CONCEPTUAL DESIGN FRAMEWORK FOR INFORMATION VISUALIZATION TO SUPPORT MULTIDIMENSIONAL DATASETS IN HIGHER EDUCATION INSTITUTIONS

AKANMU SEMIU AYOBAMI

DOCTOR OF PHILOSOPHY
UNIVERSITI UTARA MALAYSIA
2016

Awang Had Salleh
Graduate School
of Arts And Sciences

Universiti Utara Malaysia

PERAKUAN KERJA TESIS / DISERTASI
(Certification of thesis / dissertation)

Kami, yang bertandatangan, memperakukan bahawa
(We, the undersigned, certify that)

AKANMU SEMIU AYOBAMI

calon untuk Ijazah PhD
(candidate for the degree of)

telah mengemukakan tesis / disertasi yang bertajuk:
(has presented his/her thesis / dissertation of the following title):

"CONCEPTUAL DESIGN FRAMEWORK FOR INFORMATION VISUALIZATION TO SUPPORT MULTIDIMENSIONAL DATASETS IN HIGHER EDUCATION INSTITUTIONS"

seperti yang tercatat di muka surat tajuk dan kulit tesis / disertasi.
(as it appears on the title page and front cover of the thesis / dissertation).

Bahawa tesis/disertasi tersebut boleh diterima dari segi bentuk serta kandungan dan meliputi bidang ilmu dengan memuaskan, sebagaimana yang ditunjukkan oleh calon dalam ujian lisan yang diadakan pada : ***31Julai 2016.***

That the said thesis/dissertation is acceptable in form and content and displays a satisfactory knowledge of the field of study as demonstrated by the candidate through an oral examination held on: ***July 31, 2016.***

Pengerusi Viva: *(Chairman for VIVA)* — Prof. Dr. Suhaidi Hassan — Tandatangan *(Signature)*

Pemeriksa Luar: *(External Examiner)* — Prof. Dr. Abdul Azim Abd Ghani — Tandatangan *(Signature)*

Pemeriksa Dalam: *(Internal Examiner)* — Assoc. Prof. Dr. Haslina Mohd — Tandatangan *(Signature)*

Nama Penyelia/Penyelia-penyelia: *(Name of Supervisor/Supervisors)* — Prof. Dr. Zulikha Jamaludin — Tandatangan *(Signature)*

Tarikh:
(Date) ***July 31, 2016***

Permission to Use

In presenting this thesis in fulfilment of the requirements for a postgraduate degree from Universiti Utara Malaysia, I agree that the Universiti Library may make it freely available for inspection. I further agree that permission for the copying of this thesis in any manner, in whole or in part, for scholarly purpose may be granted by my supervisor(s) or, in their absence, by the Dean of Awang Had Salleh Graduate School of Arts and Sciences. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my thesis.

Requests for permission to copy or to make other use of materials in this thesis, in whole or in part, should be addressed to:

Dean of Awang Had Salleh Graduate School of Arts and Sciences
UUM College of Arts and Sciences
Universiti Utara Malaysia
06010 UUM Sintok

Ikhtisar

Visualisasi maklumat (InfoVis) digunakan dalam pelbagai aplikasi kerana keupayaannya menyelesaikan masalah lambakan maklumat dalam data institusi. Para pembuat polisi dan pembuat keputusan di institusi pengajian tinggi (IPT) juga tidak terlepas daripada masalah lambakan maklumat semasa berinteraksi dengan data pelajar disebabkan oleh kepelbagaian dimensinya. Hal ini mengekang proses pembuatan keputusan dan kerana itu satu kerangka kerja reka bentuk konsep InfoVis tersendiri diperlukan untuk menghasilkan alatan InfoVis khusus domain. Sehubungan itu, matlamat kajian ini adalah untuk mereka bentuk kerangka kerja InfoVis yang dilengkapi dengan teknik penyampaian kandungan dan proses sistematik dalam menzahirkan InfoVis khusus domain bagi menangani data pelajar IPT. Terdapat empat fasa yang terlibat dalam kajian ini: 1) kajian pengguna untuk mengkaji, mendapat dan memberi keutamaan pada data pelajar yang menjadi pilihan kepada polisi IPT. Dimensi data pelajar yang sepadan kemudiannya dikategorikan, 2) kajian tinjauan melalui analisis kandungan terhadap kepustakaan reka bentuk InfoVis, dan padanan terhadap dapatan daripada kajian pengguna untuk mencadangkan teknik visualisasi, interaksi dan herotan yang sesuai, 3) pembangunan reka bentuk kerangka kerja secara konsep yang menyepadukan model teknik tersebut dengan proses reka bentuk berpandukan penyesuaian kejuruteraan perisian dan model reka bentuk InfoVis, 4) penilaian terhadap kerangka kerja yang dicadangkan melalui tinjauan pakar, prototaip, penilaian heuristik, dan penilaian pengalaman pengguna. Bagi sebuah InfoVis yang mampu menyampai dan mewakili pilihan pengetahuan eksplisit, serta menyokong kepelbagaian dimensi data pelajar dan proses pembuatan keputusan, kajian mendapati bahawa 1) *mouse-on*, *mouse-on-click, mouse on-drag, drop down menu, push button,* kotak semak, dan *dynamics cursor hinting* adalah teknik yang sesuai untuk interaksi, 2) *zooming*, gambaran menyeluruh dengan perincian, penatalan, dan eksplorasi adalah teknik yang sesuai untuk herotan, dan 3) carta garis, *scatter plot*, *map view*, carta bar dan carta pai adalah teknik yang sesuai untuk visualisasi. Sokongan teori kepada kerangka kerja yang dicadangkan mengesyorkan bahawa ketetapan dalam *preattentive processing theory*, *cognitive-fit theory*, dan *normative and descriptive theories* mesti diikuti bagi membolehkan InfoVis membantu persepsi, kognisi dan pembuatan keputusan. Kajian ini menyumbang kepada bidang InfoVis, proses pembuatan keputusan berpandukan data, dan proses penggunaan data pelajar.

Kata Kunci: Visualisasi maklumat, Institusi pendidikan tinggi, Kerangka kerja reka bentuk konsep, teknik penyampaian kandungan, proses pembuatan keputusan berpandukan data.

Abstract

Information Visualization (InfoVis) enjoys diverse adoption and applicability because of its strength in solving the problem of information overload inherent in institutional data. Policy and decision makers of higher education institutions (HEIs) are also experiencing information overload while interacting with students' data, because of its multidimensionality. This constraints decision making processes, and therefore requires a domain-specific InfoVis conceptual design framework which will birth the domain's InfoVis tool. This study therefore aims to design HEI Students' data-focused InfoVis (HSDI) conceptual design framework which addresses the content delivery techniques and the systematic processes in actualizing the domain specific InfoVis. The study involved four phases: 1) a users' study to investigate, elicit and prioritize the students' data-related explicit knowledge preferences of HEI domain policy. The corresponding students' data dimensions are then categorised, 2) exploratory study through content analysis of InfoVis design literatures, and subsequent mapping with findings from the users' study, to propose the appropriate visualization, interaction and distortion techniques for delivering the domain's explicit knowledge preferences, 3) conceptual development of the design framework which integrates the techniques' model with its design process –as identified from adaptation of software engineering and InfoVis design models, 4) evaluation of the proposed framework through expert review, prototyping, heuristics evaluation, and users' experience evaluation. For an InfoVis that will appropriately present and represent the domain explicit knowledge preferences, support the students' data multidimensionality and the decision making processes, the study found that: 1) mouse-on, mouse-on-click, mouse on-drag, drop down menu, push button, check boxes, and dynamics cursor hinting are the appropriate interaction techniques, 2) zooming, overview with details, scrolling, and exploration are the appropriate distortion techniques, and 3) line chart, scatter plot, map view, bar chart and pie chart are the appropriate visualization techniques. The theoretical support to the proposed framework suggests that dictates of preattentive processing theory, cognitive-fit theory, and normative and descriptive theories must be followed for InfoVis to aid perception, cognition and decision making respectively. This study contributes to the area of InfoVis, data-driven decision making process, and HEI students' data usage process.

