders in the DW/BI education, with less involvement of industry and other stakeholders. More stakeholders ought to be involved, in order to grow the DW/BI education space. DW/BI professionals, business, and universities hold an equal responsibility of getting involved and growing the DW/BI education space. The two modes, *Current* and *Ideal*, pair the ninth CSH boundary question, i.e. professionals (source of knowledge). In CSH the three stakeholders (professionals) are referred to as the sources of knowledge underpinning the required change.

BI professionals, Business, and Universities, are together the actors (A) component of CATWOE. These stakeholders (actors) should perform the DW/BI education transformation (T). They are required to drive DW/BI education change using the guidelines presented in Chapter 5, Section 5.6.

7.5.1.11 Theme 11: Expertise involved

In the previous section stakeholders that must be involved in the development of DW/BI education were identified. The aim of Question 10 is to identify kinds of expertise that should be involved in DW/BI education transformation. Codes are created for various stakeholders identified.

Table 7-28 and Table 7-29 below present codes for the *Current and Ideal Expertise involved* families respectively. The most involved expertise is currently *Academic expertise*, with the highest total occurrences of 8. *DW/BI consultants* and *Research expertise* are the second most stated involved expertise, with total occurrence of 2 each. As an example, the following code excerpt shows that participant P2 says that the *Academic expertise* are involved, and also to some extent the *DW/BI consultants'* expertise. Participant P6 says that it is a challenge to get subject matter expertise in technologies.

Question 10

T: What kinds of expertise do think is currently consulted or involved in the development of DW/BI module?

P2: Currently I feel that academic expertise and a bit of consultants are involved.

P6: The academic expertise are involved but not on the technical and BI technology expertise. We struggle to get subject matter experts on those technologies too. Real life BI experts are needed.

Table 7-28: Count of codes per document for code family: *Current expertise involved*

Current expertise	Participants (P)																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Totals
Academic expertise	0	1	1	0	0	1	1	1	0	1	1	1	0	0	0	0	8
BI consultants expertise	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2
Research expertise	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
Not sure	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
Industry expertise	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Previous students expertise	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
Totals	1	2	1	1	0	1	2	1	0	1	1	1	1	1	0	0	14

Table 7-29 below presents the *Ideal Expertise* to be involved in DW/BI education. *Technical expertise*, with the highest total of 8 occurrences, should be the most involved expertise. *Academic expertise* is the second most stated involved expertise, with 7 occurrences. *Business*, *Teaching*, and *Hybrid expertise* (teaching and technical) – in this order – complete the top list of most stated ideal expertise.

Table 7-29: Count of codes per document for code family: *Ideal expertise involved*

Ideal Expertise	Participants (P)																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Totals
Technical expertise	0	0	0	1	0	1	1	1	0	0	1	0	1	1	1	0	8

Academic expertise	0	1	1	1	0	0	1	1	0	0	1	0	1	0	0	0	7
Business	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0	1	4
Teaching expertise	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	3
Hybrid expertise (teaching and technical)	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	2
BI managerial	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Business management	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
Research	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
Totals	0	2	1	4	0	2	4	2	1	1	4	0	2	2	1	1	27

7.5.1.11.1 Conclusions drawn from expertise involved

Currently, academic and technical experts from faculty members and DW/BI experts are involved in DW/BI education development. There is also some research expertise involved. Ideally, there ought to be an involvement of both technical and academic expertise. Business expertise should also be involved. Business expertise is important because DW/BI is firstly about business, and then the technicalities that support it. Teaching expertise is important because it is a means of delivering knowledge to students. The two modes, *Current* and *Ideal*, pair the eighth component (expertise) of CSH's boundary questions.

7.5.1.12 Theme 12: Reflected views

The aim of Question 12 is to establish whose views the DW/BI module should reflect according to participants. Codes are created for each identified view. Three types of stakeholders were identified already; those with decision powers (Section 7.5.1.7), the affected (7.5.1.8), and the involved (7.5.1.10). Table 7-30 and Table 7-31 and below present code families for *Current* and *Ideal reflected views* respectively. The most reflected views are those of the *Universities*, with 13 occurrences. *Industry* and *Students* are second and third, with 6 and 2 occurrences respectively. As an example, the following code excerpt shows the response of participant P1 to Question 14.

Question 14

T: According to you, DW/BI module development or education is currently representing the worldviews or perspectives of which stakeholders?

P1: Our programme or module reflects the views of the academic professionals because we have to maintain academic standards; and also the industry since we do consult few of them that we have the relationship with. Student's views are also reflected to some extent because we ask them for feedback for their reflection of the programme.

Table 7-30: Count of codes per document for code family: *Current reflected views*

Current Reflected views	Participants (P)																Totals
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Universities (performance)	1	1	1	1	0	1	1	1	0	1	1	1	1	0	1	1	13
Industry (to some extent)	1	0	0	0	0	1	0	0	0	0	1	0	1	1	1	0	6
Students	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2
Totals	3	1	1	1	0	2	1	1	0	1	2	1	3	1	2	1	21

Table 7-31 below shows that ideally *Business'* views should be the most reflected, with the highest occurrence of 9. *Industry's (BI experts)* and *Universities'* views are second and third highest most stated, with 8 and 5 occurrences respectively. For example, the following code excerpt shows the response of participant 11 to Question 14.

Question 14

T – According to you, DW/BI module development or education ought to representing the worldviews or perspectives of which stakeholders?

P11: (1) It should be business and industry (how things are done) and DW/BI trends in the industry.

Table 7-31: Count of codes per document for code family: *Ideal reflected views*

Ideal Reflected views	Participants (P)																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Totals
Business	0	0	1	1	0	0	1	1	1	1	1	1	0	0	0	1	9
Industry	1	1	1	1	0	1	0	0	0	0	1	0	1	1	0	0	8
Universities	1	0	0	0	0	0	1	0	0	1	0	0	1	0	0	1	5
Students	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	3
Totals	3	1	2	2	0	1	2	1	1	3	2	1	3	1	0	2	25

7.5.1.12.1 Conclusions drawn from reflected views

Universities' views dominate current DW/BI education. Industry's views are not much reflected. There is some indication of students' views that are reflected. It should be noted that business' views are missing in the current DW/BI education reflected views. It was highlighted before that business should define business requirements for DW/BI education, because they own the real data. Ideally business' views should be the most reflected because they contribute with industry (technical) expertise. Universities' views should be reflected because academic standards have to be upheld and maintained. It is good practice to always get feedback from former (past) students to see continuous improvement in DW/BI education. The two modes, *Current* and *Ideal*, pair the twelfth component (worldview) of CSH's boundary questions.

The guidelines created in Chapter 5, Section 5.6 were based on DW/BI literature and universities' input. Literature covers both academic and online technical articles as well as journals of DW/BI professionals. Universities' views were represented through the studying of DW/BI modules' curriculums. The guidelines are enhanced by the views of DW/BI professionals and former (past) students identified in this chapter.

7.5.1.13 Theme 13: Success measures

The purpose of Question 5 is to establish the measure of improvement for a DW/BI module. Codes are created for answers that indicated different measures identified

by the participants. Table 7-32 and Table 7-33 below present code families for *Current* and *Ideal DW/BI education success measures* respectively. Table 7-32 presents a list of codes for the *Current DW/BI education success measures*. *Practical projects* are the most stated success measures, with a total occurrence of 5. *Student pass*, *Kimball model knowledge*, *Technical abilities*, and *The impact that students have in the industry* complete the list of the most stated success measures. As an example, the following code excerpt shows the response of participant P1 to Question 5.

Question 5

T: Who or what do think currently are the measures of success or improvement for DW/BI module?

P1: when students pass and when participate and deliver practical projects. Students should be able to build a complete BI project.

Table 7-32: Count of codes per document for code family: Current success measures

Current DW/BI education success measures	Participants (P)																Totals
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Practical projects	1	0	0	0	0	0	1	1	0	0	0	1	0	1	0	0	5
Student pass	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	3
Kimball model knowledge	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	2
Technical abilities	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	2
The impact students have in the industry	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	2
Complete development life cycle knowledge	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
Knowledge and implementation of DW/BI tools	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
None	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Practical projects linked to industry	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Universities do what they think is right	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
What qualifies you to be a DW/BI developer	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Totals	2	3	1	1	0	2	1	1	1	1	1	1	1	2	1	1	20

Table 7-33 below presents a list of codes for Ideal DW/BI education success measures code family. *The impact students have in the industry* (7 occurrences) and *Students should be able to build a complete DW/BI system* (5 occurrences) emerge as the top two stated success measures. As an example, the following code excerpt shows a response of participant P8 to Question 5.

Question 5

T: What do think ought to be the measures of success (or improvement) for DW/BI module

P8: Students should be able to convert any business scenario into a BI solution. Students should be able to build a complete BI system.

Table 7-33: Count of codes per document for code family: *Ideal* success measures

Ideal DW/BI education success measures	Participants (P)																Totals
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Impact students have in the industry	0	1	0	1	0	0	1	0	0	0	0	1	1	1	1	0	7
Students should be able to build a complete DW/BI system	0	0	1	0	0	1	0	1	1	0	0	0	0	0	0	1	5
Kimball model knowledge	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
Student pass	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
BI body	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Totals	1	2	1	1	0	2	1	1	1	0	0	1	1	1	1	1	15

7.5.1.13.1 Conclusions drawn from DW/BI education success measures

The current success measures are the ability of students to do practical projects and whether they pass the exams or not. The *Ideal success measure* is the impact students have in the market. Students are prepared for industry and therefore only industry can say if they are adding value or not. The two modes, *Current* and *Ideal*, pair the third component (measure of improvement) of CSH's boundary questions.

7.5.1.14 Theme 14: Success assurer (Guarantor)

The purpose of Question 11 is to identify the DW/BI education assurer (guarantor). Codes are created for various assurers identified. Table 7-34 and Table 7-35 below present the *Current Success assurer* and *Ideal Success assurer* code families respectively. Table 7-34 presents a list of codes for the current success measures. *Tests and student practical projects* and *When universities meet industry requirements* are the current most reported guarantors, with 4 and 3 occurrences respectively. Two participants argue that there are not any success assurers. The following code except shows the response of participant P1 to Question 11, as an example of the *Tests and student practical projects* code.

Question 11

T: Who or what do you think is currently assuring successful outcome or conditions of DW/BI module?

P1: The student evaluation process such as tests and practical projects they build.

Table 7-34: Count of codes per document for code family: *Current success assurer* (guarantor)

Current Success assurer	Participants (P)																Totals
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Test and practical projects	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	4
When universities meet industry requirements	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	3

None	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2
Academic professionals	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
High academic standard and good students	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
Impact students have in the industry	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
Number of students in take into the DW/BI programme	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
Student employability (impact students have in the industry)	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
Students	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
Knowledge and implementation of DW/BI tools	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Totals	1	1	1	1	0	1	2	1	1	1	0	1	2	1	1	0	15

Table 7-35 presents the code list for the *Ideal success assurer* (guarantor). The most identified ideal guarantor, with 4 occurrences, is that students need to do *Practical projects*. For example, the following code excerpts show the responses of participants P8 and P11 to Question 11. Participant P8 identifies *practical projects* as the *Ideal success assurer*; participant P11 thinks that a *professional DW/BI body* is needed to be the assurer.

Question 11

T - Who or what do think ought to be assuring successful outcome or conditions of DW/BI module?

P8: Setting up a very large practical exam where individual students get to develop a BI solution

P11: In that case I think you have to look at professional bodies. If you look at South African institute of chartered accounts (SAICA) and that is the assurance. If somebody has those titles it means we trust this guy. BI needs such a body too. If you look at IT the exams that professionals write are not from professional bodies instead they from company specifics or vendors. We cannot use those as a body. Such exams are testing the knowledge of specific vendor products. Therefore you need an in depended body that does not look at tools.

Table 7-35: Count of codes per document for code family: *Ideal success assurer (guarantor)*

Ideal Success assurer	Participants (P)																Totals
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Practical projects	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	4
BI body	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	2
When students are employable	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	2
Academic professionals	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
Collaboration between all stakeholders involved	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
Industry leaders	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
Number of students intake into the DW/BI programme	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
Quality of DW/BI education	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
Students that can deliver DW/BI solutions	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
Tests	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
tie the project to the companies	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
To answer relevant business questions	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
When universities meet industry requirements	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
Totals	2	1	2	0	0	1	2	1	1	1	1	1	2	1	1	1	18

7.5.1.14.1 Conclusions drawn from Ideal success assurer (guarantor)

Tests and practical projects are the main success assurers reported by participants. These practical projects have to meet industry requirements. Ideally *Practical projects*, *the creation and involvement of an industry DW/BI body*, and *Student's employability* should be success assurers. The two modes, *Current* and *Ideal*, pair the ninth component (guarantor) of CSH's boundary questions.

The three SSM evaluation criteria elements, i.e. *efficacy*, *efficiency* and *effectiveness*, are used with respect to the problem situation: whether students can successfully complete practical projects (*efficacy*); and whether DW/BI education can produce DW/BI students that can complete practical projects, given their constraints (*efficiency*). The first set of constraints is based on Theme 6 (Section 7.5.1.6); it includes that *universities cannot control technology, industry input, and students employability*. The second set of constraints is based on Theme 7 (Section 7.5.1.7), i.e. industry should determine the DW/BI standards and business must control the requirements going into the DW/BI modules, and academic processes should be preserved. The third set of constraints are based on Theme 9 (Section 7.5.1.9); this includes that business and industry own the data the requirements, universities control and deliver the teaching aspects of DW/BI education, universities are identified as having the powers to control the academic, and students are the receivers of DW/BI education. The employability of students should be a long term goal of DW/BI education. This aspect supports the quality of DW/BI solutions that students develop in order to become successful DW/BI professionals; this forms part of effectiveness (*effectiveness*).

7.5.1.15 Theme 15: Prerequisites

The purpose of Question 19 is to establish what DW/BI modules' prerequisites should be. Codes are created for various *Prerequisites* identified. Table 7-36 below presents codes for the *Prerequisites* family. The sum of the occurrences of each of these codes is presented in the 'Totals' column in table. *Basic database skills* is the most stated, with 10 occurrences. *Programming, Mathematics, Understanding of computers and performance, Statistics, Problem solving, and Business knowledge* complete the top list of the most stated prerequisites. The following code excerpt is the response of participant P1 to Question 19 about prerequisites.

Question 19**T: What should be viewed as prerequisite knowledge/skills to enrolling for DW/BI module?**

P1: Basic database knowledge is key and also analytics, basic statistics and mathematics for logical thinking.

Table 7-36: Count of codes per document for code family: *Prerequisites*

Prerequisites	Prerequisites (P)																Totals
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Basic database	1	1	0	0	1	1	1	1	0	1	1	0	1	0	1	0	10
Programming	0	0	0	0	0	0	1	0	0	1	1	0	2	0	1	0	6
Mathematics	0	1	1	1	0	1	0	0	1	0	0	0	0	0	0	0	5
Understanding of computers and performance	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	1	4
Statistics	2	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	4
Problem solving	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3
Business knowledge	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	2
Should be mind-set not necessarily skills	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
Information systems understanding	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
Totals	5	2	1	1	1	3	2	1	1	3	5	0	4	1	3	3	36

7.5.1.15.1 Conclusions drawn from prerequisites

Basic database knowledge is an important prerequisite; students must have it to enrol for a DW/BI module. The skills identified here can be mapped to the table of guidelines in Chapter 5, Section 5.6, as indicated in brackets: database knowledge (Lit_5_PreGen_DBProg from Table 5-6), programming (Lit_14_PreGen_Prog), understanding of computers and performance (Lit _32_PeGen_HardwareKnow), problem solving (Lit_9_PreGen_ Analytical), and business knowledge (Lit_17_PreGen_Business) are identified prerequisites – refer to Chapter 5, Section

5.6. Mathematics and statistics are two enhancements to the prerequisites' list compiled in Chapter 5.

7.5.1.16 Theme 16: Elective subjects

The purpose of Question 20 is to identify elective subjects that can be taught in parallel with the DW/BI module. Codes are created for various *Elective subjects*; the sum of the occurrences of each of these codes is presented in the 'Totals' column in the table. Table 7-37 below presents the codes for the family *Elective subject*. *Advance database* has the highest total occurrence of 7. As an example, the following code excerpt shows participant P6's response to Question 20 about elective subjects.

Question 20

T: What general IT knowledge/skills should be taught in parallel with the DW/BI module?

P6 : DW/BI is all about "I have a problem how I solve it with the available data?", so you need:

(1) Problem solving skills (Decision support systems) (2) Advance databases (3) Requirements gathering (4) Mathematics (5) Statistics

Table 7-37: Count of codes per document for code family: *Elective subject*

Elective subjects	Participants (P)																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Totals
Advance database (SQL)	0	0	1	0	1	1	0	1	1	1	1	0	0	0	0	0	7
Business analysis/ requirements gathering	0	1	0	0	1	1	0	0	0	0	0	1	0	0	0	0	4
Decision support subjects (good reasoning skills)	0	0	0	1	0	1	0	0	1	1	0	0	0	0	0	0	4
Social skills (e.g. communication and presentation)	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	1	4
Custom development (programming)	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	3

Mathematics	1	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	3
Computer architecture and design	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	2
Statistics	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2
Advance business knowledge	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
IT and data security	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
Network infrastructure	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
Project management	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
Systems analysis	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
Understanding of computers	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
Total	2	1	1	3	2	5	1	1	3	2	5	1	0	1	4	3	35

7.5.1.16.1 Conclusions drawn from elective subjects

The most important elective subject is *Advanced database* skills. DW/BI can involve complex problems that may require students to have advance database skills. Also, complex projects require good communication, as well as requirements gathering and analysis; thus, *Business analysis* is also an important elective subject. Subjects such as *Decision support systems* for analytical thinking can help students to do complex DW/BI projects. *Custom development*, *Statistics*, and *Mathematics* are important for logical thinking of students. Other advance subjects such as *Computer architecture and design* can help with the overall picture of DW/BI solutions.

Advance database (Lit_8_AdvGen_DBA), *Business analysis/requirements gathering* (Lit_2_AdvGen_Requirements), *Decision support subjects* (Lit_6_AdvGen_Analytical) and *Computer architecture and design* (Lit_3_AdvGen_TechArchData) are already identified as elective or advance subjects Chapter 5, Section 5.6. The code in brackets indicates the associated skill in Chapter 5, Section 5.6.

Even though *Custom development (programming)*, *Mathematics*, and *Statistics* are identified as a prerequisites in the previous theme, these three were not identified as elective subjects in Chapter 5. These three subjects, including social skills (e.g. communication and presentation), are new enhancements to the elective skills.

7.5.1.17 Theme 17: Assessment component

Question 23 is aimed at establishing the overall view of the assessment component for student knowledge within the DW/BI module. Codes are created for various assessment components that participants identified. Table 7-38 below presents the codes for the *Assessment component* code family. *Practical projects* and *Tests (theory)* are the most reported assessment methods, with 11 occurrences each. The following code excerpt is the response of participant P4 to Question 23 about the DW/BI module assessment component.

Question 23

What is your overall view of the assessment component for student knowledge within the DW/BI module?

P4: I think students should get theory tests on DW principles and give students a dataset to build a solution and do an analysis on that.

Table 7-38: Count of codes per document for code family: *Assessment component*

Assessment component	Participants (P)																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Totals
Practical projects	1	0	0	1	1	1	1	0	1	1	1	0	1	0	1	1	11
Tests	1	0	0	1	1	1	1	0	1	1	1	0	1	0	1	1	11
Practical should be aligned to industry	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2
A continuous assessment method	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
No experience of it	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
Totals	2	1	1	2	2	2	2	1	2	2	2	1	2	0	2	2	26

7.5.1.17.1 Conclusions drawn from assessment components

It is imperative that students are tested for their theoretical knowledge and understanding of DW/BI concepts. The combination of theory and practice is important to show overall knowledge and understanding. The guidelines presented in Chapter 5, Section 5.6, support that students should be assessed on both theoretical and practical components.

7.5.1.18 Theme 18: Student preparation

The motivation for Question 16 of the interview is to establish what participants think of the current DW/BI curriculum, i.e. whether it is preparing students to meet industry skills requirements. Codes are created on different views given. These codes are grouped as the *Student preparation* code family listed in Table 7-39. The sum of the occurrences of each of these codes is presented in the 'Totals' column. *Students are to some extent prepared* is the view with the highest frequency (of 4). As an example, the code excerpts below show responses of participants P4 and P9 to Question 16.

Question 16

T: What is your overall perspective of what universities teaches in DW/BI modules? Do you think that the current DW/BI curriculum is better preparing students to meet industry skills requirements?

P4: They learn a tool and then they build a data warehouse. They're more focused on data analysis but no element of business insight into it. No real life business component into it. And yes universities are to some extent preparing students for industry.

P10: I have mixed answers on this one. My own experience is that I learn a lot of knowledge that is useful at university and that knowledge helped me to pick up things quickly in new technology and ways of doing things as a progress in my career. But other people that I talk to find it contrary. For example some people feel they do not see the need why they had to do advance databases at university. I think we are sitting with a problem of saying students don't know why they need to know some subject matters. How does that feature in the world of BI? A typical example is that BI developers see no need for custom development.

The second view is that *Students are better prepared by universities*, with 3 occurrences. The following code excerpts show the responses of participants P1 and P3 for Question 16, in support of this finding.

Question 16

T: What is your overall perspective of what universities teaches in DW/BI modules? Do you think that the current DW/BI curriculum is better preparing students to meet industry skills requirements?

P1: Yes. With my experience of industry and the input of few professionals we do prepare students for industry.

P13: According to feedback from different contacts in industry the DW/BI module at our campus is preparing students to meet industry skills requirements. Some of the representatives attended final presentations and were very impressed with the students and said that most of the students could be assigned to clients right away. From the feedback this is, however, not the case at most SA universities.

Table 7-39: Count of codes per document for code family: *Student preparation*

Student preparation	Participants (P)																Totals
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Students are to some extent prepared	0	0	0	1	0	1	0	0	0	1	0	0	0	0	1	0	4
Students are well prepared by universities	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	3
Students are not well prepared	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	2
Students still need to be trained after university	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	2
Students are not taught how to connect theory with technology	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Students have more theory but less practical knowledge	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
We don't have enough DW/BI professionals in the industry that are coming from universities	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Totals	1	1	2	1	0	1	1	1	1	1	1	0	1	0	1	1	14

7.5.1.18.1 Conclusions drawn from student preparation

Question 16 aims to better understand the perspectives of the participants. The responses of the participants indicate various levels of students' preparation. Four participants argue that students are to some extent prepared by universities; whilst three participants argue that students are well prepared by universities. Some participants argue that students are not well prepared by universities, and two participants argue that students still need to be trained after university.

These answers may be linked to the SSM evaluation criteria element of *efficacy*, in terms of assessing whether current DW/BI education serves its intended purpose.

7.5.1.19 Theme 19: A general perception of what universities are teaching

Another motivation for Question 16 of the interview is to establish participants' general perception of what universities are teaching. Codes are created for different perceptions given. Table 7-40 below presents the codes for the *A general perception of what universities are teaching* code family. There are three repeated perceptions, with total occurrences of 3, 3, and 2 respectively. The first perception is that *Universities spend more time on the design and development of a DW*. The second perception is that *Students do not have enough theoretical knowledge*. The third perception is that *Technology is thought to be DW/BI but it is not; it's just a tool*. For example, the following code excerpt shows the response of participant P7 to Question 4.

Question 4

T: According to you, what is currently the purpose of DW/BI module such that it is serving the interests of those who benefit from it?