Keywords: Information Visualization, Higher Education Institutions, Conceptual Design Framework, content delivery techniques, data-driven decision making process

Acknowledgement

First and foremost, I give all thanks to Allah for providing me with the spiritual and intellectual strength needed to accomplish this study. I also appreciate all individuals who contributed to the successful completion of this study. Without you all, this journey would not have been rewarding.

I would like to express my deepest gratitude to my supervisor; Prof. Dr. Zulikha Jamaludin, for her excellent mentorship, motherly care, patience and unflinching financial and moral support throughout this study.

My sincere appreciation is to Universiti Utara Malaysia. The award of Postgraduate Scholarship and PhD Incentive Schemes greatly aided my study. The effect of the financial support towards the necessary emotional stability during the scholarly journey cannot be overemphasised. Also, I want to thank all experts who shared their experiences with me in view of actualizing a reliable outcome.

On a personal note, I want to thank my parents, parents-in-law and all members of our family for their love and prayers. I would also like to extend special appreciation to my wife; Mutiat Folasade, and my kids; Habeebulah and Aadiyatulah, for their perseverance, understanding and love.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

2D	Two Dimensional
3D	Three Dimensional
ACM	Association of Computing Machineries
BI	Business Intelligence
DSS	Decision Support System
EDA	Exploratory Data Analytics
HCI	Human Computer Interaction
HEI	Higher Education Institution
HSDI	HEI Students' Data InfoVis
ICT	Information Communication Technology
IEEE	Institute of Electrical and Electronic Engineers
InfoVis	Information Visualization
JSON	JavaScript Object Notation
UCD	Users-centered Design

CHAPTER ONE
BACKGROUND OF THE STUDY

1.1 Introduction

This chapter espouses the need for Information Visualization framework to support students' data-related decision making process of higher education institutions. It provides the overview of the study, motivation, and the preliminary investigation. This chapter discusses the problem statement, enlists the research questions and the research objectives of this study. It highlights the scope which signifies the boundary of the research conducted. The research deliverables, the study's contributions, research design that explains the research framework and the thesis structure are also included. Lastly, a summary of the chapter is presented.

1.2 Overview of the Study

Information Visualization (InfoVis) is a branch of Visualization if seen as a field of study (Spence, 2007; Sheneiderman & Plaisant, 2010). In this respect, there is also scientific visualization which borders on representation of scientific data such as tensor data, unsteady flow, model and software using appropriate technique and tools (Wright, 2007).

Ware (2000), however, defines visualization as a graphical representation of concepts and data in a manner that supports decision making. Visualization, as human activity, had been in existence far before the computing era. It characterises the period of cave painting, imagery of human and all forms of ancient artistic representation of ideas to stimulate mental understanding (Ward, Grinstein, & Keim, 2010). The distinctive feature of InfoVis is its presentation and representation of the datasets in an interpretative manner to all classes of users, even to the non-tech savvies. Also, InfoVis, as a concept through artefacts called InfoVis tools or systems, supports users' interactions in ways that enhance insights about the datasets, aids exploratory data analysis and supports decision making (Sheneiderman & Plaisant, 2010; Spence, 2007; Ware, 2000). The ability to visualize, i.e. 'to form mental model or mental image of something' (Spence, 2007) is an indispensable quality of InfoVis. Figure 1.1 gives the clue on the importance of visualization.

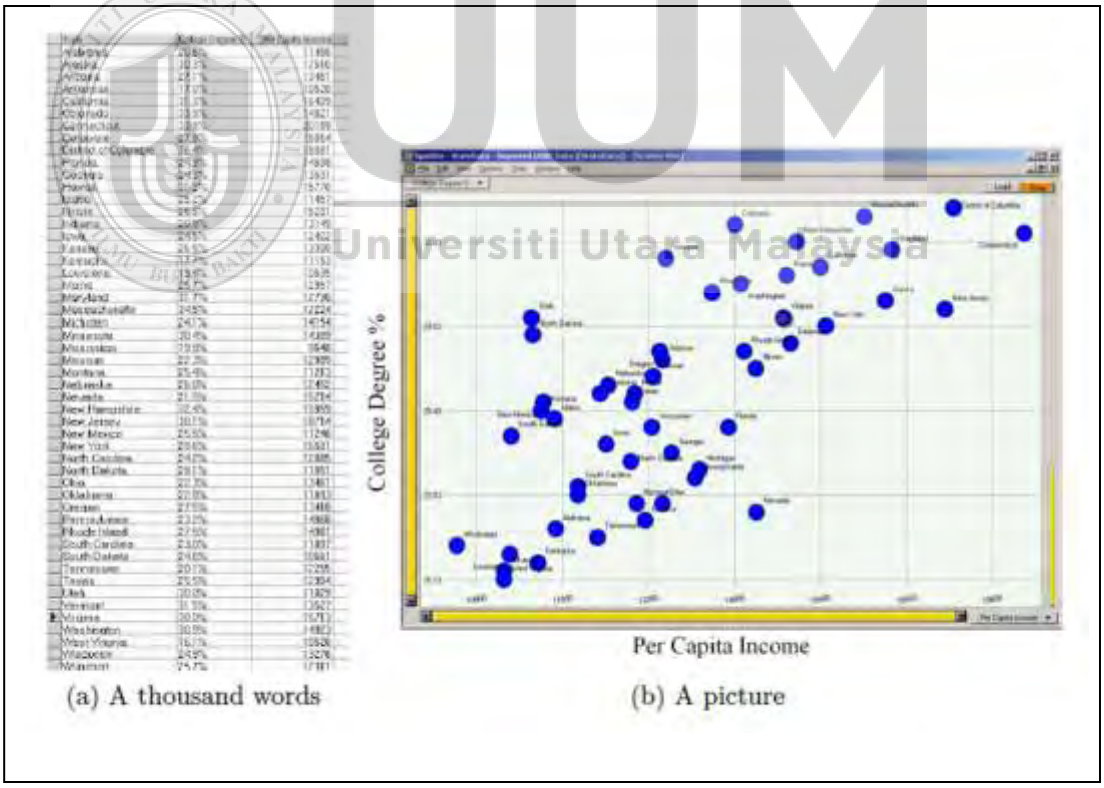


Figure 1.1. "A Picture is worth a thousand words"

(Source: Fekete, van Wijk, Stasko, & North, 2008)

The formation of mental model from data sets can only be actualized when InfoVis aligns with visual exploratory paradigm (Keim & Ward, 2002). This paradigm must also feature the Sheneiderman's information seeking mantra: Overview first, zoom and filter and then detail-on-demand (Sheneiderman, 1996). Therefore, research fields like data management, interaction design and cognitive science do overlap and come into relevance when discussing InfoVis. InfoVis also communicates to the users through statistical graphics like line chart, pie chart, histogram, bar charts and other. The visual exploration through which the users interact with statistical graphics is aided the interaction techniques. This features InfoVis as a subset of Interaction Design (IxD) domain.

Interaction design is an essential tool in the field of InfoVis due to the evolving trend from static to dynamic visualization system (Spence, 2007). Zooming, panning, colour mapping and interactive animations are some of the interaction techniques that are used in designing InfoVis. The purpose of involving interaction techniques is to enhance users' experience by aiding users' control. Also, the abilities to present graphical perception that aids insight, stimulates thinking and generates hypotheses are qualities of any functional InfoVis tool (Spence, 2007; Ware, 2000; Card, Mackinlay, & Shneiderman, 1999). The needed ability to stimulate thinking necessitates the knowledge of cognitive science in InfoVis design. Ware (2000) posits that the ability to visualize is the ability to amplify cognition.

InfoVis represents a domain of study that explores statistics, most especially using statistical graphics for a comprehensible data representation. It also uses interaction techniques to aid users' control. The output of the visual exploration done by InfoVis is expected to give insight about the large and multidimensional datasets. This is to create a mental model of the information displayed with the aid of statistical graphics. InfoVis duly harnesses cognitive science to evaluate mental model and the cognition impact of the InfoVis tool on the users.

Notably, the problem-solving nature of InfoVis has been the driving force for researches in the field. InfoVis is a research domain interested in data discovery and analysis through visual exploration (Fekete & Plaisant, 2002). Meyer (2012) and Stasko (2008) explain that InfoVis centres on making sense of data, and this is dependent of how understandably represented and visually presented is the data to the users. Stasko (2008) further asserts that researchers have the responsibility of transforming raw data to meaningful information that will assist in decision making and creation of knowledge and wisdom. This process of ensuring meaningful representation and insightful presentation of datasets has been employed in many researches basically in solving the problem of information overload.

Problems of information overload can be summarised as the dangers inherent in handling large datasets. These have been responsible for the growing trend of researches in InfoVis. Keim et al. (2008) summarizes information overload problem as getting lost in data which may be irrelevant to the task at hand, processed in an inappropriate way or presented in an inappropriate way. The need to solve the problem of information overload had been the cause of different focuses of InfoVis researches, and in different domains. It has also made InfoVis researches to be focused on identifying the suitable InfoVis techniques for the domain under study. Also, because InfoVis researches are borne to address specific problem of particular domain, novel InfoVis frameworks are designed and implemented in the development cycle of InfoVis tools. These frameworks address the befitting techniques to present analysed data that will be relevant to the task at hand with appropriate data representation techniques. It also illustrates the interaction techniques that aid users' exploration of the datasets (Ning et al., 2012; Matsui et al., 2011; Zheng et al., 2011).

It is noteworthy that datasets are of different nature and type. The two (2D) and three (3D) dimensional data, multidimensional datasets, text and web, software code and algorithm, hierarchies and graph are examples of different data types (Keim, 2002; Keim & Ward, 2002). Each of these data types has their peculiar nature that should be considered in relationship with the domain of usage and the problem to be solved by the InfoVis techniques (Lam et al. 2012). This will inform the befitting choice of the techniques.

InfoVis techniques must align with the nature of the data type, its correlations and the pattern of events to be visualized. This, in totality, determines the choice of the InfoVis techniques, or novel creation of infographics, where necessary. It also suggests that for every data visualization project, research must be conducted to identify the befitting selection and combination of techniques to be used in achieving an understandable visual data presentation and dynamic exploration. This is what InfoVis framework is expected to achieve.

Rogers, Sharp and Preece (2011) conceptualize framework as the inter-linkage between steps, questions, challenges, principles, tactics and dimensions that are structured to guide designers in scoping the user experience for which they are designing. Therefore, InfoVis framework is the modelled inter-relationship and inter-dependence between principles and tactics involved in choosing appropriate visualization and interaction techniques that are aimed to support large and/or multidimensional datasets. Figure 1.2 presents a diagrammatic representation of the overview of InfoVis. The act of supporting the multidimensional dataset is operationalized as solving the inherent problem of information overload, thus aiding the decision making process.

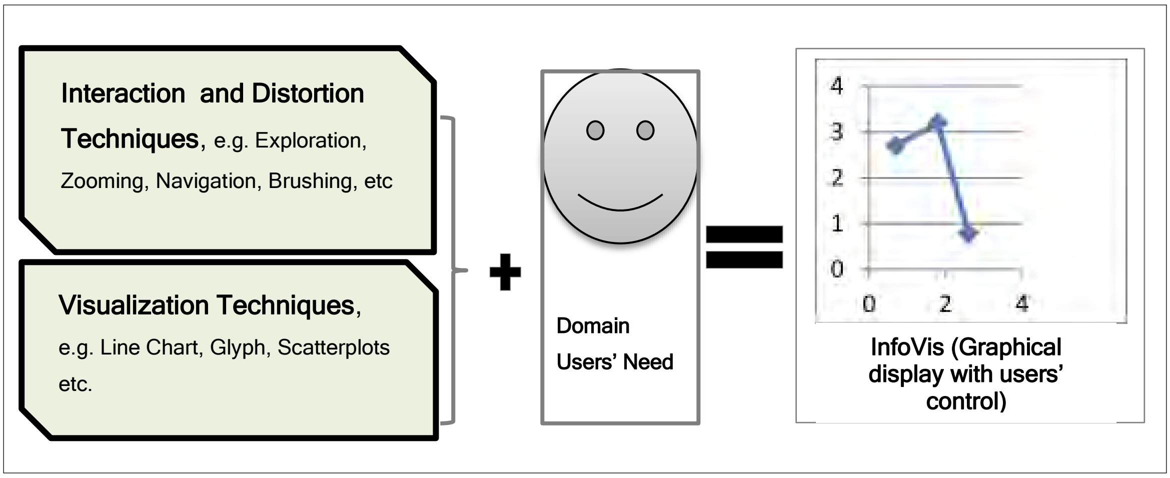


Figure 1.2. An Overview of InfoVis

1.3 Motivation of the Study

Having understood the diverse applicability and indispensable usage of InfoVis, this study is motivated by the lingering problem of information overload (Keim et al., 2010) in the cause of managing the data of higher education institutions (HEI). The HEIs' data is becoming increasingly difficult to analyse due to its fast growing trend. In this regard, getting previously unknown information that could assist policy and decision making in the education sectors is becoming challenging (Oracle Data Sheet, 2011; Yanosky, 2009; Conrad et al., 2013; Gill, Borden, & Hallgren, 2014).

The currently used data management systems by the HEI are the traditional management information systems which are faced with problems of money and time wastage, loss of scientific and industrial opportunities (Keim et al., 2008; Williams & Amin, 2006). Added to this is the usage of static statistical graphics like pie chart, bar chart and line chart. However, the multidimensionality of these datasets and its superabundant volume has limited the success of these traditional data management and data representation (Ehlers et al., 2009). Other recorded cases are occasional deployment of popular statistical tools like SAS and data visualization systems like Tableau for their operational data analysis (Bogard & Green, 2013; Conrad et al., 2013). However, the one-cut-fits-all model type of the software tools pose ineffectiveness in the decision support functionalities of the data management tools. This is because such systems are designed to fit all organisations (with a configuration by a software engineer) irrespective of the difference in their intending explicit knowledge, and consequent data models.

Nonetheless, some past studies have endeavoured to attend to the problems of information overload from the superabundant data volume faced in institutional data management of higher institutions, the vastness of the problems therein still necessitates further researches. Table 1.1 shows the limitations of these past works on the management of the HEI's data.