P7: Currently, the purpose is to expose students to the technologies that are available in the market. Not much is spent on the benefits/soft issues. More time is spent on the design and development of a DW. This is probably the reason why people have not been exposed to such an area. They are not technologically adequate and/or have very little development/programming skills

Table 7-40: Count of codes per document for code family: *Perception of what universities are teaching*

Perception of what universities are teaching	Participants (P)																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Totals
More time is spent on the design and development of a DW	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1	3
Students do not have enough theoretical knowledge	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	3
Technology is thought to be DW/BI but it is not; it's just a tool	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	2
Some students are well prepared some not depending on university they come from	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
Students are exposed to good technology	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
Students build DW/BI solution without data analysis	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
Students don't know why they do DW/BI	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
They are taught to be technical experts without business value	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
Universities just introduces a concept (BI/DW) to them	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
Universities teach transition from OLTP systems to DW/BI, not	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1

DW/BI																	
Business element is missing	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
No difference between those with DW/BI education and those who don't	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Professionals learn DW/BI on the job	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Teach theory	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
They are using case studies not real situations	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
We don't have enough DW/BI professionals in the industry that are coming from universities or that have formal DW/BI education.	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Totals	0	4	1	1	2	2	1	0	2	2	2	0	0	1	1	2	21

7.5.1.19.1 Conclusions drawn from perception of what universities are teaching

Question 16 aims to provide a better understanding of the perspectives of the participants on universities' DW/BI education. According to participants P7, P14, and P16, universities currently spend time on the design and development of a DW. Current focus is on equipping students with the ability to develop the DW solution, outside the focus of business. A business perspective must be added to the guidelines.

7.5.1.20 Theme 20: DW concepts experience

Question 17 is directed towards an understanding of what the participants' own experiences of DW concepts are. Codes are created to represent the various

experiences provided. The sum of the occurrences of each of these codes is presented in the 'Totals' column. Table 7-41 below presents the codes for the *DW concepts experience* code family. Most participants argue that DW concepts define basic standards and DW/BI language, with a total sum of 5 occurrences. *DW concepts need to be better explained with practical examples, They are immensely important for students, and They are not the law but a guide* complete the list of the most stated DW experiences, with a total sum of 2 occurrence each. As an example, the following code excerpt shows the response of participant P8 to Question 17 about DW experience.

Question 17

T: What is your own experience of DW concepts?

P8: DW concepts are extremely important because that the language of BI. Anyone who does not know BI concepts cannot have a decent BI conversation. Students must have a good solid understanding of DW concepts because we use them eve in the industry.

Table 7-41: Count of codes per document for code family: *DW concepts experience*

DW concepts experience	Participants (P)																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Totals
They define basic standards and DW/BI language	1	0	0	0	0	1	0	1	0	1	0	1	0	0	0	0	5
DW concepts need to be better explained with practical examples.	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	2
They are immensely important for students	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	2
They are not the law but a guide	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	2
They have been around for some time and they are matured and they can be taught well.	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1

There is more to what universities are teaching	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
They are based on leading minds and good for students	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
Totals	1	0	1	1	1	1	1	1	1	1	1	1	0	0	2	1	14

7.5.1.20.1 Conclusions drawn from DW concepts experience

Question 17 is also aimed at better understanding of the perspectives of the participants on DW concepts. A decent DW/BI conversation is based on good understanding of DW/BI concepts. A good knowledge and understanding of DW concepts is a good point of departure for students. DW concepts must be taught and better explained to students; however, they are not the law but rather a helpful guide.

7.5.1.21 Theme 21: What to teach DW/BI students

Universities need to be guided in terms of what to teach DW/BI students. Codes are created on views that define *What to teach DW/BI students*. Table 7-42 below presents the codes for the *What to teach DW/BI students* family. The sum of the occurrences of each of the codes is presented in the 'Totals' column in the table. The highest code sum is *DW/BI principles (primary)*. Second and third most stated codes are *Practical work* and *Tools or technology (secondary)*. For example, the following code excerpt shows the response of participant P2 to Question 17 about what to teach DW/BI students.

Question 17

T: What is your own experience of DW concepts?

P2: DW concepts define the base of BI, it is important that all BI professionals should know them. And this should start at universities.

Table 7-42: Count of codes per document for code family: *What to teach DW/BI students*

What to teach DW/BI students	Participants (P)																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Totals
DW/BI principles (primary)	1	1	1	1	1	0	2	1	1	1	1	0	1	1	1	1	15
Practical work	1	0	0	1	1	0	1	1	1	1	1	1	1	1	1	2	14
Tools or technology (secondary)	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	0	13
Business models	0	0	1	0	0	0	1	1	1	0	0	0	0	0	0	0	4
Front end of reporting	0	0	0	0	0	0	0	0	1	1	1	0	0	0	1	0	4
BI trends	1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	3
Business analysis and requirements gathering	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	3
Draw a DW/BI story that explain why do DW/BI	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	3
Analysis and trend analysis	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	2
Change group projects to individual projects	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	2
Difference between OLTP and DW/BI	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	2
Documentation	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	2
How to build ETL process	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	2
Optimisation and improvements	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	2
Push internship programmes knowledge	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	2

back into universities																	
BI architecture	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
BI infrastructure	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
BI project management	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
BI solution maintenance	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
Business engagement	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
Cross subjects area analysis	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
Data quality management	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
Data sources	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
Databases	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
Decision making	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
Delivery mechanisms	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
DW methodologies	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
How to train users to use DW/BI solution they developed	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
Look at various DW/BI definitions	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
Purpose of DW/BI	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
Scheduling jobs	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
Testing	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
Totals	4	2	4	4	4	5	5	5	14	9	11	2	2	4	12	3	90

7.5.1.21.1 Conclusions drawn from what to teach DW/BI students

The focus of DW/BI teaching should be DW/BI principles as well as practical projects and tools. Universities should invest more time teaching DW/BI principles. These principles should also be demonstrated with practical projects through DW/BI technology tools. DW/BI *practical work* should be one of the focus areas that universities must teach.

The codes developed in this theme (Table 7-42) are used again in Section 7.6 to link the content of this chapter to the content of Chapter 5.

7.5.1.22 Theme 22: Other concerns

Other concerns that may be relevant to study are identified and coded. A code family *Other concerns* is created. The sum of the occurrences of each of these codes is presented in the 'Totals' column in the table. Table 7-43 below presents codes for *Other concerns*. As can be seen, the data reveal that DW/BI *managers need DW/BI training to know its value*, with 4 occurrences. The other two main concerns are that DW/BI *should not be for post graduate only* and DW/BI *should not be for post graduate only*. As an example, the following code excerpt shows the response of participant P9 to Question 6 about decision makers. The participant also highlights a concern about DW/BI managers' lack of knowledge.

Question 6

T: How are going to achieve that? Do you think currently the same executives are doing that?

P9: No; executives don't have the decision powers because they don't know either what are the benefits of BI. We may have to teach them too. They are the ones who are supposed to be driving BI education. As an add-on we should also look at business intelligence education for executives.

Table 7-43: Count of codes per document for code family: *Other concerns*

Other concerns	Participants (P)																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Totals
BI managers	0	0	0	0	0	0	1	0	1	0	1	0	0	0	1	0	4

need DW/BI training to know its value of DW/BI																	
BI should not be for post graduate only	0	0	1	0	0	1	0	0	1	0	0	0	0	0	1	0	4
A module is not enough for a DW/BI education	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	3
BI consulting companies get involved only to recruit instead of building this educational space	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	2
At honours DW/BI students should benefit not starting to learn about it.	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
At honours it should be about specialised areas and taking DW/BI forward	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
At honours it should not be about models, but how to deliver DW/BI	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Developers are concerned about technology than DW/BI principles	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
Students and professionals give up early	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1

DW/BI career because they don't know its value																	
Totals	0	0	6	0	0	1	2	0	2	1	3	0	0	0	2	1	18

7.5.1.22.1 Conclusions drawn from other concerns

The value of DW/BI can only be fully realised if DW/BI managers know its value. There is a need for DW/BI managers to be educated on the value it brings, so that they can take full advantage of it. The value of DW/BI education is realised by those students that are willing to progress beyond undergraduate and into honours degrees. Most students that leave universities with undergraduate degrees and find themselves later in the DW/BI field learn DW/BI on the job. For this reason, there is need for DW/BI education to be introduced earlier than the 4th year. Former (past) students feel that one module is not enough to cover the breadth of DW/BI.

7.5.1.23 Theme 23: DW/BI processes

The purpose of Question 15 is to establish steps that can be taken to improve the DW/BI module. Codes are created for steps or that will lead to an improved DW/BI module. The purpose is to consolidate steps or processes that can be used to improve DW/BI education. Table 7-44 below lists these steps, but the order of execution is not defined. The following code excerpt shows the response of participant P2 to Question 15 about the DW/BI process.

Question 15

T: What do you think should be the process to develop an improved DW module? Please give detailed steps - I would appreciate if you draw some kind of activity diagram that shows dependencies.

P2: Start with the employers and those are businesses people who will give and direct their requirements. Second: get DW/BI professionals to give us guideline as to where DW/BI is going. Third: involve also the academics to determine what level of knowledge to teach and breaking it down in years or modules. These are the most important steps. I think this will be the framework.

Table 7-44: Count of codes per document for code family: *DW/BI process*

BI process	Participants (P)																Totals
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Identify and involve relevant stakeholders	1	1	0	0	0	1	1	1	0	0	1	1	1	1	0	1	10
Check DW/BI literature and trends (other independent)	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	1	4
Define industry requirements	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	3
Structure and develop the programme	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	3
Define success measures	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	2
Choose technology to teach	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	2
Start with what we have	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	2
Conduct course comparison analysis	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
Define the module prerequisites and get better students	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Draw insight from independent solutions in industry	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
Identify core skills needed	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
Incorporate big business models	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
Past students to review	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
Review periodically	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
Totals	5	2	0	4	0	3	3	2	0	1	4	1	3	4	0	2	34

7.5.1.24 Conclusions drawn from DW/BI process

The most important step in the process of improving DW/BI modules is identification and involvement of relevant stakeholders. Data reveal that *Identify and involve relevant stakeholders* is the most stated, with occurrence of 10. *Check DW/BI literature and trends (other independent)*, *Define industry requirements*, *Structure and develop the programme*, *Define success measures*, *Choose technology to teach*, and *Start with what we have* complete the list of the most stated steps to improve DW/BI modules. The next section presents guidelines from interviews.

7.6 Conclusions from the data

The researcher conducted interviews to better understand DW/BI from the perspective of former (past) students, faculty members, and industry role players. The guidelines developed in Chapter 5, as well as aspects of SSM and CSH, were used to develop interview questions. Data analysis resulted in findings on the process to be followed to design and implement a DW/BI module, as well as technical content of such a module. The aspects of CSH and SSM, identified in the analysis of the interview data, are used in the next chapter to guide the process of redesigning the module. Table 7-45 below serves as a summary of the module content reported as by the participants (Table 7-42 in Theme 21).

The content of Table 7-45 is cross referenced to the content of Chapter 5, Section 5.6. The first column on the left shows the skills identified in Table 7-42 (Theme 21). The second column shows the number that references the skill in Chapter 5. The third column gives a short description why the skill is supported or not.

Table 7-45: A list of skills that support the guidelines of Chapter 5.

Skill from answers of participants	Chapter 5 reference no	Description
DW/BI principles	Lit_1_BI_Components Lit_3_BI_DWlifecyle Lit_7_BI_DimModeling Lit_4_BI_Methodologies	DW/BI principles are covered in DW life cycle. The DW life cycle touches on DW/BI components knowledge such as Data, technology, processes, and people are covered.
Practical work	Lit_6_BI_ETL_Tools	BI practical's are covered in the ETL,

	Lit_7_BI_DimModeling Lit_8_BI_MDX Lit_9_BI_OLAP Lit_10_BI_MobileBI	Dimensional, OLAP (CUBE - MDX), and mobile DW/BI development.
Tools or technology	Lit_7_AdvGen_ETL_Tools Lit_6_BI_ETL_Tools	Tools and technology are covered and taught when students do development such as ETL solutions.
Business models	NEW GUIDELINE	Business models skills or knowledge are not identified nor refuted in Chapter 5.
Front end of reporting	Lit_16_AdvGen_DBReporting Lit_9_AdvGen_Interface Lit_4_AdvGen_DBProg	Advance reporting skills are required to enable the development of flexible reports required by DW/BI applications.
BI trends	Lit_10_AdvGen_BigData	BI trends such as DW/BIG data tools, methods and application are covered. No explicit guideline exists on new developments in the field.
Business analysis and requirements gathering	Lit_2_AdvGen_Requirements	According to Kimball <i>et al.</i> (2008d:64) business analysis skills are key in requirements collection process.
Draw a DW/BI story that explain why do DW/BI	This is an instructional preference.	The introduction to the module provided by lecturers does contain this but it is not presented as a narrative. <u>Develop a visual narrative that explains why students should do and gain DW/BI knowledge</u> is not identified nor refuted in Chapter 5.
Analysis and trend analysis	Lit_6_AdvGen_Analytical	According to Kimball <i>et al.</i> (2008d: analytical skills are essential for requirements collection process.
Change group projects to individual projects	This is an instructional preference. Lit_5_AdvGen_OLTP	This is an instructional preference. Due to the limited time available in terms of credit hours for the module, students are typically grouped. Groups are also used to simulate the industry environment of project teams. <u>This aspect requires more reflection.</u>
Difference between OLTP and DW/BI	Lit_17_OLTP_BI	Advance understanding of differences between OLTP and DW/BI systems.
Documentation	Lit_3_BI_DWlifecycle	Documentation is discusses in detail by Kimball in all the phases of the <u>DW lifecycle but more explicit emphasis must be given on the role thereof.</u>
How to build ETL process	Lit_3_BI_DWlifecycle Lit_7_AdvGen_ETL_Tools	ETL development process is part of DW life cycle.
Optimization and	Lit_8_AdvGen_DBA	Optimization can be in terms of database optimization and user

improvements		interface optimization. Database optimization is covered in Chapter 5 and interface optimization is typically covered in prerequisites modules in Data structures and Algorithms <u>These aspects must be emphasised more explicitly.</u>
Push internship programme's knowledge back into universities	This is scope extension to this study.	This is scope extension to this study. <u>This aspect requires more reflection.</u>
BI architecture	Lit_3_AdvGen_TechArchData	According to Kimball <i>et al.</i> (2008d:110) technical architecture skills is key for DW/BI technical architecture.
BI infrastructure	Lit_5_BI_TechAchitect	According to Kimball <i>et al.</i> (2008d: project management skills is essential skill to the entire DW life cycle.
BI project management	Lit_1_AdvGen_ProjectM	According to Kimball <i>et al.</i> (2008d: project management skills is essential skill to the entire DW life cycle.
BI solution maintenance	Lit_3_BI_DWlifecycle	Maintenance is discusses in detail by Kimball in all the phases of the DW lifecycle.
Business engagement	Lit_2_AdvGen_Requirements Lit_3_BI_DWlifecycle	Kimball advocates the role of a "business sponsor" throughout the lifecycle but <u>more explicit emphasis should be given to this.</u>
Cross subjects area analysis	NEW GUIDELINE	Cross subjects area analysis is not identified nor refuted in Chapter 5.
Data quality management	NEW GUIDELINE	Data quality management is part of advance database management, but this aspect requires more explicit reference in the guideline.
Data sources	Lit_3_AdvGen_TechArchData	Data sources are covered by Kimball in DW technical architecture.
Databases	Lit_4_AdvGen_DBProg Lit_12_AdvGen_Metadata Lit_15_AdvGen_DataSecurity	According to Kimball <i>et al.</i> (2008d:327) DB programming skills are key to designing the physical database and planning for performance.
Decision making	Lit_6_AdvGen_Analytical Lit_11_AdvGen_DataMining	Decision making is an analytical and end-user activity that <u>requires explicit attention.</u>
Delivery mechanisms	Lit_10_BI_MobileBI	Delivery mechanisms knowledge such as mobile is covered in Chapter 5 but this aspect should be covered in "trends" components of the

		module, identified earlier.
DW methodologies	Lit_4_BI_Methodologies	Knowledge of alternative DW Methodologies.
How to train users to use DW/BI solution they developed	Lit_14_AdvGen_UserTraining	According to Kimball <i>et al.</i> (2008d: training skills are essential in the deployment phase of a DW/BI solution.
Look at various DW/BI definitions	Lit_1_BI_Components	DW/BI principles are covered in DW life cycle. The DW life cycle touches on BI components knowledge such as Data, technology, processes, and people are covered.
Purpose of DW/BI	This is an instructional preference as well as module orientation issue.	The introduction to the module provided by lecturers does contain this but the <u>current module focuses more in DW than BI.</u>
Scheduling jobs	Lit_3_BI_DWlifecycle	Scheduling jobs is covered by the implementation phase of DW lifecycle.
Testing	Lit_13_AdvGen_Testing	Testing skills are covered in DW/BI solution testing.

By analysing the items in Table 7-45, items not previously covered in Chapter 5 are identified. These items include matters on content as well as items on the process to follow when designing the DW/BI module. The latter are addressed in the next chapter. Table 7-46 contains items identified by participants to be added to the list of aspects to include in the module, including a reference code to be used in Chapter 9. The reference code has a prefix of “Int_” in order to identify its source as interview data. All the items are identified in Theme_21_Teach.

Table 7-46: Items not previously covered in Chapter 5

Item	Code	Comment
Cross subjects area analysis	Int_1_CrossSubjAna	The module should include an element of creating of teaching students to organise reporting elements by business functions.
Data quality management	Int_2_DataQualMan	Data quality management is part of advance database management, but this aspect requires more explicit reference in the guideline.
Business models	Int_3_BusMod	The module should include an element of teaching students about different business models to

		demonstration that DW/BI can be applied almost any business that values data.
Draw a DW/BI story that explain why do DW/BI	Int_4_DrawBIStory	This is an instructional preference. The introduction to the module provided by lecturers does contain this but it is not presented as a narrative. Develop a visual narrative that explains why students should do and gain DW/BI knowledge.
Push internship programme's knowledge back into universities	Int_5_IntProg	This is scope extension to this study. This aspect requires more reflection.
Purpose of DW/BI	Int_6_BIPurpose	This is an instructional preference as well as module orientation issue.

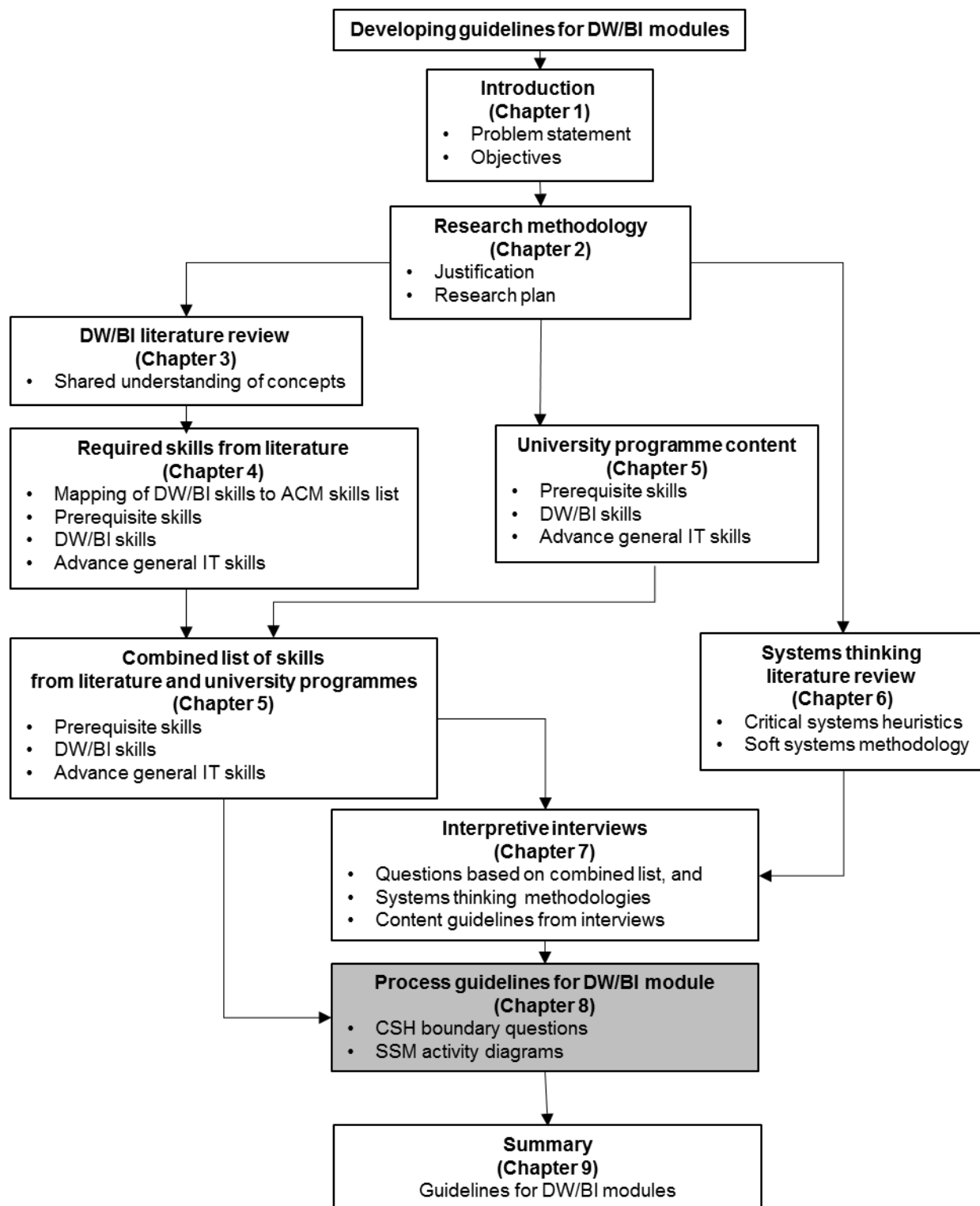
Data analysis shows no guidelines developed in Chapter 5 not supported or refuted by the data from the interviews.

In terms of prerequisites of the DW/BI module (Theme 15), mathematics and statistics are two enhancements to the prerequisites list developed in Chapter 5, Section 5.6.

7.7 Summary

In this chapter open ended interviews are used as a data collection method. The purpose of these interviews were to gain an understanding of stakeholders' perspectives of DW/BI, and the skills students should possess when they enter the industry as graduates (i.e. how to make university students DW/BI ready?). The interview development was guided by questions raised by CSH and SSM, as well as the review of universities' DW/BI programmes and modules— refer to Chapter 5.

Three groups of participants were identified and interviewed. The groups are former students, industry experts, and faculty members. Interviews were recorded and transcribed for analysis. Content analysis was applied to data and conclusions drawn from themes. A summary of findings based on data analysis is presented. Data analysis revealed items not previously covered in Chapter 5. These items include matters on content as well as items on the process to follow when designing the DW/BI module. The process items are discussed further in the next chapter.



Chapter 8 Stakeholders' perspectives on DW/BI education (continuation)

8.1 Introduction

The objective of this study is to develop guidelines for business intelligence/data warehousing modules in information technology programmes at universities; the guidelines are based on an understanding of the needs of various stakeholders. Two methodologies, i.e. critical systems heuristics and soft systems methodology, are used to understand the various stakeholders' perspectives. The findings from the empirical data – presented in Chapter 7 – are used as the foundation for this chapter. This study presents two questions thus far. The first question is: “what to teach in the business intelligence/data warehouse module content?”; and the second question is: “what process should be used to improve business intelligence/data warehouse education?”

Guidelines for business intelligence/data warehouse module content were developed from the literature – refer to Chapter 4. Data analysis supported these guidelines and some additional module content aspects were also identified – refer to Chapter 7. The combined set of guidelines for module content is presented in the summary of the study – refer to Chapter 9.

The researcher also investigates processes to improve the DW/BI module; the interview questions (in Chapter 7) are used for this purpose. Findings relating to the boundary questions of critical systems heuristics, identified through the analysis of the data (in Chapter 7) are presented in this chapter. Critical systems heuristics' boundary findings are presented in Section 8.2 and 8.3.

Soft systems methodology is used in this study to create reflective stakeholder models that are purposeful to the problem situation. The interviews (in Chapter 7) are analysed to identify elements thereof. The findings of Chapter 7 are presented as activity diagrams that are representative of stakeholders' views. The application of

soft systems methodology in this study is presented in Section 8.4.-8.9 for different perspectives.

A chapter summary is presented in Section 8.10.

8.2 Critical systems heuristics (boundary findings)

Critical systems heuristics (CSH) boundary judgments are used to guide the formulation of interview questions in this study – refer to Chapter 7. Boundary judgements are applied to determine business intelligence (BI) situational framing; questions 3 to 12 of the research questionnaire (presented in Section 7.2 of Chapter 7) focus on CSH's boundary questions. Each question should be answered in both the “*Is*” (current) and in the “*Ought to*” (ideal) mode. Table 8-1 below presents the current (“*is*”) and ideal (“*ought to*”) DW/BI education state according to the data presented in Chapter 7 (Themes 3 - 14). Only beneficiaries identified by more than one participant are considered. The twelve CSH questions are categorised into four functional groups, i.e. sources of: motivation, power, knowledge, and legitimation.

Table 8-1: Current (“*is*”) and ideal (“*ought to*”) DW/BI education state

CSH component	Chapter 6 reference	Current state	Ideal state
Source of Motivation			
Beneficiaries	Theme_3_Beneficiaries	• Students • Industry • Universities	• Industry • Students • Business
BI purpose	Theme_4_BIPurpose	• Teach DW/BI technology (tools). • Train and equip students with DW/BI knowledge and skills. • Teach students accurate reporting. • To teach Kimball methodology.	• To teach value of an enveloped solution that pushes business forward. • Make students DW/BI ready for industry.
Success measure	Theme_13_SucessMeasure	• Practical projects • Student pass rate • Kimball model knowledge • Technical abilities • The impact students have	• Impact students have in the industry. • Students should be able to build a complete DW/BI system.

		in the industry.	
Sources of Power			
Decision powers	Theme_5_ DecisionMakers	- Universities (Faculty members) - Industry (to some extent)	- Industry (experts) - University (faculty members) - Business - Students
Resources	Theme_7_ DecisionMakers Powers	- Universities control the academic component (standard and course content requirements). - Universities control the teaching aspect. - Universities have control over textbooks and other materials.	- Industry should determine DW/BI standards. - Business must control data and requirements. - Universities should control course content. - We don't have a body that recognises what a DW/BI professional is.
Environment	Theme_6_ Conditions OutsideControl	- Universities can't control DW/BI technology and its trends (hardware and software licensing costs). - Universities can't control industry input of real life problems. - Industry can't control programme requirements. - Universities can't control student's employability.	- Universities can't control DW/BI technology and its trends (hardware and software licensing costs). - Universities can't control industry input of real life problems. - Industry can't control programme requirements - Universities can't control student's employability.
Sources of Knowledge			
Involved stakeholder	Theme_10_ Involved Stakeholders	- Universities - Industry (to some extent)	- BI professionals - Business - Universities (academic professionals) - Former students
Expertise	Theme_11_ Expertise Involved	- Academic expertise - BI consulting expertise	- Technical expertise - Academic expertise - Business - Teaching expertise - Hybrid expertise (teaching and technical)
Guarantor	Theme_14_ Success Assurer	- Tests and practical projects - When universities meet industry requirements. - None	- Practical projects - BI body - When students are employable.
Sources of legitimation			
Affected	Theme_8_ Affected	Business	- Business - Students

stakeholders	Stakeholders	• Industry • Students • BI professionals	• Industry • BI professionals
Emancipation	Theme_9_ Affected Stakeholders Control	• Universities - control academic aspect. • Business - must have input into course content. • Industry - have no control. • Students - must have input into course content.	• Business must define requirements that go into DW/BI. • Industry should define business requirements. • Universities control teaching and academic aspect. • BI professional control practical projects.
Worldview	Theme_12_ Reflected Views	• Universities • Industry (to some extent) • Students	• Business • Industry • Universities • Students

The different answers to the questions in the “*is*” and the questions in the “*ought to*” modes point to unresolved boundary issues; the next section presents conclusions drawn from the findings.

8.3 Conclusions drawn from boundary findings

This section presents a summary of conclusions drawn from Table 8-1. The following practical guidelines are formulated as a result of the CSH analysis:

Distinction should be made between the roles of business and industry. Business and industry should be involved as beneficiaries of the DW/BI module; the students and universities should also be involved in the development process of the DW/BI module (Beneficiaries).

DW/BI education can be improved by teaching the value of an enveloped solution that pushes business forward and makes students DW/BI-ready for industry. A business component should be incorporated in the teaching of DW/BI technology, principles, and skills. This entails a transition from just teaching DW/BI technology, principles, and skills without the core business value, to teaching DW/BI bundled with business purpose and value (BI purpose).

Development of practical DW/BI solutions should be the measure of success for the DW/BI module. The impact that students will ultimately have in the industry should be included as a measure of success (Measure of success).

The decision powers for a DW/BI module should include business, students, universities, and industry (Decision powers).

The roles of stakeholders should be clearly defined (as discussed in Chapter 6). Industry should determine DW/BI standards; business should control data and requirements; universities should control course content; and a body that recognises what DW/BI professionals do should be formed (Resources - decision maker's powers).

There is an expectation that universities should produce students that are better prepared. However, there are some environmental aspects (that are considered to be outside of the control of DW/BI modules) that should be identified. Universities cannot control DW/BI trends, the development of DW/BI technologies, hardware and software licensing costs, and the input of industry's real life problems (Environment - conditions are not controlled by the decision taker),

Universities, industry, business, and former (past) students are the stakeholders which should provide the necessary information, practical experience, and know-how to achieve objectives (Involved stakeholder).

The expertise that should be involved in DW/BI modules' development include: academic and DW/BI consultants, technical, business, and hybrid (teaching and technical) (Expertise).

Projects involving practical work, as well as employability of students, should be the guarantee that planners (involved stakeholders) aim to achieve. A DW/BI body of knowledge comprising of all relevant stakeholders (the involved and affected) should be formed and included as a guarantor (Guarantor).

The four main stakeholders that are affected are: business, students, industry, and DW/BI professionals (Affected stakeholders).

Specific concerns (how or what will emancipate those that are affected by the DW/BI module) of the affected stakeholders should be addressed. Roles or alignment of areas of concerns should be addressed. For example, business should define business requirement aspects to be incorporated in a DW/BI module; industry should define the technical requirement aspects that should be incorporated in a DW/BI module; universities should control teaching and academic aspects; and former students should give input into course content (Emancipation).

A DW/BI module should reflect the views of universities, industry, former (past) students, and business (Worldviews).

Transforming universities' DW/BI modules (education) from the current to ideal state is the ultimate goal of the improvement process. The application of CSH helped to describe the current and ideal state of DW/BI education; in addition, it also helped to identify the input elements of the improvement process. These are unresolved boundary issues which can or should change in the current state.

8.4 Soft systems methodology application

This section presents the application of SSM in this study; data presented in Chapter 7 are used. Chapter 6 includes a discussion on SSM; SSM was developed by Checkland (2011:497). SSM enables different stakeholders to create models that represent their views on improving the problem situation. These models are then compared with the real problem situation in order to find accommodative improvement for the identified problem.

Chapter 7 includes the results of the interviews held with participants; participants are representative of three groups of people. The first group consists of former (past) students (P5, P6, P7, P8, P10, and P16) that studied DW/BI at an university. The second group consists of DW/BI professionals; they are now working in the

industry/business and may or may not have a formal DW/BI education (P2, P3, P4, P8, P9, P10, P11, P12, P13, P14, P15, P16). The third group consists of faculty members (P1, P9, P13) or academic professionals (lecturers). Participant P9 plays two roles, i.e. lecturer at university and DW/BI professional in the industry.

8.5 SSM application to this study

This section focuses on specific SSM activities: rich pictures, CATWOE, root definitions, and activity models.

8.5.1.1.1 Rich picture

The interview transcription of one participant from each group (i.e. P1, P3, and P7) is selected to use as a starting position to develop SSM diagrams. This is an implicit process as the participants did not draw these diagrams themselves; they are not knowledgeable on SSM theory. The researcher sent the diagrams to the participants to verify that they reliably represent their viewpoints; participants' enhancements were incorporated and diagrams updated accordingly.

The tacit DW/BI knowledge and understanding demonstrated by the students as well as DW/BI and academic professionals are captured and represented in a rich picture without losing context. Checkland and Poulter (2006:219) provide the following elements as a guideline of what to include in a rich picture:

Structures – refers to aspects that are slow to change, or hard, such as buildings, geographical locations, physical equipment etc.

Processes – refers to the transformations that occur in the problem situation depicted in the rich picture.

Climate – refers to the environment of the problem situation area, such as DW/BI module education environment.

People - refers to all stakeholders that are involved in or affected by the problem situation area.

Issues expressed by people – refers to opinions or views of the people that are involved in or affected by the problem situation. These include conflicting viewpoints of people in the problem situation.

The above elements are used later in this chapter to compare the three rich pictures of the selected participants. Responses of the three interviewees (participants P1, P3, and P7) are used to develop rich pictures for DW/BI education according to a student, a DW/BI professional, and a faculty member respectively. These pictures depict the perspectives on the current (“is”) state of the module. Alphabetical characters are used as labels to indicate different sections of the rich pictures in order to aid the explanation thereof.

8.5.1.1.2 CATWOE

The CATWOE mnemonic is used to define: the customer (C), actors (A), transformation process (T), Weltanschauung (the worldview) (W), ownership (O) and environmental constraints (E) (Checkland & Poulter, 2006:220). The CATWOE analysis is used to articulate the scope and components of DW/BI education.

8.5.1.1.3 Root definition

Following the presentation of rich pictures and CATWOE, the next step is to formulate the root definition (RD) of the DW/BI module being studied. A RD is a structured definition of a DW/BI education system. The PQR formula is used as the *enriching process* for the RD. The RD enrichment is achieved by answering questions such as: “what?”, “how?”, and “why?”. Checkland and Poulter (2006:219) give the transformation formula PQR as meaning the following: Do P (what?) by Q (how?) in order to help achieve R (why?). In this DW/BI education problem, the researcher wants to *improve the DW/BI syllabus (P) by involving stakeholders guided by research methodology (Q) in order to better prepare students for a career in the DW/BI industry.*

8.5.1.1.4 Activity diagram

Activity diagrams are derived from each CATWOE and RD, as well as other data presented in Section 8.2 of this chapter and in Chapter 6. The models describe what the DW/BI module in question must “do” in order to have an improved DW/BI module. The models do not describe the existing DW/BI education; rather, they are the logical set of activities that can be carried out in a DW/BI process.

In Chapter 6 the theme, i.e. Theme 23 (BI process), focuses on the steps that can be taken to improve a DW/BI module. Table 8-2 below presents steps that are recommended (Chapter 7, Section 7.5.1.23– Theme 23) to redesign or improve a DW/BI module. This process can serve as a guideline for universities.

Table 8-2: Steps or process to redesign DW/BI module

No	BI process
1.	Identify and involve relevant stakeholders
2.	Check DW/BI literature and trends (other independent sources)
3.	Define industry requirements
4.	Structure and develop the programme
5.	Define success measures
6.	Choose technology to teach
7.	Conduct course comparison analysis
8.	Define the module prerequisites and get better students
9.	Draw insight from independent solutions in industry
10.	Identify core skills needed
11.	Incorporate big business models
12.	Get former (past) students to review
13.	Review periodically

These steps are then applied in drawing up the DW/BI process activity model. Figure 8-1 below presents an activity model according to Theme 23.

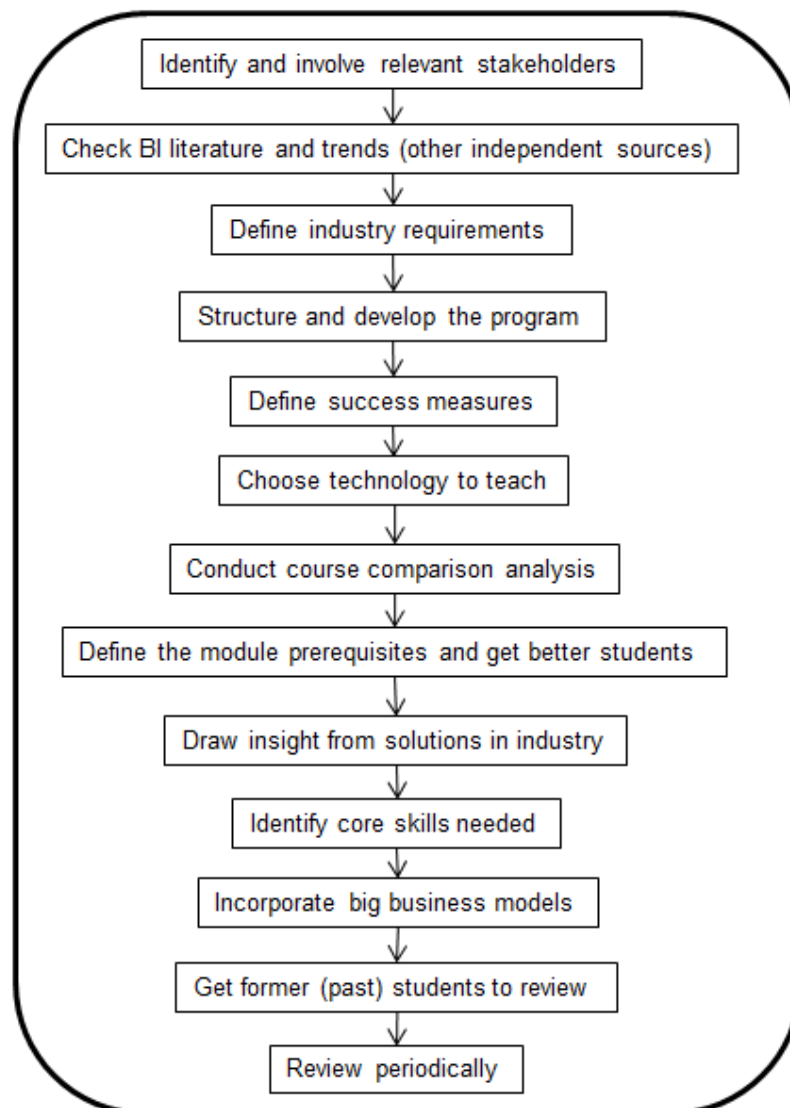


Figure 8-1: DW/BI process activity model

The next section presents the application of these elements in this study.

8.5.1.1.5 Application of SSM in this study

The elements of SSM are used in this study to describe the different viewpoints of various stakeholder groups. Table 8-3 below presents information on the application of the SMM elements of rich pictures, CATWOE analyses, RDs and activity diagrams.

Table 8-3: SSM activities

Activity	Application to study
Rich picture	Rich pictures are created from initial interviews and analysed data from Chapter 7. This step is presented in Sections 8.6.1.1.1, 8.7.1.1.1 and 8.8.1.1.1 for the three groups of stakeholders respectively.
CATWOE	CATWOE analyses are performed from the chosen interviews and enriched from analysis of the interview data of other participants representing a specific stakeholder group from Chapter 7. This step is presented in Sections 8.6.1.1.2, 8.7.1.1.2 and 8.8.1.1.2 for the three groups of stakeholders respectively.
Root definition	This is a structured and rich definition of a DW/BI education system. This activity is presented in Sections 8.6.1.1.3, 8.7.1.1.3, and 8.8.1.1.3 for the three groups of stakeholders respectively.
Activity diagram	It is the logical set of activities that can be carried out in order to improve the DW/BI module. This activity is presented in Sections 8.6.1.1.4, 8.7.1.1.4 and 8.8.1.1.4 for the three groups of stakeholders respectively.

8.6 Student perspective

This section presents a rich picture, CATWOE, RD, and activity model for the student group (P5, P6, P7, P8, P10, and P16) that studied DW/BI at universities.

8.6.1.1.1 Rich picture

Figure 8-2 below presents the current DW/BI module (education) rich picture according to a student's perspective (participant P7). A section with label "A" shows a student enrolling for a DW/BI module. A section with label "B" shows faculty members planning and developing a DW/BI module without the involvement of industry experts. A section with label "C" shows a DW/BI module lecture room environment and its activities. A lecturer is teaching the student DW/BI theory with less exposure to tools and practical work. A student writes exams in theory and practical work (i.e. develop a DW/BI solution). A section with label "D" shows a student graduating upon passing both theory and practical exams. Lastly, a section with label "E" shows the student joining the industry. According to the industry manager's assessment, the industry benefits to some extent from the graduates' knowledge and skills.

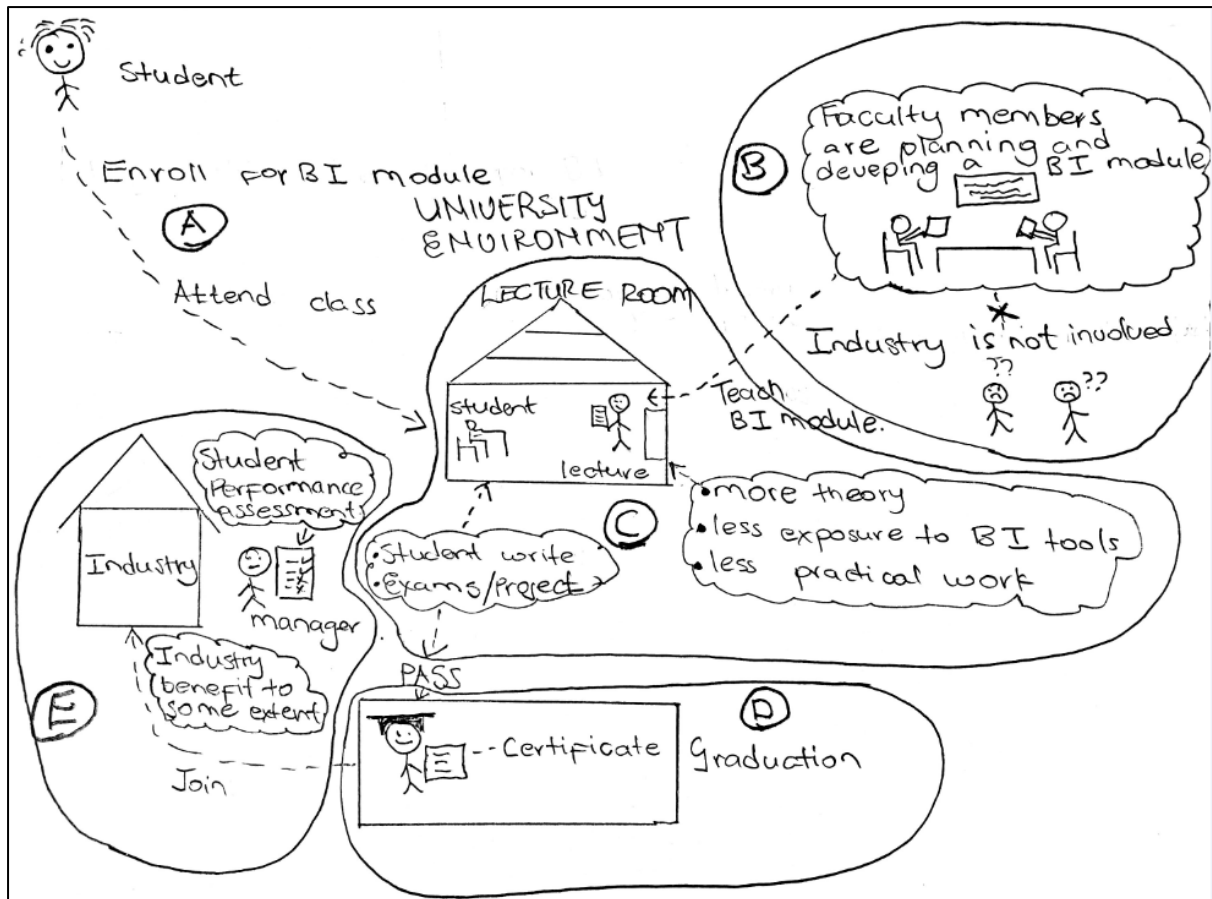


Figure 8-2: DW/BI education rich picture according to student's perspective (P7)

The SSM elements identified in the student's rich picture are presented in Table 8-4 below.

Table 8-4: SSM rich picture elements according to student

Elements	Item
Structures	University lecture room (Label C) Industry work place (Label E)
Processes	Students enrolment (Label A) DW/BI module development (Label B) DW/BI module teaching and learning (Label C) Graduation (Label D) Students joins industry (Label E)
Climate	Nothing explicitly presented
People	Student (Label A) Industry manager (Label E) Faculty members (Label B and C) Business people (Label B)
Issues expressed by people	Manager says industry benefits to some extent (Label E)

	Industry is not involved in the DW/BI module development (Label B). Student(s) is taught a lot of theory (Label C) Student(s) is taught less of practical work (Label C) Student(s) is less exposed to DW/BI tools (Label C)

The first column lists the elements; the second column contains items. Each item has an alphabetical character as label next to it, to show where it is identified in the picture. Participant P7 identifies two 'structures' items, five 'processes' items, four groups of people, and five issues expressed by people. Since the researcher draws the rich picture from participant P7's data, no climate elements are identified. The researcher does not have experience of the DW/BI module environmental climate, and nothing in the data represent this aspect.

8.6.1.1.2 CATWOE

Table 8-5 below shows a CATWOE analysis for the student groups. It is compiled from the chosen interview with participant P7 (round bullets in Table 8-5) and enriched (square bullets in Table 8-5) from analysis of the interview data of other participants representing a specific stakeholder group as per Chapter 6. The first column presents the elements of CATWOE as well as the theme used from Chapter 6. The first element is customers (C) (i.e. who are the beneficiaries of DW/BI education). For example, the customer (C) component is based on Themes 3 and 8 (*"Ideal beneficiaries"*, *"Affected stakeholders"*) in Table 6.14 and Table 6.23 of Chapter 6. The interview data of participant P7 is used to provide a basis answer. He/she identifies *Business* and *Students* as customers (round bullets). Participant P6, who is also a member of the student group, identifies *Business analyst*. *Business analyst* is added as enrichment to the customer component (square bullets). Other participants of the student group did not give a respond on *"Ideal beneficiaries"*; hence, *Business analyst* is the only enhancement to the customer (C) component.

The second element is actors (A) (who are involved in the DW/BI education?). The actor (A) component is based on Theme 10 (*"Involved stakeholders"*). The third element is transformation (T) (i.e. what transformation needs to happen in DW/BI education?). The fourth element is worldview (W) (i.e. what is the perspective from

which the root definition is formulated?). The fifth element is owners (O) (i.e. who control DW/BI education?). The last element is environment (E) (i.e. what constraints do we have to live with in DW/BI education?). The owner (O) component is based on Theme 5 (*“Decision makers”*).

Table 8-5: DW/BI CATWOE according to students' group

CATWOE	
Component	Student
C – Customers Theme_3_Beneficiaries Theme_8_AffectedStakeholders	• Business • Students ▫ Business analysts
A – Actors Theme_10_Involvedstakeholders	• Universities (faculty members) • BI professionals • Business ▫ Former students
T – Transformation process	• Developing DW/BI programme that can better prepare students for industry.
W – Worldview	• Universities can improve DW/BI module to meet the needs of industry.
O – Owners Theme_5_DecisionMakers	• Industry (experts) ▫ University (faculty members)
E – Environmental constraints Theme_6_ConditionsOutsideControl	• Universities can't control DW/BI technology and its trends (hardware and software licensing costs). • Lack the practical component. ▫ Universities can't control industry input of real life problems. ▫ Industry can't control programme requirements. ▫ Industry can't control modules to complement DW/BI. ▫ Universities can't control technical expertise.

8.6.1.1.3 Root definition

The following root definition for the student group is formulated from the rich picture and CATWOE.

Students enrolling for a DW/BI module should be better prepared with knowledge and skills. Universities (faculty members), DW/BI professionals, business, and former (past) students should each give their input in the development of a DW/BI module. Although universities do not have control over DW/BI trends, technology, lack of practical work, industry input and technical expertise; and industry do not

have control over programme requirements and modules that complement DW/BI; the module should accommodate all affected stakeholders such as business, students, industry, and DW/BI professionals. Students should have Kimball model knowledge, and be able to build a complete DW/BI solution so that their impact in the industry can be significant and satisfactory to the DW/BI manager. This process should be controlled and owned by industry (experts) and universities (faculty members).