Table 1.1

Limitations of the past works on HEI's data management

Reference	**Nature of the tool**	**Function of the tool**	**Limitations**
Pocius & Reklaitis, 2000	Web-based interface for data repository	For the institutional strategic management and planning	Absence of visualization and exploratory data analysis features.
Bresfelan et al., 2009	Data mining technology	To proffer better curricula and syllabi design	Absence of visualization and interactive features
El-Fattah, 2012	Visual Data model	To be used in distributing teaching resources	Absence of exploratory data analysis and interactive features.
Sarker et al., 2010	Data warehousing	To be used for sharing and accessing institutions' data among the stakeholders.	Absence of exploratory data analysis and interactive features.
Delavari et al., 2008	Data Mining technology	To give an analytical guide that will enhance decision making	Absence of visualization and interactive features.

From Table 1.1, Sarker et al. (2010) emphasised the role of institutional repositories in addressing the data management challenges of higher education. Their study identified that the inability to adopt new technologies, improve the quality of learning and teaching, and design befitting curriculum can be solved by providing information repository. This repository is suggested to be shareable and accessible by all institutions and education stakeholders. In the same vein, Pocius and Reklaitis (2000) asserted that information system is a strategic tool for information management and planning, and there is a cause for improving the academic studies of higher institutions. Their study thus developed a university information system using web-based interfaces linked with databases.

Bresfelean et al. (2009) suggested that with the trend of competition in the higher education environment, universities should further adopt information communication technologies (ICT). The data mining technologies which are the ICT suggested to tackle data management challenges are also opined to be effective in managing educational capacity utilization (El-Fattah, 2012). Bresfelean et al. (2009) and Delavari et al. (2008) suggested data mining tools for managerial decision making due to the enterprises' reliance on data for strategic decision and prediction.

In essence, the fast growth of data, its multidimensionality, and the consequent problem of information overload are also found in HEI generally, and its students' data specifically. Therefore, the institutional data management has been constrained by the limitations of the presently adopted technologies in the processing of the data, and the representation and presentation of the domain explicit knowledge. This, invariably, has been responsible for the unsatisfactory managerial decision and policy making process. This therefore forms the need for further studies that will attend to data management issues of HEIs' students' data.

1.4 Preliminary Study

A preliminary study is conducted to empirically support the proposition that HEI experiences information overload and decision making constraint –as deduced from the limitations of the past works on HEI's data management. The methods, analysis and findings are discussed in section 1.4.1 – 1.4.3 below.

1.4.1 Method

Information overload is measured by a 9-item survey questionnaire (Appendix B, part ii). This is developed through adaptation of items from previous related studies, and deductions from conceptual explanation made in literatures (Whelan, 2011; Yanosky, 2009; Fekete, van Wijk, Stasko, & North, 2008; Fekete & Plaisant, 2002; Haksever, 2000). Table 1.2 presents the items, their codes and the respective references.

Table 1.2

Items, Codes and the Supporting References

Code	Items **At my institution…**	Sources
IO1	Sensitive data is not secure from unauthorized access	Haksever (2000)
IO2	We are experiencing growth in our volume of data	
IO3	The growth of our data volume needs better analytical tool	Whelan (2011); Yanosky (2009)
IO4	Our data analysis tool cannot be used by some of the administrative personnel	
IO5	Employees do not understand their responsibilities in the use of data	Fekete and Plaisant (2002)
IO6	Employees cannot manipulate our datasets and transform it to their desired information	
IO7	We get maximum administrative decision support from our institutional data	
IO8	We get maximum academic (teaching and learning) value from our institutional data	Fekete, van Wijk, Stasko, & North (2008)
IO9	We get maximum business (students' enrolment, industries consulting) value from our institutional data	

The questionnaires were distributed to thirty-two (32) respondents purposively selected among HEI decision makers. The choice is made based on the fact that HEI decision makers are indefinite population (Tongco, 2007) and the aim of the preliminary study is users' validation (Tullis & Albert, 2008). Data collected are analyzed using Statistical Package for the Social Sciences (SPSS).

As shown in Table 1.2, the items designed to measure the experience of information overload in HEI were firstly coded before input into the analysis tool. Notably, items IO7, IO8 and IO9 were worded in an opposite direction as against the phenomenon that is under study. This was to check for bias in the responses gathered (Pallant, 2011). In order to avoid its negative influence on the statistical result, they were recoded and formed into another items with code REOI7, REOI8 and REOI9 respectively using the SPSS.

A reliability test was conducted to validate the internal consistency of these items and to show that they duly measured the intended phenomenon. For a stronger Cronbach's Alpha value, item IO1 is recommended for deletion. The reliability test gives a Cronbach's Alpha value of 0.712. Value of 0.6 and above for Cronbach's Alpha represents an internal consistency of the items measuring the construct (Pallant, 2011).

For the measure of central tendency, the mean value of the construct was computed as NTOI, i.e. the summation of IO2, IO3, IO4, IO5, IO6, REIO7, REIO8 and REIO9. These items give weighted sum for the assessment of the phenomenon under study; Information Overload in HEI.

1.4.2 Findings

A Cronbach's Alpha value greater than 0.6 is acceptable, while that greater than 0.7 is preferable (Pallant, 2011). Therefore, Cronbach's Alpha value of 0.712 is satisfactory. This shows that items IO2, IO3, IO4, IO5, IO6, REIO7, REIO8 and REIO9 are suitable for the assessment of information overload.

The mean value is given as 4.03 on 5-point Likert scale (1 = Strongly Disagree, 2= Disagree, 3= Neutral, 4= Agree and 5= Strongly Agree). Mean as a measure of central tendency is appropriate in measuring the popular opinion of selected participants in users' study (Tullis & Albert, 2008). The findings therefore showed that HEI data management experiences information overload, is accepted.

1.4.3 Implications of the Preliminary Findings

Basically, the findings of the preliminary study point to the same conclusion made from the review of the past literature of HEI data management issues. It shows HEIs have been experiencing increase in the volume of their datasets generally and the student data specifically. With the increase in the volume of the institutional data and limitations of previously-designed data management tools to handle the data multidimensionality, HEI decision and policy makers are not making comprehensive use of the available data. On the other hand, InfoVis is expected to bridge this gap by making data exploration an activity that can be done by all, with the aid of interactive user interface through visual exploration. From the result of item IO1, HEIs are not experiencing any significant data security problem.

In conclusion, the preliminary study provides (i) empirical evidence that HEIs are experiencing problem of information overload, and (ii) users' validation that improving data-driven decision making process in HEIs is essential and needed. From the foregoing, this study posits that since HEIs are experiencing problem of information overload due to the multidimensionality of the students' data as reported by the data administrators and the limitations of currently used tools. In view of this, a domain-specific InfoVis is necessary. The InfoVis is to be actualized through the InfoVis conceptual framework which specifies the techniques and design process suitable for the display domain's explicit knowledge preferences, and associated users' visual exploratory activities.