8.6.1.1.4 Activity diagram

Figure 8-3 depicts the conceptual model according to the student group (P5, P6, P7, P8, and P16). In Theme 23 (Chapter 6) the student group identifies seven main activities that can be carried out to improve a DW/BI module: *Identify and involve relevant stakeholders*; *identify core skills needed (i.e. industry needs)*; *conduct a comparison analysis of modules with other universities*; *structure and develop the programme*; *former (past) students to review*; *choose technology to teach*; and *check DW/BI literature and trends*. These steps are used as a basis for an activity model. Model enhancement is achieved by looking at CATWOE and RD (dark grey activities –Figure 8-3), and other themes (light grey) from Chapter 7 that addresses DW/BI processes. Figure 8-3 below shows the activity diagrams with data labels (themes) and references to Chapter 6 (in order to ground the activities in the interview data).

Other activities are sub-activities supporting the main activities that address processes. The first main activity is to *identify and involve relevant stakeholders (Theme 23)*. This activity is enhanced by the CATWOE components from Table 8-5. The actors (A), customers (C), and owners (O) are the stakeholders. The actor (A) component's activity is supported by identifying the expertise they should bring into the process, in Theme 11.

Theme 21 in Chapter 7 addresses the aspect of what content to teach in a DW/BI module in support of the *identify core skills needed (i.e. industry needs)* activity. No other themes, CATWOE, or RDs address aspects that are linked to the *conduct a comparison analysis of modules with other universities (Theme 23)* main activity.

The *structure and develop the programme (Theme 23)* main activity is supported by information from the analysis (dark grey shading) that informs the main activities. The researcher adds other activates (light grey shading) from the data addressing or supporting the main activity. This group addressed aspects such as *define prerequisites (Theme 15)*, *define guarantor (Theme 14)*, *define assessment component (Theme 14)*, *define electives (Theme 16)*, *define module content (Theme 21)*, and *define success measures (Theme 13)*.

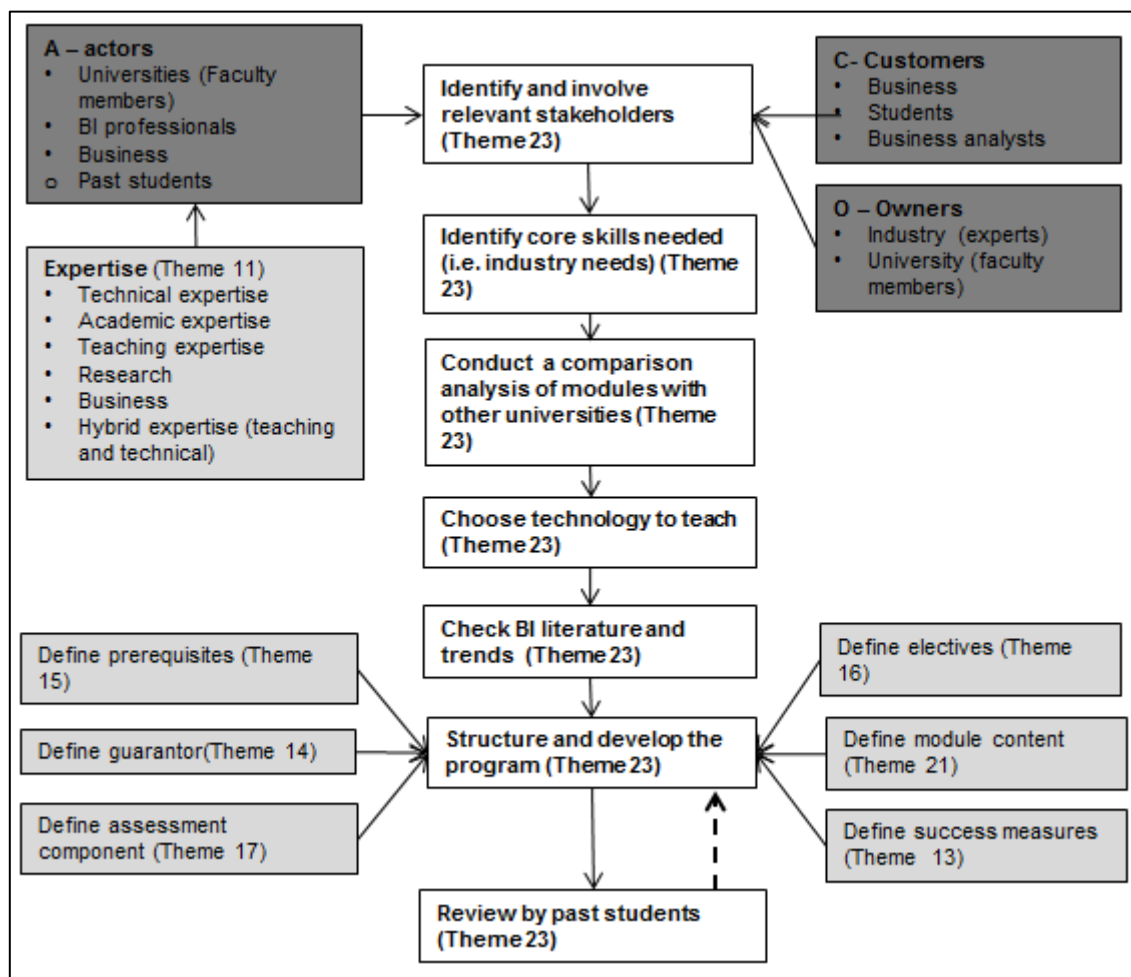


Figure 8-3: Conceptual activity module according to student group

The students' perspectives are compared with others' perspectives in Section 8.9 later in this chapter.

8.7 BI professional perspective

This section presents a rich picture, CATWOE, RD, and activity model for the DW/BI professional group (P2, P3, P4, P8, P9, P10, P11, P12, P13, P14, P15, P16) that are currently employed in the industry; they may or may not have had formal DW/BI education.

8.7.1.1.1 Rich picture

Figure 8-4 below presents the current DW/BI education rich picture according to a DW/BI professional's perspective (participant P3). A section with label "A" shows university faculty members developing a DW/BI module without the input of industry or business experts. A section with label "B" shows a student enrolling for a DW/BI module. A section with label "C" shows the lecture room environment. Students are taught DW/BI theory and practical work. The student writes theory and practical (i.e. develop DW/BI solution) exams. A section with label "D" shows the student graduating upon passing both the theory and practical exams. A section with label "E" shows industry that recruits a graduate to join industry's workforce. Industry is recruiting a graduate but does not add towards building the DW/BI module knowledge ("A"). The manager is to some extent happy. This is because he feels that it takes the graduate close to three years to know his/her role in DW/BI and to add value to business.

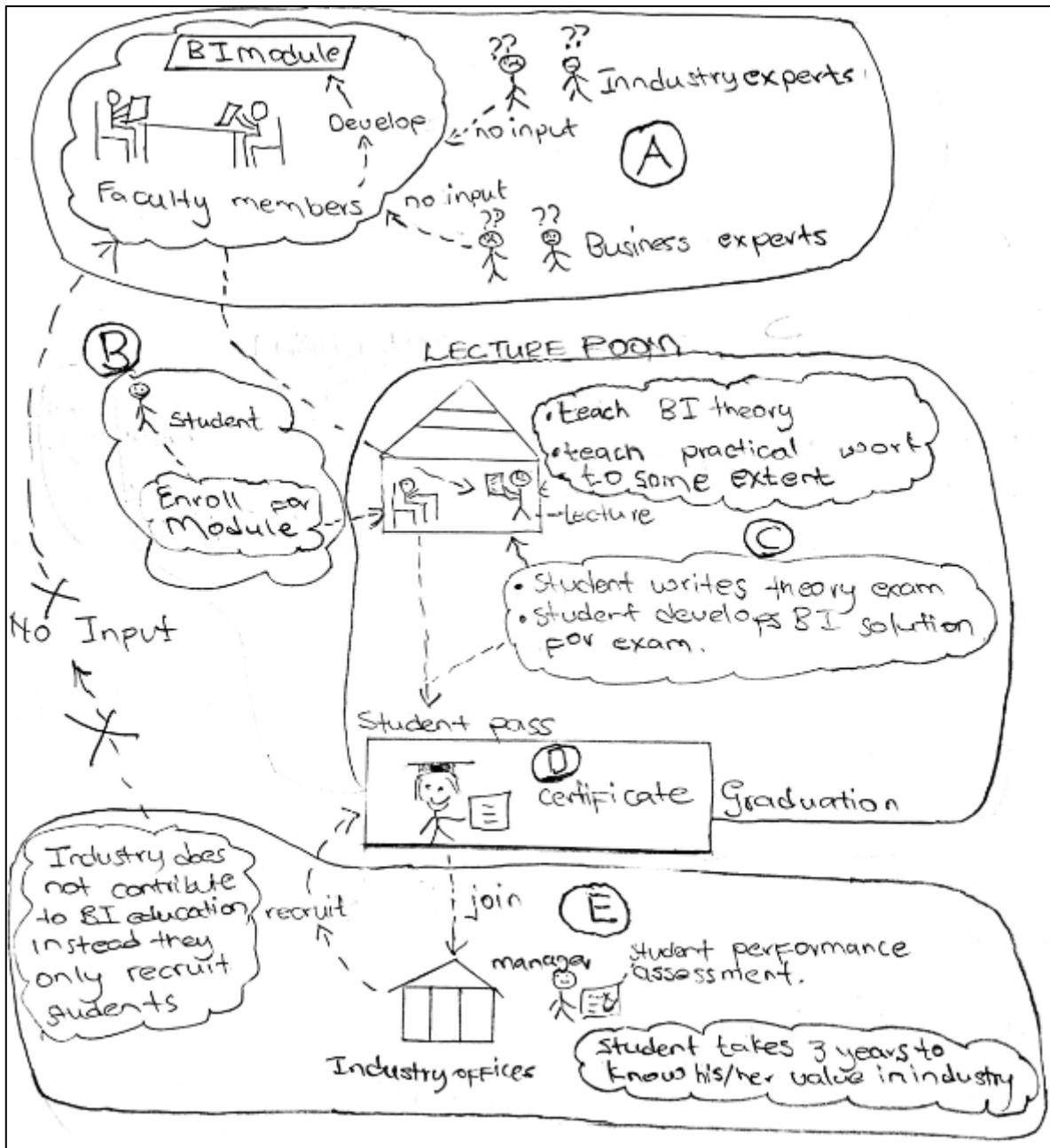


Figure 8-4: DW/BI education rich picture according to DW/BI professional (P3) perspective

The rich picture elements are presented in Table 8-6 below. Each item has an alphabetical label next to it to show where can be identified in the picture. Participant P3 identifies two 'structures' items, five 'processes' items, five groups of people, and five issues expressed by people. No climate elements are identified, because the researcher drew the rich picture from participant P3's data.

Table 8-6: SSM rich picture elements according to DW/BI professional (P3)

Elements	Item
Structures	University lecture room (Label C) Industry work place (Label E)
Processes	DW/BI module development (Label A) Students enrolment (Label A) DW/BI module teaching and learning (Label C) Graduation (Label D) Students joins and works (Label E)
Climate	Nothing explicitly presented
People	Industry experts (Label A) Business experts (Label A) Student (Label B) Faculty members (Label A and C) Industry manager (Label E)
Issues expressed by people	Industry is not involved in the DW/BI module development. (Label B) Student(s) is taught a theory (Label C) Student(s) is to some extent taught practical work (Label C) Student takes three years to know his/her value and to add value in the industry (Label D) Industry recruits graduates from universities but it is not contributing towards building the academic field (Label E)

CATWOE

Table 8-7 below shows a CATWOE for the DW/BI professional group. The researcher compiles it from the chosen interviews (round bullets in Table 8-7) and enriches it from analysis of the interview data of other participants (square bullets) from Chapter 7. The first column presents the elements of CATWOE as well as the theme of Chapter 6 used. The first element is customers (C) (i.e. who are the beneficiaries of DW/BI education). The second element is actors (A) (i.e. who are involved in the DW/BI education?). The actor (A) component is based on Theme 10 (*“Involved stakeholders”*) in Table 7-27. The third element is transformation (T) (i.e. what transformation needs to happen in DW/BI education?). The fourth element is worldview (W) (i.e. what is the perspective from which the root definition is formulated?). The fifth element is owners (O) (i.e. who control DW/BI education?). The last element is environment (E) (i.e. what constraints do we have to live with in DW/BI education?). The owners (O) component is based on Theme 5 (*“Decision makers”*) in Table 7-18.

Table 8-7: DW/BI CATWOE according DW/BI professionals group

CATWOE	
Element	BI professional
C – Customers Theme_3_Beneficiaries Theme_8_AffectedStakeholders	• Industry • Students • Business ▫ BI professional
A – Actors Theme_10_Involvedstakeholders	• Universities (faculty members) • BI professionals • Business ▫ Former (past) students
T - Transformation process	• Developing DW/BI programme that can better prepare students for industry.
W - Worldview	• Universities can improve DW/BI module to meet the needs of industry.
O – Owners Theme_5_DecisionMakers	• Industry (experts) • Business ▫ University (faculty members) ▫ Students
E - Environmental constraints Theme_6_ConditionsOutsideControl	• Universities can't control packaging of DW/BI programmes such that it adds value. ▫ Universities can't control industry input of real life problems. ▫ Universities can't control student's employability. ▫ Universities can't control technical expertise. ▫ Industry has no control how academic programmes should run. ▫ Universities can't control cost of hardware and software. ▫ Industry can't control modules to complement DW/BI.

8.7.1.1.2 Root definition

The following root definition for the DW/BI professional group is formulated from the rich picture and CATWOE.

BI education owned by industry, business, universities and students should be improved to better prepare students. This is undertaken to accommodate the needs of all affected such as business, students, and business, and former (past) students. Constraints such as that universities cannot control industry input (technical expertise) of real life problems, students' employability, and technology; also industry

doesn't have control over how academic programmes should run; should be taken into consideration. Students should be able to build a complete DW/BI solution so that their impact in the industry can be significant and satisfactory to the DW/BI manager. A professional DW/BI body to oversee the input and output of DW/BI education should serve as a success measure.

8.7.1.1.3 Activity diagram

Figure 8-5 below presents a conceptual model according to the DW/BI professional group (P2, P3, P4, P9, P11, P12, P14, and P15). This group identifies nine main activities in Theme 23 of Chapter 6 to improve DW/BI module: *Identify and involve relevant stakeholders, define industry requirements, identify core skills needed, structure and develop the programme, choose technology to teach, define success measures, check DW/BI literature and trends, draw insight from independent solutions in industry, and review periodically.*

The first main activity is to *identify and involve relevant stakeholders (Theme 23)*. This activity is enhanced by the CATWOE components from Table 8-7. The actors (A), customers (C), and owners (O) are listed as stakeholders (dark grey shading). The actors' (A) activity is enhanced by identifying the expertise they should bring into the process in Theme (11). The second main activity is: *defining industry requirement (Theme 23)*. This activity is supported by: *identify core skills need (Theme 23)*.

The *structure and develop the programme* main activity is supported by other activities that the researcher thinks addresses or supports the main activity. This group addresses aspects such as: *define prerequisites (Theme 15), define guarantor (Theme 14), define assessment component (Theme 14), define electives (Theme 16), define module content (Theme 21), and define success measures (Theme 13)*.

The last activity that addresses DW/BI processes is to *review periodically (Theme 23)*. DW/BI professionals recommend that the DW/BI development process be reviewed periodically. This will ensure that the module stays relevant to current needs of the involved and affected stakeholders.

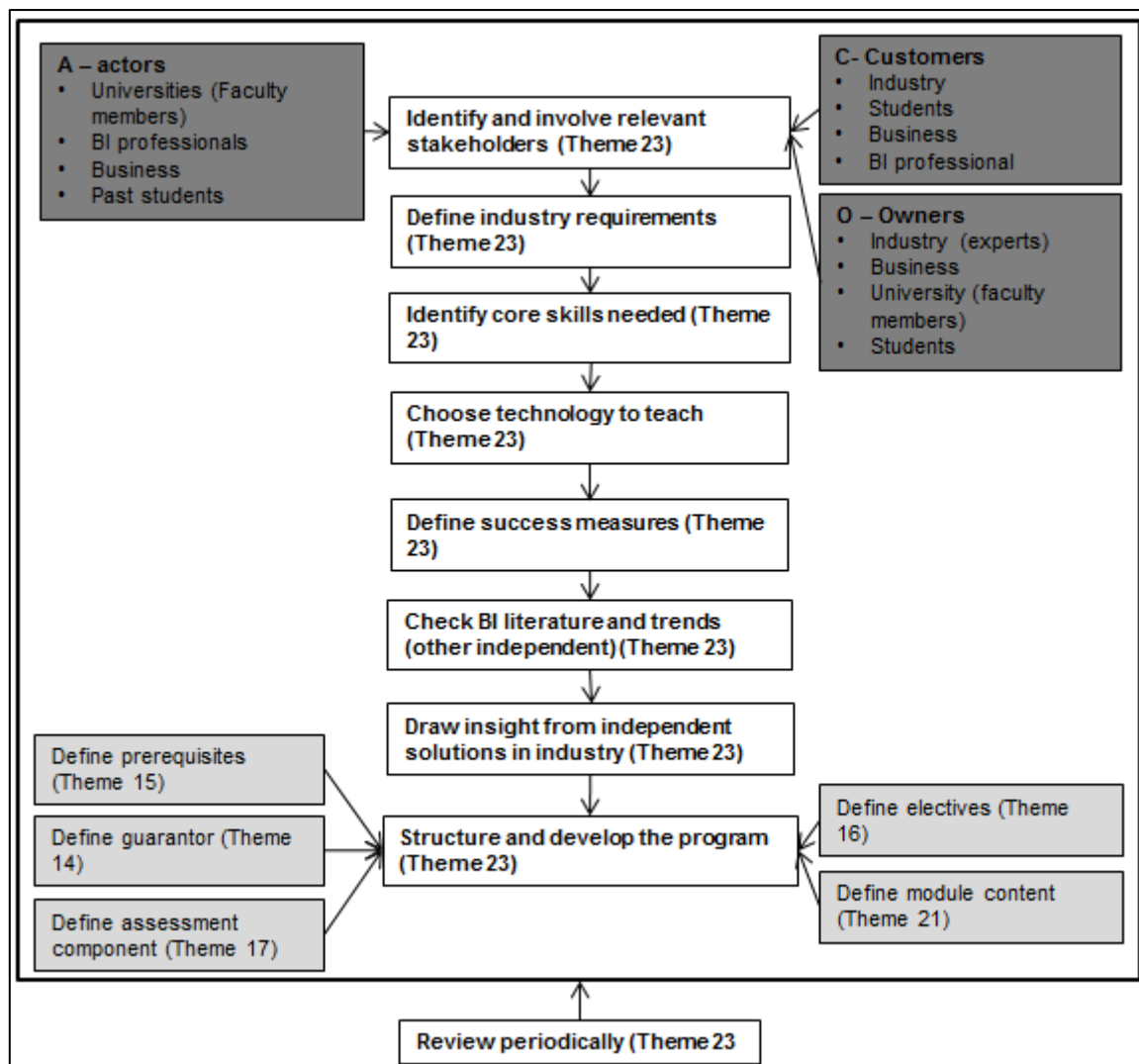


Figure 8-5: Conceptual model according to DW/BI professional

The DW/BI professionals' perspective is compared to other perspectives in Section 8.9 later in this chapter.

8.8 Faculty member perspective

This section presents a rich picture, CATWOE, RD, and activity model for faculty members group (P1, P9, P13). They are academic professionals (lecturers) .

8.8.1.1.1 Rich picture

Figure 8-6 below presents the current DW/BI education rich picture according to faculty members' perspectives (lecturer - participant P1). A section with label "A", shows an industry expert, and faculty members, developing a DW/BI module. The

faculty member (lecturer) feels that the university needs better students to enrol for the DW/BI module. A section with label “B” shows a student enrolling for a DW/BI module. A section with label “C” shows a class room or lecture environment. A student is taught DW/BI theory and practical work by a lecturer that contributed towards the development of a module. Students write both theory and practical (i.e. develop a working DW/BI solution) exams in order to pass. The section with label “D” shows students that passed, which join the industry workforce. The industry benefits from the graduate’s knowledge and skills and the manager is happy. In the section with label “E” the student gives feedback (view) on the module. This feedback is then used as an input towards improving the DW/BI module in “A”.

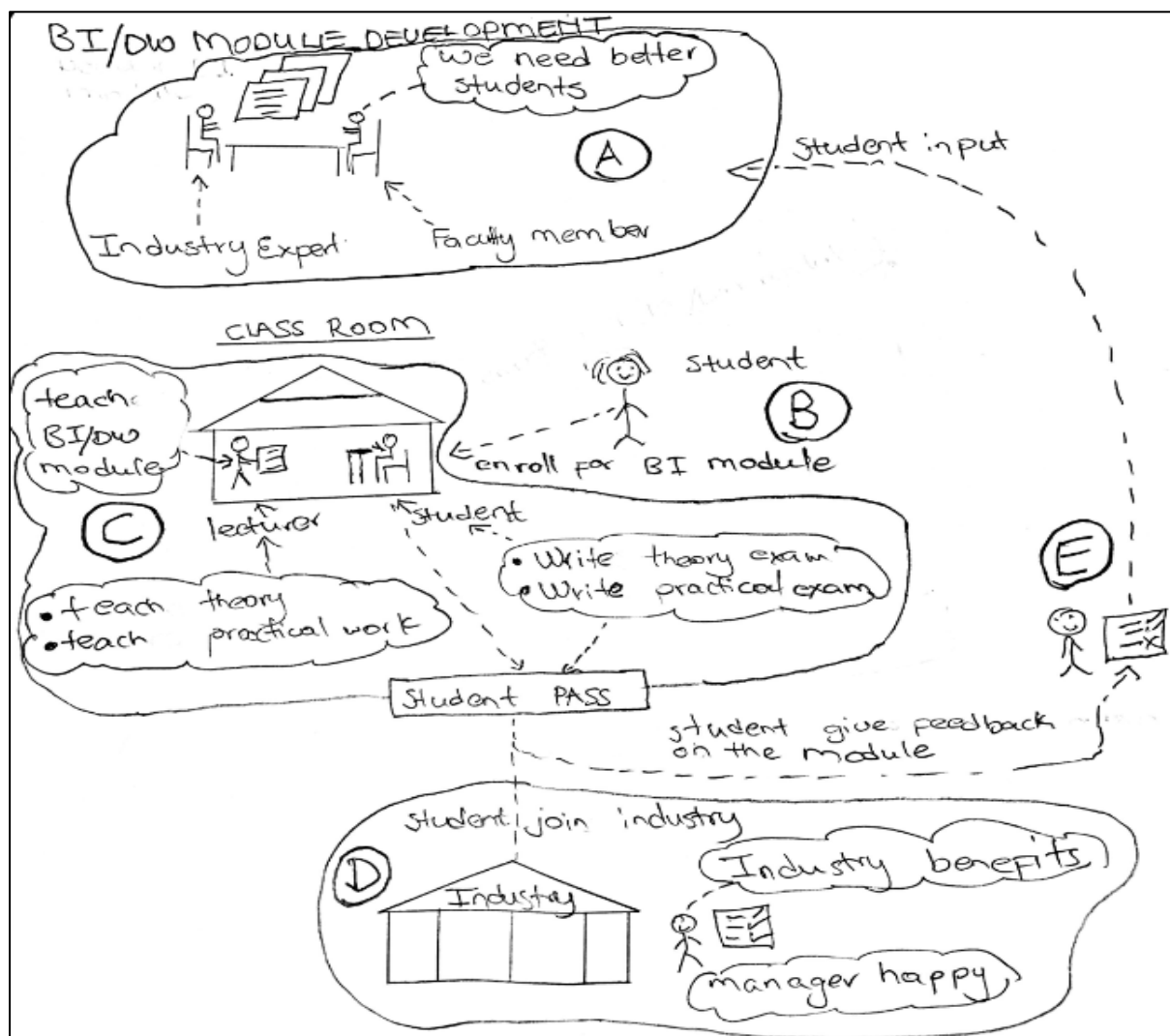


Figure 8-6: DW/BI education rich picture according to faculty member (Participant P1)

The SSM elements that are identified by participant P1's (BI professional) rich picture are presented below. Participant P1 identifies two 'structures' items, five 'processes' items, three groups of people, and five issues expressed by people. No climate elements are identified because the researcher drew the rich picture from participant P1's data. This rich picture is analysed in Table 8-8.

Table 8-8: SSM rich picture elements according to faculty member (P1 - lecturer)

Elements	Faculty member (lecturer)
Structures	University lecture room (Label C) Industry work place (Label D)
Processes	Students enrolment (Label B) DW/BI module development (Label A) DW/BI module teaching and learning (Label A) Students joins and work (Label D) Student feedback (Label E)
Climate	Nothing explicitly presented
People	Student (Label B) Industry manager (Label E) Faculty member (Label A) Industry experts (Label A)
Issues expressed by people	Faculty member (Lecture) thinks universities need better student intake (Label A) Industry is involved in DW/BI module development (Label A) Student(s) is taught theory (Label C) Student (s) is taught practical work (Label C) Industry benefits from graduates quality and the manager is happy (Label E)

8.8.1.1.2 CATWOE

Table 8-9 below shows a DW/BI module CATWOE according to the faculty member group. It is compiled from the chosen interview (P1 – round bullets in Table 8-9) and enriched from analysis of the interview data (square bullets in Table 8-9) of other participants' (square bullets) data from Chapter 7. The first column presents the elements of CATWOE as well as the theme from Chapter 7. The first element is customers (C) (i.e. who are the beneficiaries of DW/BI education?). The second element is actors (A) (i.e. who are involved in the DW/BI education?). The actors' (A) component is based on Theme 10 (*"Involved stakeholders"*) in Table 7-27.

The third element transformation (T) (i.e. what transformation needs to happen in DW/BI education?). The fourth element is worldview (W) (i.e. what is the perspective from which the root definition is formulated?). The fifth element is owners (O) (i.e. who control DW/BI education?). The last element is environment (E) (i.e. what constraints do we have to live with in DW/BI education?). The owners (O) component is based on Theme 5 (*“Decision makers”*) in Table 7-18.

Table 8-9: DW/BI module CATWOE according faculty members

CATWOE	
Component	Faculty members
C – Customers Theme_3_Beneficiaries Theme_8_AffectedStakeholders	• Industry • Students • Business
A – Actors Theme_10_Involvedstakeholders	• Industry experts • Universities (faculty members)
T - Transformation process	• Developing DW/BI programme that can better prepare students for industry.
W - Worldview	• Universities can improve DW/BI module to meet the needs of industry.
O – Owners Theme_5_DecisionMakers	• Industry (experts) ▫ Business ▫ University (faculty members)
E - Environmental constraints Theme_6_ConditionsOutsideControl	• Universities can't control DW/BI technology and its trends (hardware and software licensing costs). • Industry can't control programme requirements. • Universities can't control real data. • Universities can't control industry input of real life problems. ▫ Universities can't control cost of hardware and software.

8.8.1.1.3 Root definition

The following root definition for the faculty group is formulated from the rich picture and CATWOE.

DW/BI education owned by industry (experts), business, and university (faculty members) should be improved to better prepare students. This is undertaken to

accommodate the needs of all affected as business, students, and business, and former (past) students. Constraints such as the fact that universities can't control DW/BI technology, trends and licensing, real data, and industry real life problems should be taken into consideration. Also, industry can't control programme requirements. Students should be able to build a complete DW/BI solution and pass the DW/BI examination so that their impact in the industry can be significant and satisfactory to the DW/BI manager.

8.8.1.1.4 Activity diagram

Table 8-9 below shows activity model according to the university faculty member group (P1, P9, and P13). This group identifies 7 main activities from Theme 23 to improve a DW/BI module: *Start with what we have, identify and involve relevant stakeholders, define industry requirements, structure and develop the programme, define the module prerequisites and get better students, and check DW/BI literature and trends*. Other activities are sub-activities of main activities that address processes.

The first main activity is to *identify and involve relevant stakeholders (Theme 23)*. This activity is enhanced by the CATWOE components from Table 8-9. The actors (A), customers (C) and, owners (O) are stakeholders (greyed out) in Figure 8-7. The actors' (A) activity is enhanced by identifying the expertise that they should bring into the process in Theme (11). No other themes, CATWOE, or RDs addresses aspects that are linked to the *define industry requirements (Theme 23)* main activity.

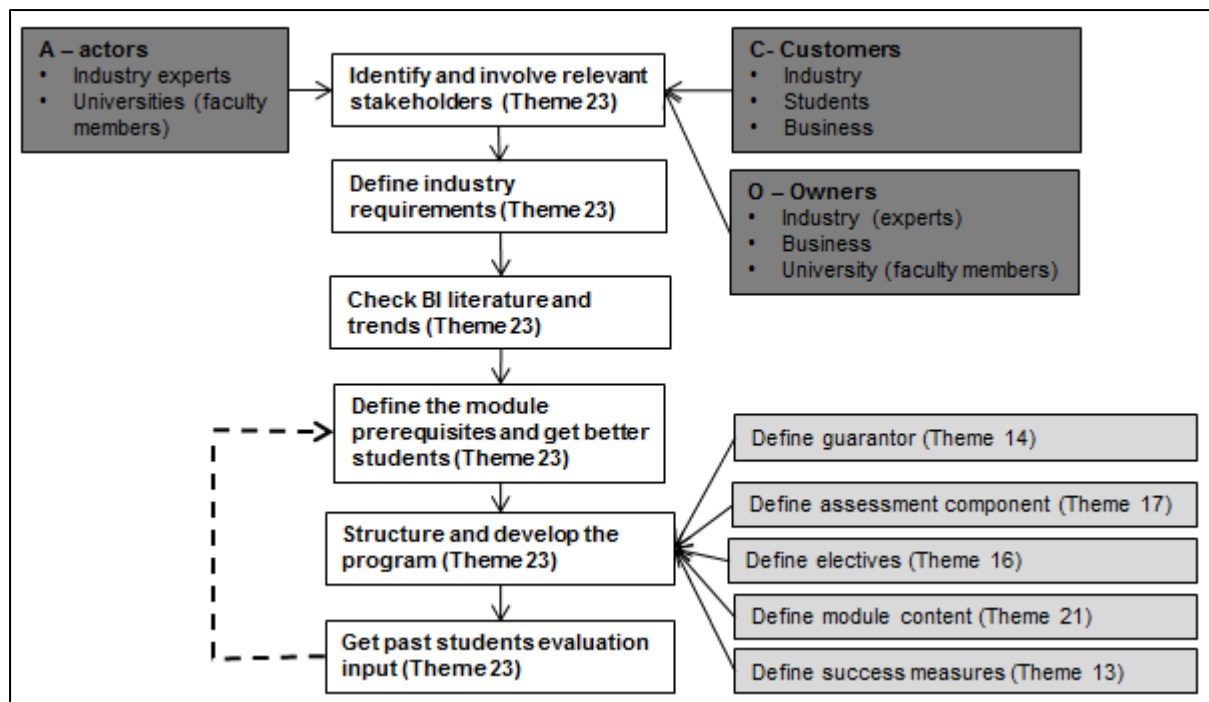


Figure 8-7: A conceptual activity model according to university faculty members (lecturer)

The activity structure and develop the programme (Theme 23) is supported by activities such as define the module prerequisites, get better students, and check DW/BI literature and trends.

The researcher adds other activates that are addressing or supporting the main activity. This group addressed the aspects: define guarantor (Theme 14), define assessment component (Theme 14), define electives (Theme 16), define module content (Theme 21), and define success measures (Theme 13). The last activity is: get former (past) student to evaluate the module.

The faculty members' perspective is compared to other perspectives in the next section.

8.9 Conclusions drawn from SSM activities

A summary of conclusions drawn from a rich pictures, CATWOE's, RD;'s and activity models of the three groups is presented in this section.

8.9.1.1.1 Rich pictures

In this section the rich pictures of participants P1 (faculty member), P3 (BI professional), and P7 (student) are compared. The comparison is done for all picture elements. Table 8-10 below presents the SSM rich picture elements according to the different perspectives. From the table the following can be noted:

Structure - All three participants identify two structures, i.e. university lecture room and industry work place.

Processes – Five processes are identified from each rich picture. DW/BI module development, student enrolment, DW/BI teaching, and student joins work place are common process to all rich pictures. Participant P1 identifies a student feedback process which the P3 and P7 do not.

People – Students, industry manager, and faculty members are common to all three rich pictures. P7 and P3 identify business people with expertise, which P1 do not identify.

Issues expressed – Their shared and contrasting views are identified among the three participants. Participant P3 and P7 share the same view that industry is not involved in the DW/BI module, while Participant P1's view is in contrast with that. Participant P1 has a view that industry benefits from the quality of students that universities produce, while P7 thinks industry benefits but only to some extent.

A common view among the three participants is that students are taught DW/BI theory and practical work.

Participants also share unique views such as that P7 thinks that the student is less exposed to DW/BI tools. Participant P3 shares a unique view that industry recruits graduates from universities but it is not contributing towards building the academic field and that students take at least three years to know their value and add value to business. Participant P1 share a unique view that universities should get better students to enrol for DW/BI education.

Table 8-10: SSM rich picture elements according to student, DW/BI professional and faculty member groups

Elements	Student (P7)	BI professional (P3)	Faculty member (lecturer – P1)
1. Structures	- University lecture room (Label C) - Industry work place (Label E)	- University lecture room (Label C) - Industry work place (Label E)	- University lecture room - Industry work place
2. Processes	- Students enrolment (Label A) - DW/BI module development (Label B) - DW/BI module teaching and learning (Label C) - Graduation (Label D) - Students joins and (Label E)	- DW/BI module development (Label A) - Students enrolment (Label A) - DW/BI module teaching and learning (Label C) - Graduation (Label D) - Students joins and works (Label E)	- Students enrolment - DW/BI module development - DW/BI module teaching and learning - Students joins and work - Student feedback
3. Climate	Nothing explicitly presented		
4. People	- Student (Label A) - Industry manager (Label E) - Faculty members (Label B and C) - Business people (Label B)	- Industry experts (Label A) - Business people (Label A) - Student (Label B) - Faculty members (Label A and C) - Industry manager (Label E)	- Student - Industry manager - Faculty member
5. Issues expressed by people	- Manager says industry benefits to some extent. (Label E) - Industry is not involved in the DW/BI module development. (Label B) - Student(s) is taught a lot of theory (Label C). - Student (s) is taught less of practical work (Label C). - Student (s) is less exposed to DW/BI tools (Label C).	- Industry is not involved in the DW/BI module development. (Label B) - Student(s) is taught a theory (Label C). - Student (s) is to some extent taught practical work (Label C). - Student takes three years to know his/her value and to add value in the industry (Label D). - Industry recruits graduates from universities but it is not contributing towards building the academic field (Label E).	- Faculty member (Lecture) thinks universities need better student intake (Label A) - Industry is involved in DW/BI module development. - Student(s) is taught theory. - Student (s) is taught practical work. - Industry benefits from graduates quality and the manager is happy.

It is evident that the three participants share most views; there are, however, some unique or contrasting viewpoints.

8.9.1.1.2 CATWOE

A summary of CATWOE for students, DW/BI professionals, and faculty members is presented in Table 8-11 below. From this table it can be seen that students, industry, and business are the three main customers (C) identified by the three groups. Business analysts and DW/BI professionals are uniquely identified by the student group and DW/BI professionals respectively.

The student group and DW/BI professional group identify four actors (A) each, whilst faculty members identify only two. Universities are the only common actor (A) that the three groups identify for the actor (A) component. Business and former students are identified by both the student group and the DW/BI professional group.

The three groups share the same transformation goal and that is to develop a DW/BI programme that can better prepare students for industry. They also share the same worldview, i.e. that universities can improve DW/BI modules to meet the needs of industry.

Industry and universities are identified by all three groups to be the owners of DW/BI education. DW/BI professionals and faculty members share the same view, i.e. that business should be the owner of DW/BI education. DW/BI professionals uniquely identify students as additional owners of DW/BI education. The researcher's perception to the choice of students as owners by DW/BI professionals is that former students can add value to improve DW/BI education.

All three groups identify environmental constraints; these can be summarised as follows: technical expertise, data input, and software licencing are the major constraints for universities to deliver better DW/BI education.

Table 8-11: CATWOE of DW/BI education according students, DW/BI professionals and faculty members

CATWOE	Student	BI professional	Faculty members
C- Customers Theme_3_ Beneficiaries Theme_8_ AffectedStakeholders	• Business • Students ▫ Business analysts	• Industry • Students • Business ▫ BI professional	• Industry • Students • Business
A – Actors Theme_10_ Involvedstakeholders	• Universities (faculty members) • BI professionals • Business ▫ Former (past) students	• Universities (Faculty members) • BI professionals • Business ▫ Former (past) students	• Industry experts • Universities (faculty members)
T - Transformation process	• Developing DW/BI programme that can better prepare students for industry.	• Developing DW/BI programme that can better prepare students for industry.	• Developing DW/BI programme that can better prepare students for industry.
W - Worldview	• Universities can improve DW/BI module to meet the needs of industry.	• Universities can improve DW/BI module to meet the needs of industry.	• Universities can improve DW/BI module to meet the needs of industry.
O – Owners Theme_5_ DecisionMakers	▫ Industry (experts) ▫ University (faculty members)	• Industry (experts) • Business ▫ University (faculty members) ▫ Students	• Industry (experts) ▫ Business ▫ University (faculty members)
E - Environmental constraints Theme_6_ Conditions OutsideControl	• Universities can't control DW/BI technology and its trends (hardware and software licensing costs). • Lack the practical component. ▫ Universities can't control industry input of real life problems. ▫ Industry can't control programme requirements ▫ Industry can't control modules to complement DW/BI. ▫ Universities can't control technical	• Universities can't control packaging of DW/BI programmess such that it adds value. ▫ Universities can't control industry input of real life problems. ▫ Universities can't control student's employability. ▫ Universities can't control technical expertise. ▫ Industry has no control how academic programmes should run. ▫ Universities can't	• Universities can't control BI technology and its trends (hardware and software licensing costs). • Industry can't control programme requirements. • Universities can't control real data. ▫ Universities can't control industry input of real life problems. ▫ Universities can't control cost of hardware and software.

	expertise.	control cost of hardware and software. ▫ Industry can't control modules to complement DW/BI.	

The next sections present the RDs for each group (students, DW/BI professionals, and faculty members (lecturers)).

8.9.1.1.3 Root definition

The three RDs (student group, DW/BI professional group, and faculty member group) are compared in this section. All three groups have the same transformation goal, i.e. to better prepare students with skills and knowledge. The differences are found in owners, beneficiaries, affected stakeholders of DW/BI education, and environmental constraints. For example, the student group identifies that DW/BI education is owned by *industry (experts)*, and *university (faculty members)*. The DW/BI professional group identifies *industry, business, universities and students*. The faculty member group identifies *industry, business and universities*.

The following comprehensive RD is formulated from the three RDs presented in the previous section.

Universities (faculty members), DW/BI professionals, business, former (past) students and business analysts should collaborate to develop a DW/BI module that will benefit all affected stakeholders (Industry, students, business, DW/BI professional). The module must be developed within the academic standards and other constraints (programme requirements, DW/BI trends, real life problems, real data, technical expertise input, programme requirements, students' employability, and technology and licensing). This process should be controlled and owned by industry (experts), business, university (faculty members), and students. Students should be able to build a complete DW/BI solution and pass DW/BI module examination so that their impact in the industry can be significant and satisfactory to

the DW/BI manager. A professional DW/BI body to oversee the input and output of DW/BI education should serve as a success measure.

8.9.1.1.4 Activity diagrams

Checkland and Poulter (2006:11) argue that the goal of SSM is to seek accommodation rather than consensus; this accommodation must be desirable and culturally feasible to the all parties involved. For example, *identify and involve relevant stakeholders*, and *structure and develop the programme*, are the two common activities among the three activity diagrams. Other activities are not common to all the three groups. For example, the activity *review periodically* (*Theme23*), is only identified in the DW/BI professional's conceptual model (Figure 8-5), and not in the other two models (Figure 8-3 and Figure 8-7). The differences between the three models create the need for the fourth conceptual model that will be accommodative of the three groups' views. Figure 8-8 below presents an accommodative model that is drawn from the models of the three groups.

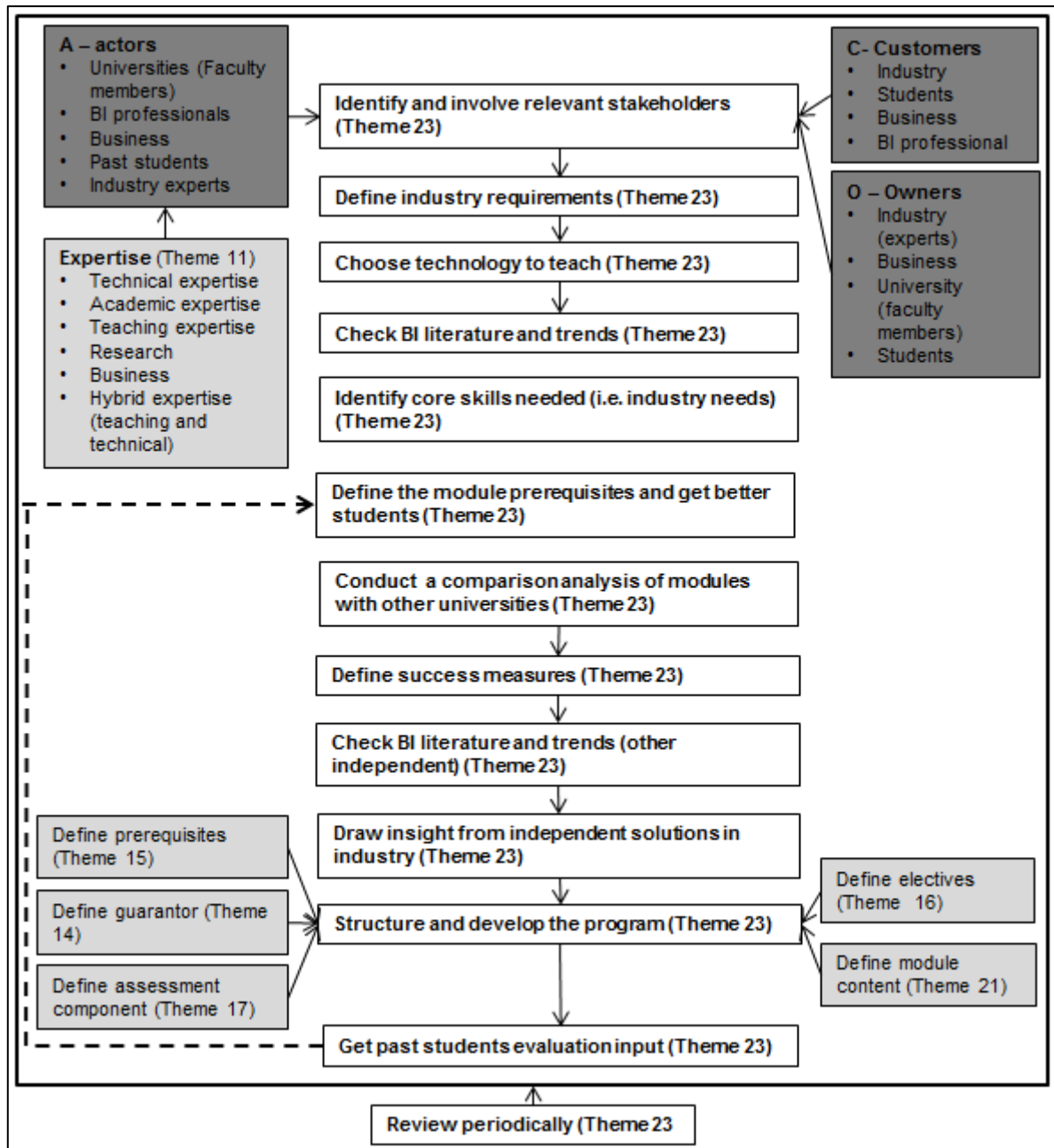


Figure 8-8: Accommodative activity model

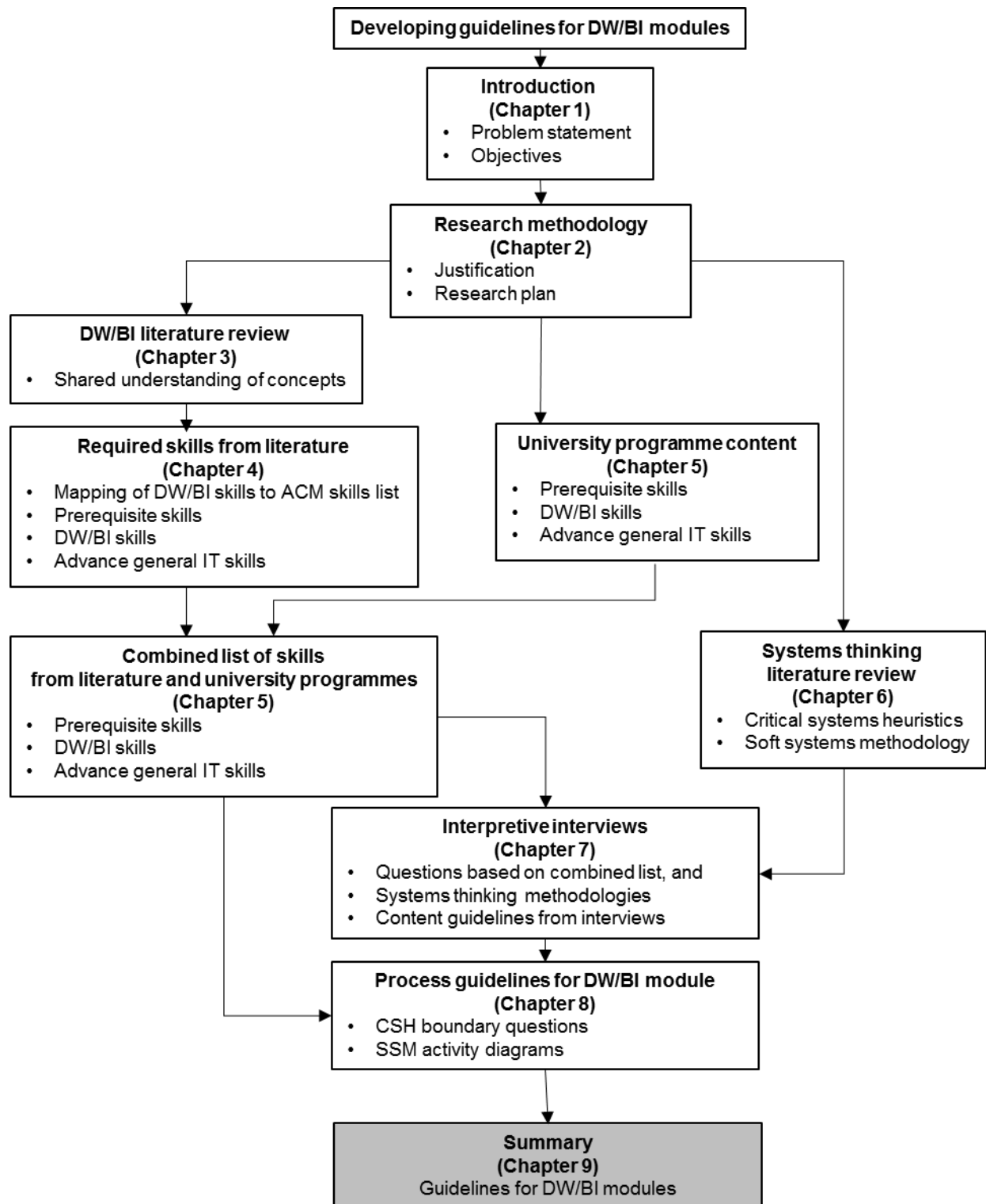
8.10 Summary

This chapter is a continuation of Chapter 7. Data gathered, analysed, and the presentation of findings – refer to Chapter 6, are used as the basis of this chapter. In this chapter CSH and SSM models are used to frame or defined the boundaries of what is and is not considered important. CSH is used to create a picture of the current and ideal DW/BI module state. The boundaries are categorised into four

functional groups, i.e. sources of motivation, power, knowledge, and legitimation. DW/BI modules should be driven more towards the ideal state.