1.5 Problem Statement

The experience of getting lost in data that are irrelevant to the current task at hand, processed and presented in inappropriate ways are inherent in HEIs' students' data because of its multidimensionality (Karine, Diane, & Deborah, 2012; Pinto et al., 2012; Bresfelean et al., 2009; Williams & Amin, 2006). This causes the problem of information overload (Keim, 2008). The previous works done to address this problem have only delivered static statistical graphic (Gelman & Unwin, 2012; El-Fattah, 2012; Sarker et al., 2010; Ehlers et al., 2009; Delavari et al., 2008; Pocius & Reklaitis, 2000), as shown in Table 1.1. Notably, none of these studies provided the comprehensive inclusion of visualization, interation, and distortion techniques. And, studies like Cristina et al. (2016), Sabri, Hela, & Mounir et al. (2015), Boo et al. (2015), Sop et al. (2012), and Pinto et al. (2012), among others, did not consider the students' data specific domain explicit knowledge preferences. The domain explicit knowledge preferences are the output of data transformation done by the data analysis tools which must align with the domain users' preferences.

The static representation of data through the statistical graphics does not support users' control, especially its incompatibility with the Sheneiderman's information seeking mantra: Overview first, zoom and filter and then detail-on-demand (Xie, Hofmann, & Cheng, 2014; Gelman & Unwin, 2012; Battista & Cheng, 2011; Fekete et al., 2008). The non-inclusion of the HEI policy makers' explicit knowledge preferences and absence of InfoVis features in the presently-used HEI data tools are responsible for loss in the data that are irrelevant to the task at hand, processed and presented through inappropriate means (Pinto & Carvalho, 2014; Eduventure, 2013; Karine, Diane, & Deborah, 2012; Keim et al, 2011).

An exploratory pre-design study was therefore conducted to validate the findings from literature that HEI decision makers are facing problem of information overload while interacting with students' data because of its multidimensionality. This is earlier reported in section 1.4. In view of the multidimensionality the students' data and the unavailability of user-controlled visual exploration that attends to the domain policy makers' explicit knowledge preferences, there is decision making constraint and under-utilization of available information at the disposal of HEI policy and decision makers. Therefore, there is a need for a development of InfoVis conceptual design framework that specifies the systematic choice and combination of InfoVis techniques (visualization, interaction and distortion). The conceptual design framework would address the specifics of HEI students' data, and include a process model which depicts the stage-by-stage and phase-by-phase approach of translating conceptual design to physical artifact (i.e. prototyping). The process model would complement the currently available InfoVis user-centred models like Meyer (2012), Simon et al. (2011), Koh et al. (2011), and Robinson et al. (2005) which did not contain the process of prototyping. The InfoVis actualized by the conceptual design framework would also address the identified domain explicit knowledge preferences.

1.5.1 Research Gaps

The unavailability of InfoVis conceptual design framework that explicitly states the visualization, interaction and distortion techniques that can represent and present the HEI policy makers' explicit knowledge, with due consideration to the data type and the tasks involved in the work practices. This has always formed the basis of researches in InfoVis (Stasko, 2008; Rogers et al., 2011; Sheinerderman & Plaisant, 2010; Koh et al., 2011; Tan et al., 2006).

There is insufficient theoretical understanding and theory-based scholarly works for InfoVis because it is an evolving research domain (Yi, 2010; Teets et al, 2010; Fekete et al., 2008). Particularly, evaluating InfoVis' decision support has not been well attended to through a provision of a theory-based evaluation framework. This study hopes to collectively attend to these gaps.

1.6 Research Questions

The central question of this study is how can visualization, interaction, and distortion techniques be combined for the representation and presentation of multidimensional (HEI's students) data and the explicit knowledge preferences of HEI policy makers?

In answering this central question, these three (3) questions are put forward:

1. What are the explicit knowledge preferences (in relationship with students' data) of the HEI policy makers?
2. What are the visualization, interaction and distortion techniques that can represent and present HEI students' dataset with due consideration of the domain policy makers' explicit knowledge preferences?
3. Would the proposed visualization, interaction and distortion techniques and the consequent InfoVis tool assist in supporting decision making based on the explicit knowledge identified?

1.7 Research Objectives

The main objective of this study is to propose a conceptual design framework that specifies how visualization, interaction, distortion techniques can be combined for a befitting representation, and presentation of multidimensional (HEI's students) data, and appropriately for related decisions making.

To achieve this main objective, the following are the sub-objectives:

1. To categorise the appropriate data dimensions of the HEIs' students' data based on the identified and prioritised explicit knowledge preferences.
2. To analyse the visualization, interaction and distortion techniques which are suitable for the representation and presentation of the identified explicit knowledge preferences.
3. To design an HEI students' data-specific InfoVis conceptual framework that is suitable for the representation and presentation of the identified explicit knowledge preferences.
4. To evaluate the InfoVis conceptual design framework.

1.8 Scope of the Study

This study aims at designing and evaluating an InfoVis conceptual design framework within the following scope:

i. The components of the framework address the visualization, interaction and distortion techniques for the representation and presentation of the HEIs' policy makers' explicit knowledge. HEI is one of the domains that are yet to have specifically-tailored InfoVis domain research for the understanding of the students' data. The choice and combination of these techniques are based on the findings of the pre-design study and supporting InfoVis design literatures.

ii. Understanding the HEIs' students' data, as highlighted in (i) above, is by the ability to support the decisions borne out of the varieties of explicit knowledge identified.

The InfoVis conceptual design framework is a framework that links the data dimensions, and by extension, associated explicit knowledge, with its InfoVis (visualization, interaction and distortion) techniques (Rogers et al., 2011; Matsui et al., 2011; Zheng et al., 2011; Ning et al., 2012). These techniques are the components that are responsible for the functionalities of any InfoVis tool (Yi et al., 2011; Spence, 2007; Sheinerderman & Plaisant, 2010; Stasko, 2008).

iii. The users' study is to identify the tasks and work practices that support the varieties of decision made by HEI policy makers. The following are the groups of users studied:

a. Data Administrator
b. Data analyst
c. HEI Decision and Policy makers.

The HEI decision and policy makers are any members of the HEI management team, either at the school, faculty, or university's senate level. The sampled respondents reflected positions of registrar, Assistant Registrar, Deans, Deputy Deans and Deputy Vice Chancellor. The users' study at the pre-design phase exclusively employed HEI decision makers from Universiti Utara Malaysia. However, the users' experience evaluation phase involved HEI decision makers from institutions from Malaysia, South Africa, Nigeria, and United States of America.

The exploratory data analysis, representation and presentation must meet the needs of the identified users. In compliant with user-centred design (UCD) approach, the users to be inquired are staff members of HEIs who rely on data of the students' profiles for administrative steps and decision making (Lam et al. 2012; Ziemkiewicz et al., 2012; Sedlmair et al., 2012; Koh et al., 2011; Roth et al., 2010).

iv. The nature of the decisions covered by this study is purely non-curriculum design related, and statistical modelling is excluded.

v. The prototype from the HEI-focused InfoVis conceptual framework is developed based on the findings of the users' study.

vi. This study centres on evaluating the usability, correct data representation and decision support of the designed InfoVis.

1.9 Research Deliverables

This study delivers the following deliverables:

i. HEI students' data-focused InfoVis Conceptual Design Framework

ii. An InfoVis Prototype for HEI students' data.

iii. Conceptual Framework for Data-driven decision making in HEI.

iv. Instrument for Experimental Evaluation Study of InfoVis System.

1.10 Significances of the Study

This study aimed designing an InfoVis conceptual design framework based on the findings of the HEI-focused pre-design study. The unique characteristic of the framework encompasses UCD guidelines, Human-Computer Interaction (HCI), visualization theories, and cognition models. It identifies the essence of users' study in designing a functional InfoVis, and finally delivers a design guide for InfoVis designers. In essence, the identified research gap is attended to by this work. It attempts to solve the problem of information overload experienced in HEI by supporting the students' dataset through a specifically-designed InfoVis.