The application of SSM helped in defining the process (purposeful activity model) to be followed in order to improve DW/BI modules. Reflective stakeholder models that are purposeful to the problem situation are created. This is achieved by getting input from the three groups (students, DW/BI professionals and university faculty members). Each group's input is used to create their view of DW/BI module. A comprehensive purposeful activity model that accommodates different views could be created.

The next chapter serves as summary and reflection for the study.



Chapter 9 Summary

9.1 Introduction

The problem addressed by this study is the difficulty that traditional universities experience to provide the data warehouse/business intelligence industry with graduates that are able to contribute to organisations without substantial further training. The complexity of the situation arises from the different goals of the involved and affected stakeholders. The universities' faculty members are responsible to develop modules in data warehouse/business intelligence that satisfy both university programme requirements and the needs of the data warehouse/business intelligence industry. In order to provide a solution to this problem the researcher developed guidelines for data warehouse/business intelligence modules in information technology programmes. These guidelines cover both content and process matters.

The purpose of this chapter is to present these guidelines and summarise the process of their development. The research objectives are revisited in terms of the content of the various chapters. The final set of guidelines for module content and the process that should be used to improve data warehouse/business intelligence education are presented. A brief overview of the implications of this study is presented. Research limitations that are evident are listed, and recommendations for further study are made. This chapter concludes with some suggestions for future research.

This chapter is organised as follows: The research objective of this study is revisited in Section 9.2. Final conclusions for the study (i.e. a combined set of guidelines) are presented in Section 9.3. Further study and recommendations are presented in Section 9.4. Section 9.5 presents the assumptions and limitations of the study. Section 9.6 provides a summary for the chapter.

9.2 Research objectives revisited

The main objective of this study was to develop guidelines for data warehouse/business intelligence (DW/BI) modules in information technology (IT)

programmes at universities; the guidelines were based on an understanding of the needs of various stakeholders. Two methodologies, i.e. critical systems heuristics (CSH) and soft systems methodology (SSM), were used to understand the various stakeholders' perspectives.

9.2.1 Theoretical objectives

The following theoretical objectives were set for this study:

1. Demonstrate an understanding of what constitutes a guideline as well as an understanding of general research methodology in order to justify the selection of the interpretive research paradigm.

Information on the nature of guidelines was provided in Chapter 2. Principles (as discussed in Chapter 2; refer also to Table 2-1) were applied in this study. The researcher based the justification and selection of the interpretive research design on the ontological and epistemological assumptions of popular research paradigms – refer to Chapter 2. A discussion of interpretive research, and the application thereof in this study, was provided in Chapter 2, Table 2-3. The discussion on interpretive research methods led to the presentation of the research plan of the study – refer to Figure 2-9. This figure was repeated on the adjacent pages at the beginning of each chapter to provide context to the study.

2. Demonstrate an understanding of DW/BI components within the IT discipline.

Business Intelligence (BI) is presented in this study as an umbrella term that includes the theory of data warehousing as well as other tools to support business decision making. Since the inclusion of data warehousing in the definition is not done by all scholars, it was decided to use the label “DW/BI” to represent this holistic view of DW/BI. This is in accordance with the work of Kimball *et al.* (2008a:3). Chapter 3 provided a detailed description of DW/BI architecture and the DW/BI lifecycle components. The main objective of Chapter 3 is to provide a shared understanding in order to aid understanding of the responses of the participants, as was reported in

Chapter 7. After completion of the literature review of DW/BI, initial guidelines were developed for the module content – refer to Chapter 4.

3. Demonstrate an understanding of SSM and CSH in the context of systems thinking.

The complex nature of the problem situation demanded a holistic approach where the perspectives of different stakeholders could be represented. Systems thinking methodologies provide tools for holistic thinking and stakeholder identification. So, in Chapter 6 the SSM and CSH methodologies were presented to create a shared understanding of their underlying concepts. These concepts, and the literature review of DW/BI, guided the interview question (presented in Chapter 7). CSH and SSM were used in Chapter 8 to present different worldviews of different groups of stakeholders in terms of the process to be followed when designing a module in DW/BI. In reflection of the study, the researcher is of the opinion that the systems thinking methodologies used in the study aided his holistic view of the problem environment. However, the lack of knowledge of the participants regarding the methodologies created specific challenges.

9.2.2 Empirical objectives:

The following empirical objectives are presented for this study:

- 1. Investigate what DW/BI modules and programmes universities currently teach, by collecting and analysing document data from universities' DW/BI programmes.*

The aim of Chapter 5 was to investigate the current DW/BI programmes and modules at universities. Data were collected using interpretive document data collection as a method. The documents were coded and they were compared to the guidelines developed in Chapter 4 from the DW/BI literature. The result of Chapter 5 was three guidelines that each contained a list of skills that students should acquire: prerequisite general IT skills; DW/BI specific skills; and advanced general IT skills.

2. Investigate the perspectives of stakeholders, using CSH and SSM, as well as understand the needs of the stakeholders in terms of DW/BI module content and the process to design a DW/BI module. Although the aim of this objective is to identify stakeholders, it is intuitively assumed that the stakeholders will at least include industry, universities, and former students.

Concepts of systems thinking methodologies (presented in Chapter 6) were used, together with guidelines from literature and universities' DW/BI modules (presented in Chapter 5), to develop an interpretive interview; this interview aimed to understand the perspectives of different stakeholders. DW/BI professionals, former students, and faculty members took part in the study. Data were analysed using content analysis. Findings were linked to codes developed in the data. The data collection and analysis process was detailed in Chapter 7. The result of Chapter 7 was an updated list of guidelines on the content of DW/BI modules. The process of developing DW/BI was rooted in analysis of interview data in terms of CSH and SSM. This process and its findings were presented in Chapter 8.

3. Develop guidelines representative of the needs of all identified stakeholders, grounded in the literature review, universities' programmes, and from interview data analysis.

The final set of guidelines are provided in next section of this chapter.

9.3 Combined set of guidelines

This section presents the combined set of guidelines from this study. The main research objective was addressed by focusing on “what to teach in the DW/BI content” and “what process should be used to improve DW/BI education?” The guidelines focus on skills and knowledge that DW/BI modules content should address. Guidelines are categorised into four groups: prerequisites; DW/BI specific skills; advanced general IT skills; and processes to improve the DW/BI module.

9.3.1 First proposed guideline (Prerequisites)

Data analysis (presented Chapter 5, Section 5.6) pointed out that a DW/BI module should have prerequisites. Data analysis performed in Chapter 7 identified Knowledge on Mathematics and Statistics as additional prerequisites. ***The first proposed guideline is that: students enrolling for DW/BI module must have done modules that cover skills that are presented in Table 9-1 as prerequisite skills.*** The process for the development of this table is demonstrated in Chapter 5, Section 5.6.1.

Table 9-1 Advanced general IT skills prerequisite

Prerequisite general literature Skill	Source of skill	Description
Project management	Lit_1_PreGen_ProjectM	Project management of OLTP systems including project management of technical systems such as ETL.
Requirements collection (gathering)	Lit_2_PreGen_Requirements	Requirements collection and analysis for OLTP systems and end-user applications focussed on reporting.
Technical architecture	Lit_3_PreGen_TechArchData	Data management and technical architecture for OLTP systems based on ERD data models.
Database concepts	Lit_4_PreGen_DBFunc	Database concepts and functionality of relational databases.
Database programming	Lit_5_PreGen_DBProg	Database programming of relational databases
Database modelling (ERD modelling)	Lit_6_PreGen_ERD	Database modelling of relational data in ERD models.
Online transactional processing systems	Lit_7_PreGen_OLTP	Overall modelling and management of data in OLTP systems.
Data flow processes	Lit_8_PreGen_Processes	Data flow and other operational processes in OLTP systems.
Analytical skills	Lit_9_PreGen_Analytical	Analytical skills in terms of all components of OLTP systems including requirements analysis and technical analysis.
Programming skills	Lit_10_PreGen_ETL_Tools	Programming skills to do basic data manipulation.

Data structures	Lit_11_PreGen_DataStructures	Knowledge and manipulation skills of general data structures such as objects, link lists and arrays.
Database connections	Lit_12_PreGen_DBConnections	Creation of connections between different types of databases.
Database administration skills	Lit_13_PreGen_DBA	DBA skills for relational databases.
General programming	Lit_14_PreGen_Prog	General programming skills.
Reporting	Lit_15_PreGen_Reporting	General reporting skills.
Source systems	Lit_16_PreGen_Source	Knowledge of different source systems and platforms including ERP systems.
Business processes	Lit_17_PreGen_Business	Knowledge of typical business processes
Database reporting	Lit_18_PreGen_DBReporting	Reporting linked to database systems.
DW/BI Tools knowledge and skills	Lit_19_PreGen_ReportTools	Use of reporting tools.
System development life cycle (SDLC)	Lit_20_PreGen_SDLC	Good knowledge of SDLC of OLTP systems.
Interface programming	Lit_21_PreGen_Interface	Development of user interfaces for OLTP systems.
Microsoft Excel	Lit_22_PreGen_MSEExcel	Microsoft Excel knowledge including pivot tables.
Big Data and unstructured data	Lit_23_PReGen_BigData	Introductory knowledge of Big Data and unstructured data.
Data Mining	Lit_24_PreGen_DataMining	Introductory knowledge of Data Mining.
Online transaction processing vs DW/BI systems	Lit_25_PreGen_OLTP/BI	Knowledge of the comparison of OLTP and DW/BI systems.
Metadata	Lit_26_PreGen_Metadata	Knowledge of metadata.
Deployment	Lit_27_PreGen_Deployment	Knowledge of the deployment of new systems.
Testing	Lit_28_PreGen_Testing	Good testing knowledge.
User training	Lit_29_PreGen_UserTraining	Knowledge of user training and support.
Security (information)	Lit_30_PreGen_DataSecurity	Awareness of security risks and procedures.
Networking	Lit_31_PreGen_Networking	Knowledge and skills of network topologies and their implementation.

Hardware	Lit_32_PeGen_HardwareKnow	Knowledge of hardware systems such as distributed computing.
Mathematics	Interviews Theme 15	Basic Mathematics knowledge
Statistics	Interviews Theme 15	Statistics knowledge

9.3.2 Second proposed guideline (DW/BI specific skills)

The data analysis showed that there are specific core skills that a DW/BI module should teach. The second proposed guideline is that DW/BI module should cover the specific skills that are presented in Table 9-2. These skills are grounded in data analysis that was carried out in Chapter 6 (Section 6.6). These skills were identified through data analysis, and supported by guidelines from the literature review and universities' programmes. The empirical data analysis also identified other skills that were not identified from the literature review and universities' programmes.

The table was initially developed in Chapter 4, Section 4.4.12, and then enriched by universities' programme/module data in Chapter 5, Section 5.6.2. This combined table was then extended by items identified during analysis of the interview data (in Section 7.6 of Chapter 7).

Table 9-2 DW/BI specific skills

DW/BI specific skills	Source of skill	Description
BI components knowledge	Lit_1_BI_Components	BI components include data, technology, processes, and people.
DW architecture	Lit_2_BI_Achitecture	Knowledge and understanding of DW/BI architecture in terms of sources systems, presentation server and end-user applications.
DW lifecycle	Lit_3_BI_DWlifecycle	Knowledge and application of the DW/BI lifecycle in terms of business requirements, technology choices, data modelling, end-user applications, deployment, growth, and project management.
DW Methodologies.	Lit_4_BI_Methodologies	Knowledge and application of alternative DW Methodologies designed by authors such as Ralph Kimball, Bill Inmon, Linstedt and others.
Technical architecture	Lit_5_BI_TechAchitect	Knowledge of the technical architecture in terms of product selection and acquisition,

		infrastructure implications.
ETL tools	Lit_6_BI_ETL_Tools	Knowledge and application of vendor specific tools that enable better ETL processing as well as the underlying principles applied by these tools.
Dimensional modelling	Lit_7_BI_DimModeling	Knowledge and application of dimensional modelling in terms of introductory and advanced concepts.
Multidimensional programming language	Lit_8_BI_MDX	Knowledge and application of multidimensional dynamics expression (MDX) language compared to traditional database manipulation.
Analytical processing (OLAP)	Lit_9_BI_OLAP	Knowledge and application of online analytical processing (OLAP) cubes and their role in support of business decisions.
Mobile DW/BI	Lit_10_BI_MobileBI	Knowledge and application of Mobile DW/BI technology and their application.
Cross subjects area analysis	Int_1_CrossSubjAna	The module should include an element of creating of teaching students to organise reporting elements by business functions.
Data quality management	Int_2_DataQualMan	Data quality management is part of advance database management, but this aspect requires more explicit reference in the guideline.
Business models	Int_3_BusMod	The module should include an element of teaching students about different business models to demonstration that DW/BI can be applied almost any business that values data.
Draw a DW/BI story that explain why do DW/BI	Int_4_DrawBIStory	This is an instructional preference. The introduction to the module provided by lecturers does contain this but it is not presented as a narrative. Develop a visual narrative that explains why students should do and gain DW/BI knowledge.
Push internship programme's knowledge back into universities	Int_5_IntProg	This is scope extension to this study. This aspect requires more reflection.
Purpose of DW/BI	Int_6_BIPurpose	This is an instructional preference as well as module orientation issue.

9.3.3 Third proposed guideline (Advance general IT skills)

According to data analysed advanced general IT skills are a requirement for DW/BI students. These skills may be part of the DW/BI module, or covered by advanced modules on other IT topics concurrently done by DW/BI students. **The third**

proposed guideline is that DW/BI students should master advanced general IT skills. These are presented in Table 9-3. These skills were identified in Chapter 6 (Section 6.6) from universities' programmes/module documentation. The empirical data analysis did not yield additional skills that were not identified from the literature review and university programmes.

Table 9-3 Advanced general IT skills

Advanced general IT skills	Source of skill	Description
Project management	Lit_1_AdvGen_ProjectM	According to Kimball <i>et al.</i> (2008d:339) project management skills is essential skill to the entire DW lifecycle.
Requirements gathering	Lit_2_AdvGen_Requirements	According to Kimball <i>et al.</i> (2008d:64) business analysis skills are key in requirements collection process. This should include skills to develop a requirements bus matrix.
Technical architecture	Lit_3_AdvGen_TechArchData	According to Kimball <i>et al.</i> (2008d:110) technical architecture skills is key for BI/DW technical architecture.
Database programming	Lit_4_AdvGen_DBProg	According to Kimball <i>et al.</i> (2008d:327) DB programming skills are key to designing the physical database and planning for performance.
Data modelling	Lit_5_AdvGen_OLTP	According to Kimball <i>et al.</i> (2008d:370) data modelling skills are essential for ETL and DW/BI application development in a DW/BI solution.
Analytical skills	Lit_6_AdvGen_Analytical	According to Kimball <i>et al.</i> (2008d:348) analytical skills are essential for requirements collection process.
ETL Tools	Lit_7_AdvGen_ETL_Tools	According to Kimball <i>et al.</i> (2008d:110) knowledge of DW/BI tools is essential to especially.
Database administration	Lit_8_AdvGen_DBA	According to Kimball <i>et al.</i> (2008d:327) DBA skills are key to designing the physical database and planning for performance.
Application programming	Lit_9_AdvGen_Interface	According to Kimball <i>et al.</i> (2008d:506) application programming is important for user front development and configuration.
Big data	Lit_10_AdvGen_BigData	The historical amounts of data accumulated thus far by many organisations and its current growth rate led to the coining of "Big data" term (Tankard & Pathways, 2012:5). This growth has given rise in volume (amount of

		data), velocity (speed of data) and variety (range of data).
Data Mining	Lit_11_AdvGen_DataMining	According to Kimball <i>et al.</i> (2008d:475) data mining is one of DW/BI application used to analyse data.
Metadata	Lit_12_AdvGen_Metadata	According to Kimball <i>et al.</i> (2008d:64) metadata knowledge and handling is key in requirements collection process and the entire DW/BI system.
Testing	Lit_13_AdvGen_Testing	According to Kimball <i>et al.</i> (2008d:364) system testing and quality assurance skills are essential in the deployment phase of a DW/BI solution.
User Training	Lit_14_AdvGen_UserTraining	According to Kimball <i>et al.</i> (2008d:15) training skills are essential in the deployment phase of a DW/BI solution.
Data Security	Lit_15_AdvGen_DataSecurity	According to Kimball <i>et al.</i> (2008d:327) information security skills are key to designing the physical database and planning for performance.
Database reporting	Lit_16_AdvGen_DBReporting	Advance reporting skills are required to enable the development of flexible reports required by DW/BI applications.
Online transactional processing and DW/BI systems	Lit_17_OLTP_BI	Advance understanding of differences between OLTP and DW/BI systems.

The guideline concerning the process to improve a DW/BI module is presented in the next section.

9.3.4 *Fourth proposed guideline (Processes to improve the DW/BI module)*

The first two guidelines concern the content of a DW/BI module and was mostly developed from DW/BI literature and DW/BI programmes/modules at universities; they were substantiated by interview data. ***The fourth guideline is to include the involved and the affected parties and their expertise in the process as identified by the CSH analysis – presented in Table 9-4. A process similar to the SSM activity diagram (Figure 9-1) is proposed.***

This guideline focuses on two aspects: the CSH analysis for improving DW/BI; and a SSM activity diagram.

In Chapter 8 the focus was on processes to improve the DW/BI module. CSH boundary judgements were used to also help identify the input elements of the improvement process. The following elements, grounded in the CSH findings (refer to Chapter 8, Section 8.3) supported the guideline to improve DW/BI modules.

Table 9-4 CSH analysis of problem situation

Element	Description
Beneficiaries	DW/BI module should benefit: business industry students universities
BI purpose	Teach the value of an enveloped solution that pushes business forward. These include: Business component DW/BI technology DW/BI principle Skills
Success measures	Students should be able to develop complete DW/BI solutions The impact students have in their various employment areas (industry).
Decision powers	Stakeholders that should possess the decision powers in DW/BI modules are: Business Students universities industry
Decision makers powers	The decision makers powers should: Business must control data and requirements. Universities should control course content. A body that recognises what DW/BI professional should be formed. Industry should determine DW/BI standards. Students have the decision powers to enrol for DW/BI module and to give a module review report back.
Conditions outside the control of universities	The objective to improve DW/BI module should consider the following constraints that universities can't control: DW/BI technology and trends (hardware and software licensing costs) Real life problems input by industry
Involved stakeholders	The following stakeholders should be involved in the development of a DW/BI module: Universities Industry Business Past students
Expertise to involve	The following expertise should be involved in the development of a DW/BI module: Consulting Technical business

	hybrid (teaching and technical)
Guarantor	The guarantee that DW/BI module objectives will be achieved should be: Students should do the practical work projects. Students employability A DW/BI body of knowledge comprising of all relevant stakeholders (the involved and affected) should be formed and included.
Affected stakeholders	The development of a DW/BI module should consider the following stakeholders that should be affected: Business Students Industry BI professionals
Emancipation	Specific concerns (how or what will emancipate those that are affected by DW/BI module) of the affected stakeholders should be addressed: Business should define business requirements aspect that goes into DW/BI module. Industry should define the technical requirements aspects that should got into a DW/BI module. Universities should control teaching and academic aspect. Students should have input into course content
World view	A DW/BI module should reflect the views of: universities industry past students business

Figure 9-1 below presents a SSM activity diagram that can guide action to improve a DW/BI module. This diagram is discussed in Section 8.9.1.1.4.

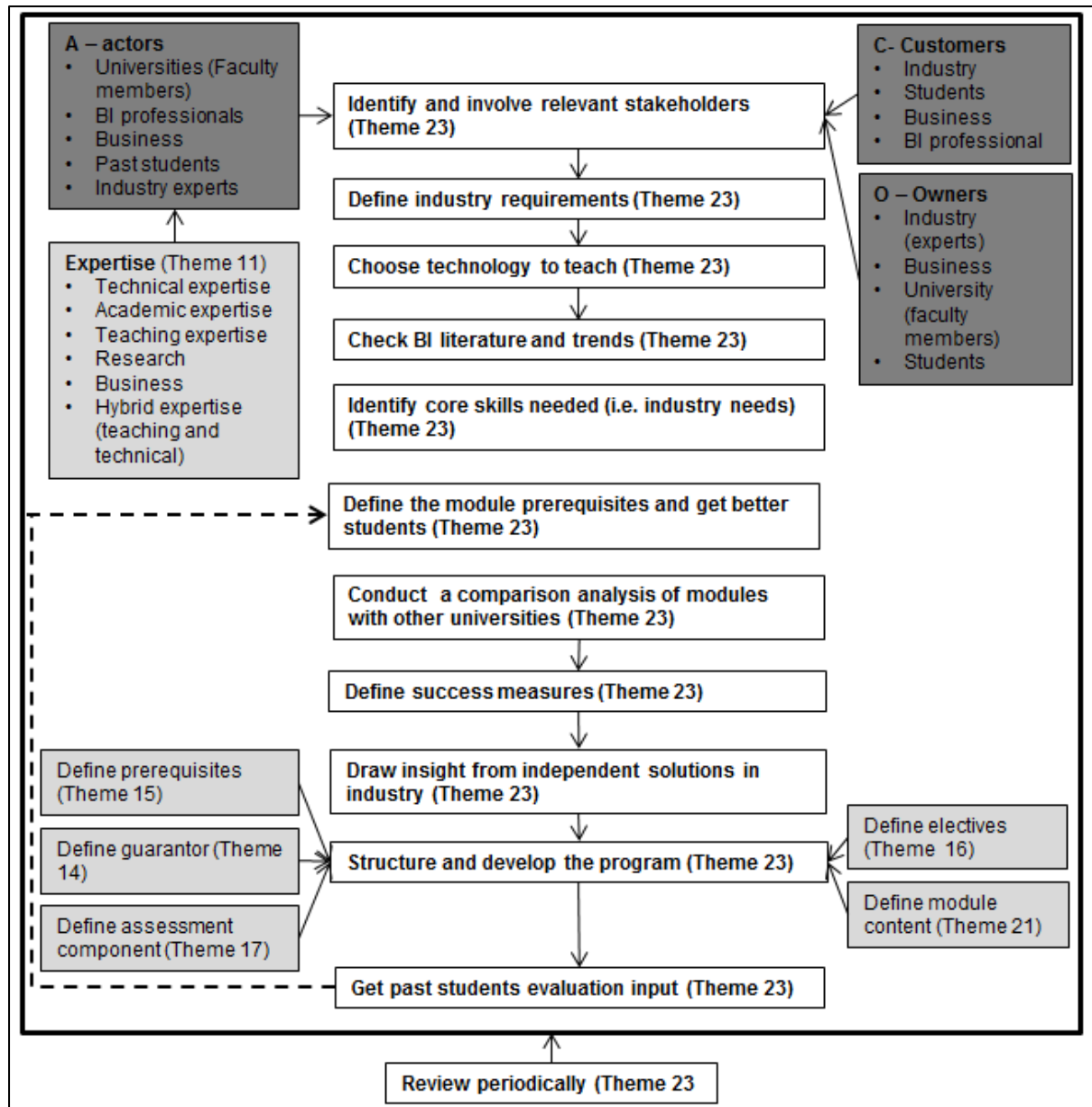


Figure 9-1: Activity model for improving a DW/BI module

Recommendations for further study are presented in the next section.

9.4 Recommendations for further study

This study concerned the development of guidelines for DW/BI modules in IT programmes at universities, based on an understanding of the needs of various stakeholders, using CSH and SSM.

From analysis of the interview data the practical application of DW/BI concepts was often highlighted by participants. Further research may focus on the specific design of the practical components in support of the guidelines proposed by this study.