Also, InfoVis can be categorized as a typical example of a Decision Support System (DSS). The prototype's ability to aid decision making among the policy and decision makers of HEIs is important. It also provides a software development platform for InfoVis software companies in developing an InfoVis system –of organisational capacity –that specifically solves the HEI students' data management challenges.

Theories, concepts and design approaches as explained in this study (Chapters 2 & 3) are relevant for InfoVis designers, information managers and researchers. Specifically, the context of policy and decision making as discussed in this study is an eye opener for necessary policy formulation to strengthen HEI data management. It signifies the importance of data-driven decision making and emphasises the inevitability of a sound data model with regards to HEI students' data.

In summary, this study adds to the pool of current literatures through its presentation of theoretical and research framework. This research framework is adoptable and adaptable by future researches to help in critically examining related theories, concepts and issues.

1.11 Research Design

The research design of this study is the map that shows how this study is conducted. The research design comprises of five phases: Definition of Problem, Suggestion, Development, Evaluation, and Conclusion. Figure 1.3 represents the said research design.

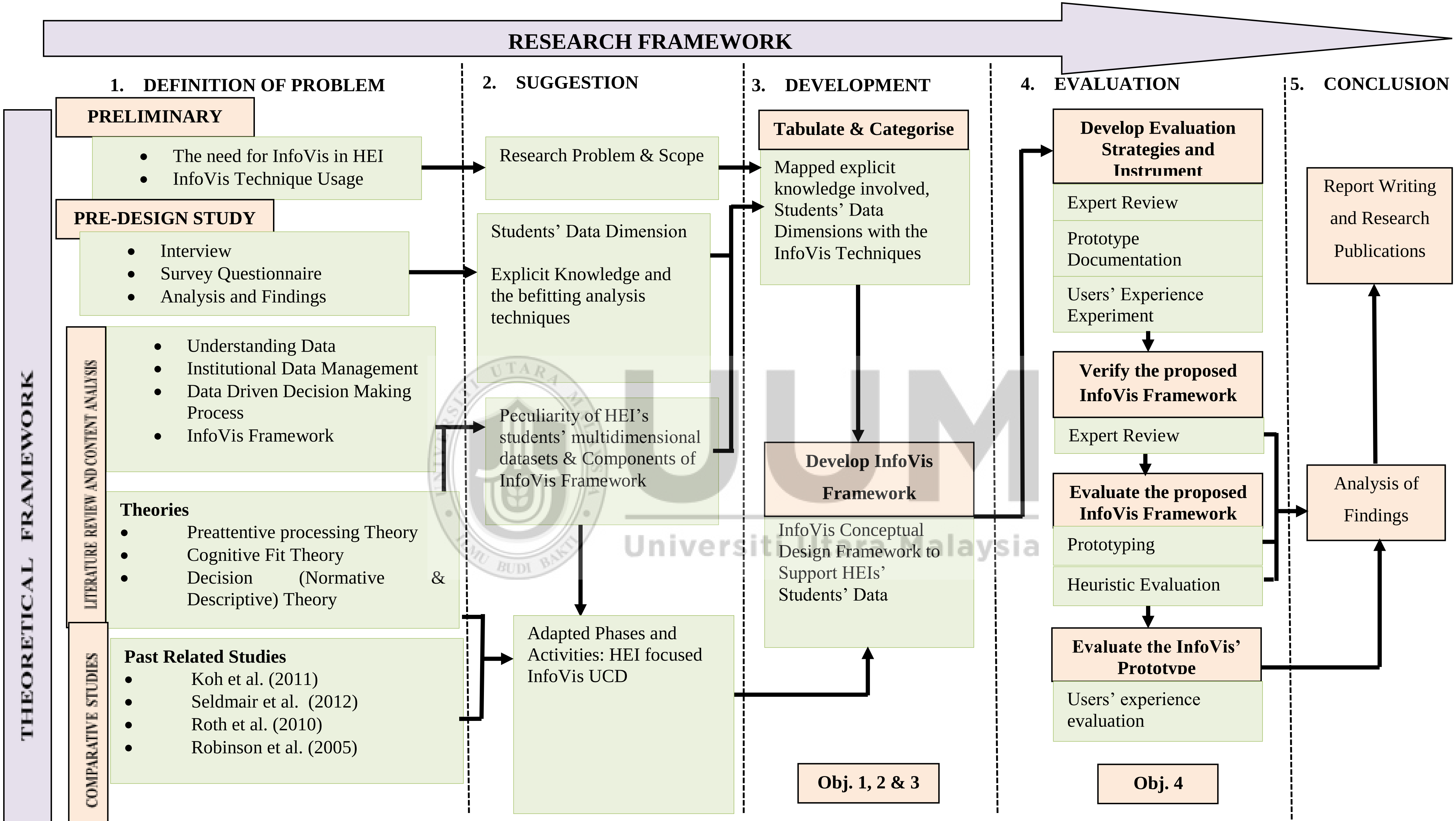


Figure 1.3. Research Design

1.12 Thesis Structure

Chapter 1 of this thesis provides the background, the scope of the study and the study's objectives. Chapter 2 then reviews related literatures, and chapter 3 explains the research methodology, in connection with the research objectives as listed in chapter 1. Based on the stated-objectives of the study, chapter 4 addresses objectives 1, 2 and 3. Chapter 5 and 6 address objective 4, i.e. the evaluation of the proposed framework using expert review, prototyping, heuristic evaluation and users' experience evaluation. Chapter 7 thus concludes, discussing each of the objectives, and the limitations experienced. Figure 1.4 shows the thesis structure.

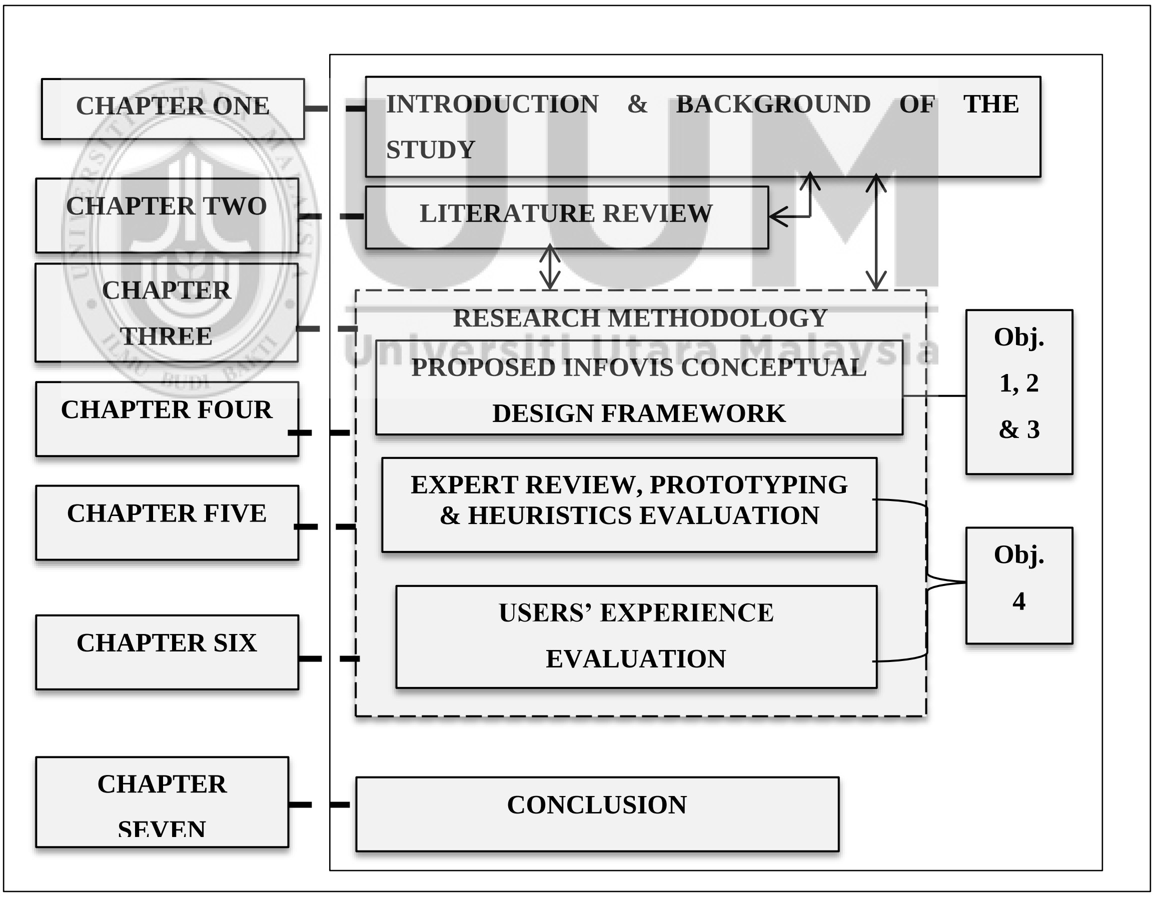


Figure 1.4. Thesis Structure

1.13 Summary of the Chapter

This chapter gives a general overview of InfoVis. It explains that the multidimensionality of students' data has resulted in the problem of information overload. Subsequently, making sense from these datasets has become a difficult task. As a development from the currently used statistical graphics, InfoVis combines visual presentation, dynamic exploration and interactive representation of data to support its multidimensionality. The InfoVis conceptual design framework is the framework that suggests the specific InfoVis' techniques that address the domain explicit knowledge preferences. It also shows how these techniques can be combined and implemented to achieve the domain explicit knowledge preferences. This study is generally motivated by the institutional data management problem in HEIs, and managing its students' data specifically. The research questions and objectives are focused on categorising the appropriate HEI students' data dimensions and analysing the visualization, interaction and distortion techniques suitable of the students' data. The research deliverables, and the study's contributions, are InfoVis conceptual design framework, an InfoVis prototype, data-driven decision making framework, and instrument for experimental evaluation study. The chapter essentially serves as the introduction to the study and shows the research design which highlights the processes involved in this study.

CHAPTER TWO
LITERATURE REVIEW

2.1 Introduction

This chapter presents an extensive review of literatures related to this study. Sections discussed are: Data-driven decision making process in HEIs, HEIs' students' data and InfoVis, InfoVis' techniques, and designing InfoVis framework. Understanding InfoVis from theories and review of past related studies on InfoVis design frameworks and visualization models are also presented in this chapter. This is done to highlight the purpose and process of achieving the stated objectives of this study and its main contribution which is InfoVis conceptual design framework to support multidimensional data of HEI students.

2.2 Data-driven Decision Making Process in Higher Education Institutions (HEIs)

HEIs are one of the leading users of data for decision making process (Bogard & Green, 2013; Oracle Data Sheet, 2011; Yanosky, 2009; Microsoft Education Series, 2004). This is evident in series of studies conducted to attend to varieties of issues in HEIs' data usage. Pocius and Reklaitis (2000) asserted that university information system is a strategic tool for information management and planning. The study also developed web-based interfaces which are linked with HEI databases for exploration of school data. Bresfelean et al. (2009) and Delavari et al. (2008) stated that, with the trend of competition in the higher education environment, universities should further adopt information communication technologies (ICT). Both studies suggested data mining tools for managerial decision making due to the enterprises' reliance on data for strategic decision and prediction. Data-driven conceptual frameworks to guide the usage, analysis and security of HEIs' data were also proposed by Jackson (2013) and Gill, Borden, and Hallgren (2014).

HEIs are part of leading organisations that rely on data for institutional operations (Borden, & Hallgren, 2014; Bogard & Green, 2013; Jackson, 2013; Bresfelean et al., 2009; Yanosky, 2009; Gill, Delavari et al., 2008). Apart from this, the core of the HEI's business values is managing data of the members of staff, students' profiles and records of research and development for making managerial decisions. This therefore posits that HEIs' data-driven decision making processes is a research-demanding niche, especially regarding the statistical and logical handling of corresponding data to meet the organisational demand (Balakrishnam et al., 2012). Table 2.1 gives a summary of past studies on HEIs' data-driven decision making, highlighting the focus of the decision making process and the nature of the corresponding data.

Table 2.1

Summary of Past Studies on HEIs' Data-driven decision making

Focus of the Decision Making Process	**Nature of the Corresponding Data That Aid The Decision**	**Author(s)**
Measuring students' skills	Students' enrolment data	Jackson (2013)
Students' enrolment and	Student enrolment	Bogard & Green (2013),

admission	demographics, students' retention history	Abd El-Fattah (2012)
Raising students' achievement and improving college readiness	Student learning (yearly graduating students' grades (cgpa)) and demographic data, school process and demographic data and Perception data, Student-Lecturer ratio.	Gill, Borden & Hallgren (2014), Lange et al. (2013), Abd El-Fattah (2012), Delavari & Phon-Amnuaisuk (2008), Wagner & Ice (2012), Goss & Hunter (2015)
Instructional Improvement	Students' data including their demographic and disabilities (physical and learning)	Protheroe (2009), Kaufman et al. (2014), Collier (2015)
Curriculum design and alignment with the employment reality, professional development and continuing education.	Students' academic records, Students retention and employability rate, Superintendents' and potential students' ideas about program structures and knowledge needs	Sarker et al. (2010), Hollingworth et al. (2012), Mandinach et al. (2012)
Program design for accreditation and improved university ranking	Enrolments /degrees awarded graduation and retention rates, student-faculty ratios	Ewell (2012), IHEP (2009)
Increasing the minority serving institutions' (MSIs) Learning, retention and racial and gender equality	Educational data comprising descriptive information about the minorities, Students' demographic data	Conrad et al. (2013), Microsoft Inc. (2004)

Students' attrition prediction, retention and recruitment	Students' Retention and Students' Recruitment-related data	Eduventures (2013)

Sarker et al. (2010) emphasized the role of institutional repositories in addressing the data management challenges of HEIs. Their study focused on curriculum design, just as Hollingworth et al. (2012) and Ewell (2012). However, while Sarker et al. (2010) focused on how to achieve curriculum structure that aligned with the employment reality, both Hollingworth et al. (2012) and Ewell (2012) aimed accreditation and improved ranking. This is similar to IHEP (2009), but with preference on strategic positioning of staff and the institution research quality. Abd El-Fattah (2012) gauged the institutions' resources for the purpose of determining the number of fresh students. Others education data-focused researches are on minority serving institutions (Conrad et al., 2014), instructional support for data extraction, professional development and continuing education (Mandinach et al., 2012), and racial and gender equity in admission quota (Microsoft Inc., 2004).