No participant made reference to data vault modelling as alternative to dimensional modelling. The adoption of this method in South Africa is an interesting research topic.

Another possible research project is the creation of a body of concerned parties for the DW/BI industry in South Africa. Such a body may develop standards and work with the DW/BI industry and higher education sectors to improve the collaboration in terms of programme and internship content.

9.5 Assumptions and limitations of the study

A major assumption of this study was that all participants would answer the survey questions truthfully. The researcher could not secure face-to-face interviews with all the participants; as a result the following were identified as limitations of the study:

1. For participants who responded in writing it was assumed that they would understand the questions as intended and would answer all questions.
2. Written responses did not grant the researcher the opportunity to follow up for clarity where it was needed. However, attempts were made to reach the participants telephonically.

Some participants have enrolled for a DW/BI module during their study, but they are not currently actively working in DW/BI industry. As a result, they could not answer all questions in full.

It would also be beneficial to consider adding more participants so that a broader range of answers can be considered.

The application of system thinking methodologies such as CSH and SSM requires knowledge of these concepts. The participants did not have this knowledge and

therefore could not develop the diagrams themselves. The researcher developed the SSM component diagrams presented in Chapter 8; he then sent it to the participants for verification. It would have been better if the participants were able to construct these diagrams themselves.

The extent (scope and depth) to which the skill sets are covered by the universities is not clear from their general curriculums.

9.6 Summary

This chapter concludes the research study to develop guidelines for DW/BI modules in IT programmes at universities, based on an understanding of the needs of various stakeholders, using CSH and SSM. This chapter demonstrates how the objectives stated in Chapter 1 were achieved.

This chapter presents the four guidelines concerning the content and development process for a DW/BI module. The guidelines are:

1. The first proposed guideline is that: students enrolling for a DW/BI module must have done modules that cover skills that are presented in Table 9-1.
2. The second proposed guideline is that: a DW/BI module should cover the specific skills that are presented in Table 9-2.
3. The third proposed guideline is that: DW/BI students should master advance general IT skills presented in Table 9-3.
4. The fourth guideline is to include the involved and the affected parties and their expertise in the process, as identified by the CSH analysis and presented in Table 9-4. A process similar to the SSM activity diagram in Figure 9-1 is proposed.

The chapter concludes with a suggestion for future research and a reflection on the limitations of the study.

References

ACKOFF, R. 1971. Towards a System of Systems Concepts. *Management Science*, 12, 661-662.

AL-QIREM, R. 2012a. Enhancing Business Intelligence Tools: A systems Thinking Approach. *IJCSNS International Journal of Computer Science and Network Security*, 93-95.

AL-QIREM, R. 2012b. Enhancing business intelligence tools: a systems thinking approach. *International journal of computer science and network security*, 12, 93-95.

ALLIANCE, U. 2015. *Essential skill sets for careers in business intelligence* [Online]. Available: http://www.villanovau.com/resources/bi/essential-business-intelligence-skills/#.VW9EQ5PD_LI [Accessed 3 November 2015].

ALTRICHTER, H., KEMMIS, S., MCTAGGART, R. & ZUBER, O. 2002. The concepts of action research. *Emerald insight*, 9, 125- 131.

ANUPINDI, N. & COADY, G. 2005. *Inmon vs. Kimball - an analysis* [Online]. Available: <http://www.nagesh.com/publications/technology/173-inmon-vs-kimball-an-analysis.html> [Accessed April 15 2014].

ARIYACHANDRA, T. & WATSON, H. 2010. Key organizational factors in data warehouse architecture selection. *Decision support systems*, 49, 200– 212.

ATKINSON, P. 1998. *Earthnography and participation observation*, London, Sage.

BALASUBRAMANYAM, M. 2012. *ETL Developers: You must master these skills* [Online]. Available: <http://www.softwarespecialists.com/etl-developers-must-master-skills/> [Accessed 28 June 2015].

BALLARAT. 2010a. *What is a policy, procedure and guideline?* [Online]. Available: http://www.ballarat.edu.au/__data/assets/pdf_file/0019/82162/What_is_a_Policy_Procedure_Guideline.pdf [Accessed October 6 2013].

BALLARAT. 2010b. *What is a policy, procedure and guideline?* [Online]. Available: http://www.ballarat.edu.au/__data/assets/pdf_file/0019/82162/What_is_a_Policy_Procedure_Guideline.pdf [Accessed 6 October 2013].

BALLARD, C., HERREMAN, D., SCHAU, D., BELL, R., KIM, E. & VALENCIC, A. 1998. *Data modeling techniques for data warehousing*, California, IBM Corporation.

BARLOW, R. 2013. Making clinical data analytics count. *Health management technology*, 34, 8- 11.

BARTLETT, G. 2001. Systemic thinking a simple thinking technique for gaining systemic focus. *The international conference on thinking breakthroughs 2001*. Probsolv International.

BASKERVILLE, R. & MYERS, M. 2004. Special issue on action research in information systems: making is research relevant to practice-foreword. *MIS Quarterly*, 28, 329-336.

BAWDEN, R. 1991 Systems thinking practice. *journal of dairy science* 74, 2362-2373.

BEEKMAN, G. 2015. *Tomorrow's Technology and you*, Upper Saddle River: Pearson, Prentice Hall.

BERGAMASCHI, S., GUERRA, F., ORSINI, M., SARTORI, C. & VINCINI, M. 2011. A semantic approach to ETL technologies. *Data & Knowledge Engineering*, 70, 717–731.

BERTALANFY, L. V. 1976. *General systems theory: Foundations, development, applications*, New York, George Brazillier.

BITSCH, V. 2005. Qualitative research: A Grounded theory example and evaluation criteria. *Journal of Agribusiness*, 23, 75- 91.

BJERKE, O. 2008. Soft systems methodology in action: A case study at a purchasing department. Göteborg, Sweden University of Göteborg.

BLACKSTONE, A. *Inductive or Deductive? Two Different Approaches* [Online]. Available: <http://2012books.lardbucket.org/books/sociological-inquiry-principles-qualitative-and-quantitative-methods/s05-03-inductive-or-deductive-two-dif.html> [Accessed August 25 2014].

BLUMBERG, R. & ATRE, S. 2003. The problem with unstructured data. *DM Review*, 42-46.

BODY, M., MIQUEL, M., BERDARD, Y. & TCHOUNIKEN, A. 2013. Handling evolution of multi-dimensional structures. *Institute of electrical and electronics engineers (IEEE)*, 581 - 591.

BOELL, S. & CECEZ-KECMANOVIC, D. 2010. Literature reviews and the hermeneutic circle. *Australian Academic and Research Libraries*, 41, 129-144.

BOON, K. & TAN, H. 2006. Clinical practice guidelines: a critical review. *Emerald Group Publishing Limited*, 195-220.

BRADBOURNE, M. & ROUSE, C. 2011. *Educating the next generation of business intelligence professionals* [Online]. Available: <http://www.b-eye-network.com/view/15246> [Accessed May 7 2014].

BRESLIN, M. 2004. Data warehousing battle of the giants: comparing the Basics of the Kimball and Inmon models. *Business intelligence journal*, 6-20.

BUCHER, T., GERICKE, A. & SIGG, S. 2009. Process-centric business intelligence. *Business process management*, 15, 408-429.

CAGDAS, V. & STUBKJAER, E. 2011. Design research for cadastral systems. *Computers environment and urban systems*, 37, 77–87.

CALVANESE, D., GIACOMO, G. D., LENZERINI, M., NARDI, D. & ROSATI, R. 1999. A principled approach to data integration and reconciliation in Data warehousing. *CEUR-WS publications*, 19, 16:1 - 16:11.

CARCARY, M. 2009. The Research Audit Trial – Enhancing Trustworthiness in Qualitative Inquiry. *Electronic Journal of Business Research Methods*, 7, 11- 24.

CARRUTHERS, G. 1999. A critical review of research paradigms and strategies and their role in chiropractic research. *The British journal of Chiropractic*, 3, 54- 82.

CASSIE, C. 1997. Marketing decision support systems. *Industrial Management & Data Systems*, 97, 293–296.

CHAN, S. 2001a. *Complex adaptive system, esd.83 research seminar in engineering systems* [Online]. Available:
<http://web.mit.edu/esd.83/www/notebook/Complex%20Adaptive%20Systems.pdf>
[Accessed 15 February 2015].

CHAN, S. 2001b. Complex adaptive systems. *Research seminar in engineering systems*, 1-9.

CHAN, S. Complex adaptive systems. Research seminar in engineering systems, 2001c. 1- 9.

CHECKLAND, P. 1997a. International Encyclopedia of Business & Management. *Thompson Business Press*, 667-673.

CHECKLAND, P. 1997b. Systems. *International Encyclopedia of Business & Management*, 667-673.

CHECKLAND, P. 2000. Soft Systems Methodology: A Thirty Year Retrospectivea. *Systems Research and Behavioral Science* 17, S11–S58.

CHECKLAND, P. 2011. Systems movement autobiographical retrospectives. *International journal of general systems*, 40, 487–512.

CHECKLAND, P. 2013. *System, system, system, soft systems methodology* [Online]. Available: <http://pdinspire.wordpress.com/tag/peter-checkland/> [Accessed September 10 2013].

CHECKLAND, P. & POULTER, J. 2006. *Learning for action: A short definitive account of the soft methodology and its use for practictioners, teachers and students*, England, Wiley 2006.

CHEN, H., CHIANG, R. & VEDA, S. 2012. Business intelligence and analytics: From big data to big data impact. *MIS Quaterly*, 36, 1165-1188.

CHEN, Y., DONG, G., HAN, J., PEI, J., WAH, B. & WANG, J. 2002. *Online analytical processing stream data: Is it feasible?* [Online]. Available: <http://www.bell-labs.com/user/minos/DMKD02/Papers/chen.pdf> [Accessed October 27 2013].

CHENA, C., CHENE, K., HSUC, C.-Y. & LI, Y.-C. 2011. Developing guideline-based decision support systems using protégé and jess. *Elsevier Ireland*, 288–294.

CHERMACK, T. & LYNHAM, S. 2004. Scenario planning in critical science research. *Futures Research Quarterly Summer*, 44-60.

CHIANG, R., GOES, P. & STOHR, E. 2012. Business intelligence and analytics education, and program Development: A unique opportunity for the information systems discipline. *ACM transactions on management information systems*, 3.

CHIODO, F., MENDITTO, A., FOUTOMATA OUENE-KEITA & CAROLI, S. 1994. Management of information: storage and retrieval of data. *Inn.Ist super sanita*, 30, 433- 437.

CHOUNTAS, P., VASILAKIS, C., EL-DARZI, E., PETROUNIAS, I. & TSENG, A. 2003. Data warehouses-TOLAP-decision Making. *Systems, man and cybernetics, 2003. Institute of electrical and electronics engineers (ieee) international conference on*, 4, 3876 - 3880.

CLARK, D. 2004a. *Understanding and Performance* [Online]. Available: <http://www.nwlink.com/~donclark/performance/understanding.html> [Accessed October 8 2013].

CLARK, D. 2004b. *Understanding and performance* [Online]. Available: <http://www.nwlink.com/~donclark/performance/understanding.html> [Accessed May 5 2016].

COBBIN, M. 2012. *BI dashboard – the good, the bad, and the ugly* [Online]. Available: <http://michaelcobbins.wordpress.com/category/business-analytics/> [Accessed May 8 2014].

COLLIER, K. 2011. *Introducing agile analytics: A value-driven approach to business intelligence and data warehousing* [Online]. Available: <http://www.informit.com/articles/article.aspx?p=1743274&seqNum=3> [Accessed 24 June 2015].

COWIE, J. & BURSTEIN, F. 2007. Quality of data model for supporting mobile decision making. *Decision Support Systems*, 43.

CRINSON, I. & LEONTOWITSCH, M. 2011. *Qualitative methods* [Online]. Available: <http://www.healthknowledge.org.uk/public-health-textbook/research-methods/1d-qualitative-methods> [Accessed 19 April 2016].

CROSS, N. 2001. Designerly ways of knowing: design discipline versus design science. *Design Issues*, 17, 49–55.

CUMBLEY, R. & CHURCH, P. 2013. Is “Big Data” creepy? *Computer security review*, 19, 601- 609.

DALRYMPLE, P. 2011. Data, information, knowledge:the emerging field of health informatics. *Bulletin of the american society for information science and technology*, 37, 41- 44.

DARKE, P., SHANKS, G. & BROADBENT, M. 1998a. Successfully completing case study research: combining rigour, relevance and pragmatism. *Info Systems* 8, 273-289.

DARKE, P., SHANKS, G. & BROADBENT, M. 1998b. Successfully completing case study research: combining rigour, relevance and pragmatism. *Info Systems*, 8, 273-289.

DATTA, A. & THOMAS, H. 1999. The cube data model: a conceptual model and algebra for on-line analytical processing in data warehouses. *Decision Support Systems*, 27, 289–301.

DELEN, D. & DEMIRKAN, H. 2013. Data, information and analytics as services. *Decision support systems*, 55, 359–363.

DETTMER, W. 2006. An introduction to the systems approach. *Goal systems international*, 1- 4.

DEVLIN, A. & MURPHY, P. 1988. An architecture for a business and information system. *IBM systems journal*, 27, 60 - 80.

DOVENPORT, T. 2006. Competing on analytics. *Harvard business review*, 1- 10.

EL-SAPPAGH, S., HENDAWI, A. & BASTAWISSY, A. E. 2011. A proposed model for data warehouse ETL processes. *Journal of king saud university – computer and information sciences*, 23, 91–104.

EL-SAPPAGHA, S., HENDAWIB, A. A. & BASTAWISSYB, A. H. E. 2011. A proposed model for data warehouse ETL processes. *Journal of king saud university - computer and information sciences*, 23, 91–104.

ELLIOT, R. & JANKELL-ELLIOT, N. 2003. Using ethnography in strategic consumer research. *Qualitative market research: An international journal*, 6, 215-223.

ELO, S., KÄÄRIÄINEN, M., KANSTE, O., PÖLKKI, T., UTRIAINEN, K. & KYNGÄS, H. 2014. Qualitative Content Analysis: A Focus on Trustworthiness. *SAGE Journals*, 1- 6.

ELO, S. & KYNGAS, H. 2007. The qualitative content analysis process. *Jan research methodology*, 107-115.

FACIONE, P. 1990. *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*, Millbrae, CA, The California Academic Press.

FACIONE, P. 2000. The disposition toward critical thinking: Its character, measurement, and relation to critical thinking skill. *Informal Logic*, 20, 61–84.

FEW, S. 2005. *Effective communicating numbers: selecting the best means and manner of display* [Online]. Available: http://www.perceptualedge.com/articles/Whitepapers/Communicating_Numbers.pdf [Accessed October 8 2013].

FIRESTONE, J. 1998. *Dimensional modeling and e-r modeling in the data warehouse* [Online]. Available: <http://www.dkms.com/papers/dmerdw.pdf> [Accessed April 14 2014].

FLETCHER, G. & WYSS, C. 2006. Data mapping as search. *Springer*, 95- 111.

FUCHS, C. & HOFKIRCHNER, W. 2009. Autopoiesis and critical social systems theory. In *Autopoiesis in organization theory and practice*. *Emerald*, 111 - 129.

GEORGE, S. 2012. *Inmon vs. Kimball: Which approach is suitable for your data warehouse?* [Online]. Available: <http://searchbusinessintelligence.techtarget.in/tip/Inmon-vs-Kimball-Which-approach-is-suitable-for-your-data-warehouse> [Accessed May 5 2016].

GILE, K. 2006. Grading BI reporting and snalysis solutions. *The Forrester Wave*, 1 - 14.

GILES, L. 1910. *The art of war by Sun Tzu*, The puppet press.

GLASER, B. & STRAUSS, A. 1967. *The discovery of grounded theory: Strategies for qualitative research*, New York, Aldine de Gruyter.

GREGOR, S. & HEVNER, A. 2013. Positioning and presenting Design science for maximum impact. *MIS Quaterly*, 37, 337-355.

HAJMOOSAEI, A., KASHFI, M. & KAILASAM, P. 2011. Comparison plan for data warehouse system architectures. *Institute of electrical and electronics engineers (IEEE)*, 290- 293.

HAMMERSLEY, M. 2012. *Methodological paradigms in educational research* [Online]. Available: http://www.google.co.za/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=2&cad=rja&sqi=2&ved=0CDwQFjAB&url=http%3A%2F%2Fwww.bera.ac.uk%2Fsystem%2Ffiles%2FMethodological%2520Paradigms.pdf&ei=9XFFUqWuN6aJ4AS0m4DACw&usg=AFQjCNGS9Rz1Khz0RT7dPUE7Wa7niTgivw&sig2=n_ [Accessed September 27 2013].

HANLIN, Q., XIANZHEN, J. & XIANRONG, Z. 2012. Research on Extract, Transform and Load(ETL) in Land and Resources Star Schema Data Warehouse *IEEE computer society*, 120- 123.

HARPER, J. 2013. *The future of business intelligence* [Online]. Available: <http://www.dataversity.net/the-future-of-bi/> [Accessed March 1 2013].

HARRIS, L. & BROWN, G. 2010. Practical Assessment, Reseach Evaluation. *Electronic Journal*, 15, 1-19.

HART, D. & CACERES, P. 2012. Using systems heuristics to guide “second order” boundary critique in a systemic intervention: Evaluating the environmental impact of mining in southern peru. *UK academy for information systems conference proceedings*. AIS Electronic Library.

HARVEY, L. 1990. *Critical social research*, London, Unwin Hyman.

HEDGEBETH, D. 2007a. Data-driven decision making for the enterprise: an overview of business intelligence applications. *The journal of information and knowledge management systems*, 37, 414-420.

HEDGEBETH, D. 2007b. Data-driven decision making for the enterprise: an overview of business intelligence applications. *The journal of information and knowledge management systems*, 37, 414-420.

HEVNER, A., MARCH, S., PARK, J. & RAM, S. 2004. Design Science in Information Systems Research. *MIS Quarterly*, 28, 75-105.

HEY, J. 2004. *The data, information, knowledge, wisdom chain: The metaphorical link* [Online]. Available: <http://faculty.unlv.edu/anjala/BGES/DIKW.pdf> [Accessed October 8 2013].

HOLSTI, O. 1968. *The handbook of social psychology*, New Delhi, Amerind Publishing Co. .

HOSTMANN, B., RAYNER, N. & HERSCHEL, G. 2009. intelligence and planning in midsize companies. *Gardner*, 1-20.

HSIEH, H.-F. & SHANNON, S. 2005. Three Approaches to Qualitative Content Analysis. *SAGE Journals*, 15, 1277-1288.

HUA, J.-S., HUANG, S.-M. & YEN, D. 2012. Architectural support for business intelligence: a push-pull mechanism. *Emerald*, 36, 52-71.

IMHOFF, C. & WHITE, C. 2011. Self-service business intelligence. The data warehousing institute.

INMON, B. 2007. *About Bill* [Online]. Available: <http://inmoncif.com/about/> [Accessed April 4 2014].

INMON, B. 2008. *ETL: Extract, transform and load for data warehousing* [Online]. Available: <http://www.b-eye-network.com/view/6680> [Accessed 2 November 2015].

INMON, B. W. 2000. *building the data warehouse: Getting started* [Online]. Available: <http://inmoncif.com/inmoncif-old/www/library/whiteprs/ttbuild.pdf> [Accessed April 28 2014].

INMON, W. 2002. *Building The Data Warehouse*, Canada, Robert Ipsen.

INMON, W. 2005. *Building the data warehouse*, Indianapolis, Wiley Publishing.

ITNEXT. 2013a. *Top 4 business intelligence trends for 2014* [Online]. Available: <http://www.itnext.in/itnext/articles/18268/business-intelligence-trends-2014> [Accessed May 5 2016].

ITNEXT. 2013b. *Top 4 business intelligence trends for 2014* [Online]. Available: <http://www.itnext.in/itnext/articles/18268/business-intelligence-trends-2014> [Accessed March 1 2014].

JACKSON, A., HETTINGA, D., MEADB, J. & MERCER, C. 2009. Using consensus methods in developing clinical guidelines for exercise in managing persistent low back pain. *Elsevier Ireland*, 302–313.

JACKSON, M. 1985. Social systems theory and practice: the need for a critical approach. *International Journal for General Systems*, 10,135–151.

JACKSON, M. 2001. Critical systems thinking and practice. *European Journal of Operational Research* 128 233 - 244.

JAKOBSEN, T. 2012. *Theory of science – what is positivism?* [Online]. Available: <http://www.popularsocialscience.com/2013/02/15/theory-of-science-what-is-positivism/> [Accessed August 13 2014].

JOHNSON, C. 2011. *Ethical decision making and behavior* [Online]. Available: http://www.sagepub.com/upm-data/39590_Chapter7.pdf [Accessed September 11 2013].

JONES, M., KRIFLIK, G. & ZANKO, M. Grounded theory: a theoretical and practical application in the Australian film industry. Qualitative research convention, 2005 Malaysia. Qualitative research of Malaysia.

JOUBISH, M., KHURAAM, M., AHMED, A., FATIMA, S. & HEIDER, K. 2011. Paradigms and characteristics of a good qualitative research. *World applied science journal*, 12, 2082- 2087.

KABOUB, F. 2008a. Positivist Paradigm. *SAGE Publications*, 2, 786-787.

KABOUB, F. 2008b. Positivist Paradigm. *SAGE Publications*, 2, 786-787.

KAMOOR, B. & SHERIF, J. 2012. Human resources in an enriched environment of business intelligence. *Emerald*, 41, 1625-1637.

KANICLIDES, T. & KIMBLE, C. 1994. *Executive information systems:A framework for their development and use* [Online]. Available: <http://www.cs.york.ac.uk/ftpdireports/94/YCS/247/YCS-94-247.pdf>.

KAPOOR, B. & SHERIF, J. 2012. Human resources in an enriched environment of business intelligence. *Emerald*, 41, 1625-1637.

KARAGIANNIS, A., PANOS, V. & ALKIS, S. 2013. Scheduling strategies for efficient ETL execution. *Information Systems*, 38, 927–945.

KEITH, M., DEMIRKAN, H. & GOUL, M. 2013. Service-oriented methodology For systems development. *Journal of management information systems*, 30, 227–259.

KIMBALL, R. 2013. *Kimball Dimensional Modeling Techniques* [Online]. Available: <http://www.kimballgroup.com/wp-content/uploads/2013/08/2013.09-Kimball-Dimensional-Modeling-Techniques1.pdf> [Accessed April 15 2014].

KIMBALL, R. 2004. *The 38 subsystems of ETL* [Online]. Available: <http://www.kimballgroup.com/2004/12/the-38-subsystems-of-etl/> [Accessed 5 November 2015].

KIMBALL, R. 2008. *Fact tables* [Online]. Available: <http://www.kimballgroup.com/2008/11/05/fact-tables/> [Accessed May 2016].

KIMBALL, R. 2014. <http://www.kimballgroup.com/> [Online]. Available: <http://www.kimballgroup.com/> [Accessed April 9 2014].

KIMBALL, R. 2015. *ETL Architecture in depth* [Online]. Available: <http://www.kimballgroup.com/wp-content/uploads/2015/05/Kimball-University-ETL-Architecture-in-Depth-Course-Description-2015.05.pdf> [Accessed 1 June 2015].

KIMBALL, R., BECKER, B., THORNTHWAITE, W., MUNDY, J. & ROSS, M. 2008a. *The Data Warehouse Lifecycle Toolkit*, Canada., John Wiley & Sons.

KIMBALL, R. & MARGY, R. 2002. *The data warehouse toolkit: The complete guide to dimensional modeling*, Canada, John Wiley and Sons.

KIMBALL, R., MARGY, R., THORNTHWAITE, W., MUNDY, J. & BECKER, B. 2008b. *The Data Warehouse Lifecycle Toolkit*, Canada, John Wiley & Sons.

KIMBALL, R., ROSS, M., THORNTHWAITE, W., MUNDY, J. & BECKER, B. 2008c. *The Data Warehouse Lifecycle Toolkit*, Canada, John Wiley & Sons.

KIMBALL, R., ROSS, M., THORNTHWAITE, W., MUNDY, J. & BECKER, B. 2008d. *The Data Warehouse Lifecycle Toolkit*, Canada, John Wiley & Sons.

KLEIN, H. & MYERS, D. 2011. A set of principles for conducting critical research in Information Systems. *MIS Quarterly*, 35 17-36.

KOMLA, M. & OFFEI-ANSAH, C. 2011a. Linking tertiary institutions to industries: Evidence from the Vocational and Technical Education Department of the University of Cape Coast. *International Journal of Vocational and Technical Education*, 2, 53-60.

KOMLA, M. & OFFEI-ANSAH, C. 2011b. Linking tertiary institutions to industries: Evidence from the vocational and technical education department of the university of cape coast. *International journal of vocational and technical education*, 2, 53-60.

KRAUSS, S. 2005a. Research paradigms and meaning making. *A Primer. The qualitative report*, 758-770.

KRAUSS, S. 2005b. Research Paradigms and Meaning Making: A Primer. *The Qualitative Report*, 758-770.

KUHN, T. 1970. *The Structure of Scientific Revolutions*, Chicago, The university of Chicago press.

KUHN, T. 1996. *The structure of scientific revolutions*, Chicago, University of Chicago Press.

LEGG, S. & HUTTER, M. 2007. Universal intelligence: a definition of machine intelligence. *Minds and Machines*, 17, 391-444.

LEONARDO, Z. 2004a. Critical social theory and transformative knowledge: The functions of criticism in quality education. *Educational Researcher*, 33, 11–18.

LEONARDO, Z. 2004b. Critical Social Theory and Transformative Knowledge: The Functions of Criticism in Quality Education. . *SAGE Journals*, 33, 11-18.

LEVENE, M. & LOIZOU, G. 2003. Why is the snowflake schema a good data warehouse design? *Information Systems*, 28, 225–240.

LINCOLN, S. & GUBA, G. 1985. *Naturalistic inquiry*, Beverly Hills, CA.

LINGARD, L., ALBERT, M. & LEVINSON, W. 2008. Qualitative research: Grounded theory, mixed methods, and action research. *BMJ*, 337, 459 - 461.

LIVESEY, C. 2006. *Sociology* [Online]. Available: <http://www.sociology.org.uk/revgrm5.pdf> [Accessed September 1 2013].

LOUKOPOULOS, T., KALNIS, P., AHMAD, I. & PAPADIAS, D. 2001. Active caching of on-line-analytical-processing queries in www proxies. *Institute of electrical and electronics engineers (IEEE)*, 419 - 426.

LUNT, B., EKSTROM, J., GORKA, S., HISLOP, G., KAMALI, R., LAWSON, E., LEBLANC, R., MILLER, J. & REICHGELT, H. 2008. Curriculum Guidelines for Undergraduate Degree Programs in Information Technology. *Association for Computing Machinery (ACM), IEEE Computer Society*.

MA, C., CHOU, D. & YEN, D. 2000. Data warehousing technology, assessment and management. *Industrial management & data systems*, 100, 125- 134.

MAQBOOLRAO, N., BUTT, M., ZAMAN, M. & BAKSHI, W. 2013. Distributed data warehouse architecture: An efficient priority allocation mechanism for query formulation. *International journal of engineering and innovative technology (IJEIT)*, 2, 157- 159.

MARJANOVIC, O. 2011. Addressing the ICT-related Challenges of Business Intelligence Education. *Int. Conf. on Information Technology Interfaces*, 151-156.

MARSHALL, L. & HARPE, R. D.-L. 2009. Decision making in the context of business intelligence and data quality. *South african journal of information management*, 11, 1- 15.

MARTIAU, H. 1896. *The positive philosophy of Auguste Comte*, London, George Bell & Sons.

MATTHIAS, J. & YANNIS, V. Data warehouse quality: a Review of the foundations of data warehouse quality project. 1997 Cambridge. Foundations of data warehouse quality, 299-313.

MCFADDEN, F. Data warehouse for EIS: Some issues and impacts. Proceedings of the twenty-ninth hawaii international conference on, 1996 Wailea. Institute of electrical and electronics engineers, 120 - 129.

MCPECK, J. 1990. Critical thinking and subject specificity: A reply to Ennis. *Educational Researcher*, 19, 10–12.

MEYER, C. 2001. A Case in Case Study Methodology. *SAGE publications*, 13, 329-352.

MICROSOFT. 2014. *Subscriptions and delivery (reporting services)* [Online]. Available: <http://technet.microsoft.com/en-us/library/ms159762.aspx> [Accessed January 10 2014].

MINGERS, J. 2001. Combining IS research methods: towards a pluralist methodology. *Information systems research*, 12, 240–259.

MIRJI, A. 2011. *Tips for developing a BI roadmap* [Online]. Available: <http://www.saama.com/tips-for-developing-a-bi-roadmap/> [Accessed September 17 2013].

- MOTTIER, V. 2005. The Interpretive Turn: History, Memory, and Storage in Qualitative Research Forum. *Qualitative social research*, 6, 1-9.
- MRDALJ, S. 2007. Teaching an applied business intelligence course. *Issues in Information Systems*, VIII, 134-338.
- MULCAHY, R. 2007. *Business intelligence definition and solutions* [Online]. CIO. Available: <http://www.cio.com/article/2439504/business-intelligence/business-intelligence-definition-and-solutions.html> [Accessed April 20 2016].
- MUNAWAR, N. S. & ROLIANA, I. 2011. Towards Data Quality into the Data Warehouse Development. *Institute of electrical and electronics engineers (IEEE) conference publications*, 1199 - 1206.
- MYERS, M. & KLEIN, H. 1999. A Set of principles for conducting and evaluating interpretive field studies in information systems. *MIS Quarterly*, 23 Issue 1, 67–94.
- NADEEM, M. & JAFFRI, A. H. 2004. Application of Business Intelligence In Banks (Pakistan). *In The Computing Research Repository (CoRR)*, 1- 6.
- NEGASH, S. 2004. Business intelligence. *Communications of the Association for Information Systems*, 13, 177-195.
- NELSON, G. 2007. Planning for and Designing a Data Warehouse: A Hands on Workshop *SAS Global Forum*, 111, 1 - 16.
- NGUYEN, T. M., TJOA, M., NEMEC, J. & WINDISCH, M. 2007. An approach towards an event-fed solution for slowly changing dimensions in data warehouses with a detailed case study. *Data & Knowledge Engineering*, 63, 26–43.
- NGWENYAMA, O. The critical social theory approach to information systems: problems and challenges. 1990 Denmark. IFIP WG 8.2 Working Conference.

NHMRC 1999. *A guide to the development, implementation and evaluation of clinical practice guidelines*, Canberra, Commonwealth of Australia.

NHMRC. 2011. *How NHMRC develops its guidelines* [Online]. Available: <http://www.nhmrc.gov.au/guidelines/how-nhmrc-develops-its-guidelines> [Accessed November 6 2013].

NILAKANTA, S., SCHEIBE, K. & RAI, A. 2008. Dimensional issues in agricultural data warehouse designs. *computers and electronics in agriculture*, 60, 263- 278.

NORI, M. 2013. How Mobile Changes the BI Experience. *SAS Global Forum*, 053, 1-11.

O'LEARY, D. 1999. REAL-D: A schema for data warehouses. *Journal of Information Systems*, 13, 49–62.

OLSEN, D. & BRYANT, P.-D. 2012. Business Intelligence And Information Systems: Enhancing Student Knowledge In Database Courses. *Review of Business Information Systems (RBIS)*, 16, 1-14.

ORACLE. 2002. *Extraction in data warehouses* [Online]. Available: http://docs.oracle.com/cd/B10501_01/server.920/a96520/extract.htm [Accessed 22 June 2014].

ORACLE. 2013. *Introduction to data warehousing concepts* [Online]. Available: http://docs.oracle.com/cd/E16655_01/server.121/e17749/concept.htm#DWHSG001 [Accessed April 28 2014].

OU, L. & PENG, H. 2006. Knowledge and Process Based Decision Support in Business Intelligence System. *Institute of electrical and electronics engineers (IEEE)*, 2, 780 - 786.

OWEI, V., BADA, A. & ANIEBONAM, M. 2006. Addressing the information technology skills shortage in developing countries: Tapping the pool of disabled users. *Comm & ethics in society*, 4, 77-89.

PAN, X., VALERDI, R. & KANG, R. 2013. Systems Thinking: A comparison between chinese and western approaches. *Science direct*, 16.

PAPAKONSTANTINOY, Y., PETROPOULOS, M. & VASSALOS, V. 2002. QURSED: Querying and reporting semistructured data. *ACM SIGMOD International conference on management of data*. Madison: ACM SIGMOD.

PARENT, C. & SPACCAPIETRA, S. 2000. Database integration the key to data interoperability. *The MIT Press*, 1- 30.

PARKER, S. 2000. *Towards a practical understanding of operational boundaries using critical systems heuristics*. Master, University of Auckland.

PATEL, N. 1995. Application of soft systems methodology to the real world process of teaching and learning. *International Journal of Educational Management*, 9, 13-23.

PAYANDEH, F. 2013. *BI vs. Big Data vs. Data analytics by example* [Online]. Available: <http://www.datasciencecentral.com/profiles/blogs/bi-vs-big-data-vs-data-analytics-by-example> [Accessed December 12 2013].

PEREZ, R. 2012. *Guest post: The future of data visualization* [Online]. Available: <http://www.tableausoftware.com/about/blog/2012/04/guest-post-future-data-visualization-16578> [Accessed May 7 2014].

PERFORMANCE, U. R. B. A. 2015. *Academic performance, university ranking, top universities in south africa* [Online]. Available: <http://www.urapcenter.org/2014/country.php?ccode=ZA&rank=all> [Accessed 10 August 2015].

PERKINS, A. 2003. Critical success factors for data warehouse engineering. Available: <http://www.visible.com/Company/whitepapers/dwcsf.pdf>.

PITRE, P. 2013. *Business Intelligence: What are the pros and cons of being a bi developer?* [Online]. Available: <http://www.quora.com/Business-Intelligence/What-are-the-pros-and-cons-of-being-a-BI-developer> [Accessed 24 June 2015].

POWER, D. 2007. *A brief history of decision support systems* [Online]. Available: <http://dssresources.com/history/dsshhistory.html> [Accessed October 15 2013].

PYSTER, A., TURNER, R., HENRY, D., LASFER, K. & BERNSTEIN, L. 2009. Master's degrees in software engineering: An analysis of 28 university programs. *Institute of electrical and electronics engineers computer society*, 26, 94 - 101.

RAHMATIAN, S. 2003. Transaction processing systems. *Encyclopedia of information systems*.

RALPH, K. 2008. *Fact tables* [Online]. Available: <http://www.kimballgroup.com/2008/11/05/fact-tables/> [Accessed December 13 2013].

RAMAKRISHNAN, K. & YASIN, N. 2011. Higher Learning Institution - Industry Collaboration: A necessity to improve teaching and learning process. *The 6th International Conference on*. Singapore: IEEE.

RANJAN, J. 2008. Business justification with business intelligence. . *The journal of information and knowledge management systems*, 38, 461-475.

RANJAN, J. 2009a. Business intelligence: concepts, components, techniques and benefits. *Journal of theoretical and applied information technology*, 9, 60-70.

RANJAN, J. 2009b. Business intelligence: concepts, components, techniques and benefits. *Journal of Theoretical and Applied Information Technology*, 60-70.

RAO, R. 2003. From unstructured; data to actionable; intelligence. *IEEE Computer Society*, 29- 35.

REEVES, J. 2015a. *Master of science in data analytics* [Online]. Available: <http://www.nu.edu/OurPrograms/SchoolOfEngineeringAndTechnology/ComputerScienceAndInformationSystems/Programs/MasterofScienceinDataAnalytics.html> [Accessed June 15 2015].

REEVES, J. 2015b. *Master of Science in Data Analytics* [Online]. Available: <http://www.nu.edu/OurPrograms/SchoolOfEngineeringAndTechnology/ComputerScienceAndInformationSystems/Programs/MasterofScienceinDataAnalytics.html> [Accessed 15 June 2015].

RIFAIE, M., KIANMEHR, K., ALHAJ, R. & RIDLEY, M. J. 2008. Data warehouse architecture and design. *Information reuse of integration*. Las Vega: Institute of Electrical and Electronics Engineers.

RMIT. 2015. *Course title: Knowledge and data warehousing* [Online]. Available: <http://www1.rmit.edu.au/courses/004171> [Accessed 12 June 2015].

RODGERS, P. Data warehouse an OLTP DBA's nightmare. 1996. Oracle User Resource, Inc.

ROSS, M. 2005. *Slowly changing dimensions are not always as easy as 1, 2, 3* [Online]. Available: <http://www.kimballgroup.com/2005/03/10/slowly-changing-dimensions-are-not-always-as-easy-as-1-2-3/> [Accessed May 5 2016].

ROSS, M. 2009. *Kimball university: the 10 essential rules of dimensional modeling* [Online]. Available: http://www.informationweek.com/software/information-management/kimball-university-the-10-essential-rules-of-dimensional-modeling/d/d-id/1080009?page_number=1 [Accessed December 20 2013].

ROSS, M. 2013a. *Design tip #152 slowly changing dimension types 0, 4, 5, 6 and 7* [Online]. Available: <http://www.kimballgroup.com/2013/02/05/design-tip-152-slowly-changing-dimension-types-0-4-5-6-7/> [Accessed June 23 2014].

ROSS, M. 2013b. *Design tip #152 slowly changing dimension types 0, 4, 5, 6 and 7* [Online]. Available: <http://www.kimballgroup.com/2013/02/05/design-tip-152-slowly-changing-dimension-types-0-4-5-6-7/> [Accessed May 5 2016].

ROWLANDS, B. 2005a. Grounded in practice: using Interpretive research to build theory. *The electronic journal of business research methodology*, 81-92.

ROWLANDS, B. 2005b. Grounded in practice: using Interpretive research to build theory. *The electronic journal of business research methodology*, 3, 81-92.

RUD, O. P. 2009. *Business intelligence success factors: Tools for aligning your business in the global economy*, New Jersey, John Wiley & Sons, Inc.

RUSSOM, P. 2001. Big data and analytics. *TDWI research*, 1- 6.

RUTHELLEN, J. 2004. The hermeneutics of faith and the hermeneutics of suspicion. *SAGE publications inc*, 14, 1-28.

SAHAY, B. & RANJAN, J. 2008. Real time business intelligence in supply chain analytics. *Information management & computer security*, 16, 28- 48.

SALDANA 2008. An Introduction to Codes and Coding. *SAGE Journals*, 1- 31.

SARGEANT, J. 2012. Qualitative research part II: Participants, analysis, and qualitative assurance. *Journal of graduate medical education*, 4, 1-3.

SCHADEWITZ, N. & TIMOTHY, J. 2007. Comparing inductive and deductive methodologies for design patterns identification and articulation. *iasdr07*, 1-19.

SCHLEGEL, K., SALLAM, R., YUEN, D. & TAPADINHAS, J. 2013. Magic quadrant for business intelligence and analytics platforms. *Gartner*, 2- 64.

SCHNEIDER, M. 2008. A general model for the design of data warehouses. *Science direct*, 112, 309–325.

SCHUTZ, C. 2010. *Extending data warehouses with hetero-homogeneous dimension hierarchies and cubes*.

SEAMAN, C. 1999. Qualitative Methods in Empirical Studies of Software Engineering. *IEEE transactions on software engineering*, 25, 557-572.

SENGE, P. 1990. *The Fifth Discipline: The Art And Practice Of The Learning Organization*, New York, Doubleday.

SEUFERT, A. & SCHIEFER, J. 2005. Enhanced business intelligence - supporting business processes with real-Time business analytics. *Institute of electrical and electronics engineers (IEEE)*, 919 - 925.

SHARMA, R. & DJIAW, V. 2011. Realising the strategic impact of business intelligence tools. *Emerald*, 41, 113-131.

SHARMILA, S. & SUBRAMANI, A. 2013. Impact of business intelligence tools in executive information systems. *International journal of computer engineering & technology (ijcet)*, 4, 01-07.

SIMITSIS, A. & VASSILIADIS, P. 2008. A method for the mapping of conceptual designs to logical blueprints for ETL processes. *Decision support systems*, 45, 22–40.

SIMON, H. 2006. *The Sciences of the Artificia*, Cambridge, MIT Press.

SIMON, K.-H. 2003. Critical systems thinking. *Eolss sample chapters systems science and cybernetics*, 2, 1- 6.

SOLER, E., TRUJILLO, J., FERNÁNDEZ-MEDINA, E. & PIATTINI, M. 2008. Building a secure star schema in data warehouses by an extension of the relational package from CWM. *Computer Standards & Interfaces*, 30, 341–350.

SONDEREGGER, P. & STRATEGIST, C. 2012. *Big data killed the data modeling star* [Online]. Available: <http://tdwi.org/articles/2012/02/07/big-data-killed-data-modeling-star.aspx> [Accessed 28 June 2015].

STANGARONE, J. 2012. *6 big Business Intelligence trends of the near future* [Online]. Available: <http://www.mrc-productivity.com/blog/2012/06/6-big-business-intelligence-trends-of-the-near-future/> [Accessed May 7 2014].

STEPHENS, R. & PLEW, R. 2001. *Database design*, Indianapolis, SAMS.

STUDER, R., BENJAMINS, R. & FENSEL, D. 1998. Knowledge engineering: principles and methods. *Data & Knowledge Engineering*, 25, 161-197.

TALEGHANI, M. 2011. Executive information systems development lifecycle. *Institute of electrical and electronics engineers (IEEE)*. Harbin.

TALWAR, K. & GOSAIN, A. 2012. Hierarchy classification for data warehouse: a survey. *Procedia Technology*, 6, 460 – 468.

TANKARD, C. & PATHWAYS, D. 2012. Big data security. *Network Security*, 2012, 5-8.

TAYLOR, P. & MEDINA, M. 2011. Educational Research Paradigms: From Positivism to Multiparadigmatic. *College Research Journal*, 9-23.

THOO, E., FRIEDMAN, T. & BEYER, M. 2013. Magic quadrant for data integration tools. *Gartner*, 3, 1- 19.

THORNTHWAITE, W. 2006. *Standard reports: Basics for business users* [Online]. Available: <http://www.kimballgroup.com/2006/02/standard-reports-basics-for-business-users/> [Accessed 5 June 2015].

TRAUTH, M., DEINO, A., BERGNER, A. & STRECKER, M. 2003. East African climate change and orbital forcing during the last 175 kyr BP. *Elsevier Science*, 206.

TURKETT, A. 2004. Part 1: Qualitative research sampling - the very real complexities. *Nurse researcher*, 12, 47-61.

ULRICH, W. 1983. *Critical heuristics of social planning*, Chichester England, John Wiley & Sons Ltd.

ULRICH, W. 1996. *A primer to critical systems heuristics for action researchers*, Broughton, South Humberside, UK, Werner Ulrich.

ULRICH, W. 2001. *A mini-primer of critical systems heuristics* [Online]. Available: <http://wulrich.com/csh.html> [Accessed 12 December 2014].

ULRICH, W. 2003. *A brief introduction to critical systems heuristics (csh)* [Online]. Available: http://projects.kmi.open.ac.uk/ecosensus/publications/ulrich_csh_intro.pdf [Accessed 13 April 2013].

ULRICH, W. 2005a. *A Brief Introduction to Critical Systems Heuristics (CSH)* [Online]. Available: http://projects.kmi.open.ac.uk/ecosensus/publications/ulrich_csh_intro.pdf [Accessed September 7 2014].

ULRICH, W. 2005b. *A brief introduction to critical systems heuristics (csh)* [Online]. Available: http://www.wulrich.com/downloads/ulrich_2005f.pdf [Accessed 15 March 2015].

ULRICH, W. & REYNOLDS, M. 2010a. Critical systems heuristics. in: reynolds, martin and holwell, sue eds. *systems approaches to managing change: a practical guide*.

ULRICH, W. & REYNOLDS, M. 2010b. *Critical systems heuristics*. In: Reynolds, Martin and

Holwell, Sue eds. *Systems Approaches to Managing Change: A Practical Guide*, London, Springer.

VASSILIADISA, P., QUIXB, C., VASSILIOUA, Y. & JARKE, M. 2001. Data warehouse process management. *Information systems*, 26, 205–236.

VERKOOIJ, K. & SPRUIT, M. 2013. Mobile business intelligence: key considerations for implementations projects. *Journal of Computer Information Systems*, 54, 23-33.

WAHYUNI, D. 2012a. The research design maze: Understanding paradigms, cases, methods and methodologies. *Jamar*, 69-80.

WAHYUNI, D. 2012b. The research design maze: Understanding paradigms, cases, methods and methodologies. *Jamar*, 69 - 80.

WALSHAM, G. 2006. Doing interpretive research. *European journal of information systems*, 15, 320–330.

WANG, H. C. & GUO, J.-L. 2013. Constructing a water quality 2.0 OLAP system in Taiwan. *Journal of Cleaner Production* 40, 40-45.

WATSON, H. & ARIYACHANDRA, T. 2005. Data warehouse architectures: factors in the selection decision and the success of the architectures. Available: http://www.terry.uga.edu/~hwatson/DW_Architecture_Report.pdf [Accessed May 05].

WEST, W. 2011. *Guidelines versus policies* [Online]. Available: <http://medicaleconomics.modernmedicine.com/medical-economics/news/modernmedicine/modern-medicine-now/guidelines-versus-policies> [Accessed November 6 2013].

WILLIAMS, P. 2012. *A short history of data warehousing* [Online]. Available: <http://www.dataversity.net/a-short-history-of-data-warehousing/> [Accessed April 4 2014].

WOLNIEWICZ, R. 2003. *Using predictive analytics to battle churn and increase profits* [Online]. Available: <http://www.billingworld.com/articles/2003/11/using-predictive-analytics-to-battle-churn-and-in.aspx> [Accessed June 23 2014].

YEASMIN, S. & RAHMAN, K. 2012a. Triangulation research method as the tool of social science research. *BUP Journal*, 1, 154 - 163.

YEASMIN, S. & RAHMAN, K. 2012b. Triangulation research method as the tool of social science research. *BUP Journal*, 1, 154 - 163.

YENICE, N. 2011. Investigating pre-service science teachers' critical thinking dispositions and problem solving skills in terms of different variables. *Educational Research and Reviews*, 6, 497-508.

YUAN, L. & ZOU, H. 2006. Fast rollup on recursive hierarchy in OLAP. *Institute of electrical and electronics engineers (IEEE)*, 1019 - 1028.

ZAIRI, M., OAKLAND, J. & CHANG, S.-H. 1998. Achieving a successful EIS: linking TQM and best. *Integrated Manufacturing*, 9, 50–61.

ZENG, L., XU, L., SHI, Z., WANG, M. & WU, W. 2006. Techniques, process, and enterprise solutions of business intelligence. *Institute of electrical and electronics engineers (IEEE)*, 4722- 4726.

ZHANG, Y. & WILDEMUTH, B. 2009a. Qualitative Analysis of Content. *Applications of social research methods to questions in information and library science*, 308–319.

ZHANG, Y. & WILDEMUTH, B. 2009b. Qualitative analysis of content. *Applications of social research methods to questions in information and library science*, 308–319.

ZHAO, H.-L. Application of olap to the analysis of the curriculum chosen by students. Anti-counterfeiting, security and identification, 2008 Guiyang Institute of Electrical and Electronics Engineers, 97 - 100

ZIMLICHMAN, E. & MEILEK-WEISS, A. 2004. Clinical guidelines as a tool for ensuring good clinical practice. *IMAJ*, 6, 626-627.

ZINS, C. 2007. Conceptual approaches for defining data, information, and knowledge. *Journal of the american society for information science and technology*, 58, 479–493.

ZUBER-SKERRITT, O. & FLETCHER, M. 2007. The quality of an action research thesis in the social sciences. *Emerald insight*, 15, 413-436.