On another hand, students' performance target, its rating and improved readiness (Gill, Borden & Hallgren, 2014; Delavari & Phon-Amnuaisuk, 2008; Wagner & Ice, 2012), instructional improvement (Collier, 2015; Protheroe, 2009), students' enrolment, retention and attrition (Bogard & Green, 2013; Abd El-Fattah, 2012; Eduventures, 2013; Bresfelean et al., 2009) and students' learning goal (Goss & Hunter, 2015; Lange et al., 2012) are researches that focused exclusively on students' data-driven decision making. Each of these studies suggested appropriate data that matches with the decision making focuses. However, none is comprehensive enough to accommodate the domain's students' data decision making interest because of their choices of the related explicit knowledge. This study therefore argues the need for a comprehensive HEIs' students' data-driven decision making (DDDM) approach. This approach encompasses all the decision making focuses and corresponding data –as observed from previous studies. It achieves a summative output that attends to limitations of these previous works, and generate the associated explicit knowledge. This explicit knowledge is the usable extracted information from the relationship of the HEIs' students' data variables. Based on the reviewed past studies on HEI students' DDDM and summarized in Table 2.1, this study presents a HEI decision makers data usage model, shown in Figure 2.1.

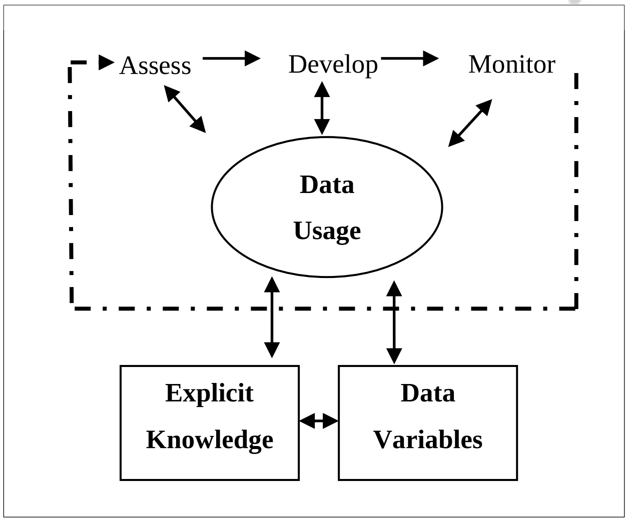


Figure 2.1. HEI Decision Makers' Data Usage Model

(Source: Semiu & Zulikha, 2015c)

The HEI decision makers' data usage model (Semiu & Zulikha, 2015c) suggests that HEI data usage are using data to assess (assessing impact of a policy, or state of an action), to develop (policy, blueprints etc.), and to monitor (developmental progress, etc.). The data variables and the explicit knowledge are used for the DDDM purpose. They are equally essential in designing InfoVis which is a data visualization tool. Therefore, identifying the explicit knowledge preferences of HEI for DDDM, as illustrated in the HEI decision makers' data usage model, is essential in categorizing the appropriate HEI students' data dimensions. This is the background for the first research question and first objective of this study.

2.2.1 Higher Education Institutions' Students' Data and Information Visualization

HEIs' students' data are multidimensional datasets. These are structured datasets of more than three (x, y & z) dimensions (Hao et al., 2010; Card et al 1999), and with at least one dimension being binary (Healey, 1996). Connolly and Begg (2010) explain multidimensional data as a numerical measurement (e.g. property sales revenue data) with associated dimensions which define the attributes of the data element (e.g. location of the property, time of the property sale). The HEI's students' data is a multidimensional structured datasets because many dimensions are associated with its every single data instance.

Notably, HEIs' management of its data has experienced usage of online analytical processing (OLAP), data mining tools (Bresfelan et al., 2009; Delavari et al., 2008), and data warehousing (Sarker et al., 2010). However, the adoption of these tools was only able to attend to the growing size of the students' data, integration and its process automation. The act of making sense and gaining insight of the students' data multidimensionality is still unattended to because the deployed tools lack data exploratory, interactive and visualization features which are the strengths of an InfoVis tool. This situation of handling multidimensional dataset without InfoVis is responsible for the often-experienced information overload by the HEI decision makers while interacting with the students' data (Sarker et al., 2010). This suggests the need for the domain specific InfoVis.

InfoVis is employed when gaining insight and discovering previously unknown explicit knowledge of a multidimensional dataset is of core importance (Sheneiderman & Plaisant, 2010; Spence, 2007). The InfoVis' characteristics enable its support in gaining insight by stimulating questions through exploration and filtering. It has the ability to proffer more narrow part of the data, and allows users to study it more carefully (Sheneiderman, 1996). This allows InfoVis to support multidimensionality of HEI students' data. The users can relate one dimension to the other; explore the dimension of interest, and so on, through the designed InfoVis techniques and frameworks. This is essentially the need for InfoVis conceptual design framemworks which are guidelines that specify the domain specific InfoVis techniques that deliver the domain explicit knowledge preferences.

2.3 InfoVis' Techniques

InfoVis' techniques consist of visualization, interaction and distortion techniques (Keim, 2002). These techniques are expected to be mapped with the nature of the data to be visualized (Chi, 2000), and aligned with the explicit knowledge preferences of the domain decision makers. This is essential to ensure InfoVis supports the decision making process of the domain (Semiu & Zulikha, 2013). There are different types of visualization, interaction and distortion techniques, with different applicability and suitability which are often determined by the domain's data types and the explicit knowledge to be displayed (Keim, 2002; Roth et al., 2010; Robinson et al. 2005).

The choice of appropriate InfoVis' techniques is to achieve undistorted information representation (Spence, 2007; Sheneiderman, 1996). Therefore, choosing a befitting visualization technique is key to the presentation of information that will duly support the decision making process of the users. The interaction and distortion techniques are techniques that aid users' exploratory activities and control, hence, enhance users' experience. However, there is need to analyse each of the types of the InfoVis' techniques in view of identifying their suitability in the representation and presentation of HEI's decision makers' explicit knowledge preferences.

2.3.1 Visualization Techniques

Taxonomy of visualization techniques, following different metrics, have been done by Wright (2007), Keim (2002), Sheneiderman (1996), Spence (2007), and Card, Mackinlay and Sheneiderman (1999). This study, however, categorised visualization techniques to standard 2D/3D display, geometrically-transformed display, iconic display, dense pixel display and stacked display using the Keim's (2002) taxonomy. Although the review of the techniques deals with all varieties, the choice of Keim (2002) taxonomy is because it proffers clearest characterization of the visualization techniques.

This review provides the basis for the choice of the visualization techniques made by this study, as shown in Chapter 4 of this thesis. The chosen visualization techniques are analysed and found to be the best that appropriately presented the explicit knowledge that supports the DDDM process, with specific attention to HEIs' students' data.

2.3.1.1 Standard 2D/3D Techniques

The standard 2D/3D techniques are line graph, bar chart and scatter plots (Chan et al., 2013; Keim et al., 2009). SAS (2011) posited that line graph is appropriate for showing relationship between two variables, and often used to track changes or trends over a particular period of time. Bar chart is said to be the best and mostly used when comparing quantities of different categories like the values of the categories are represented with bar configured in vertices or horizon.

On the other end, scatter plots had been a traditional visualization technique mainly used for x/y representation of multidimensional data. It allows the identification of the relationship between one variable and the other and mostly used for visual exploration of data when used with interaction techniques (Sheneiderman & Plaisant, 2010; Spence, 2007; Tan et al., 2006; Janetzko et al., 2013). Janetzko et al. (2013) further observed that scatter plots is easy to represent data and aid users' perception by observing how a variable is affected by the other i.e. showing a correlation or association among two variables. Presently, there are enhancements with the aid of interactive web application library like JavaScript Object Notation (JSON). This makes scatterplot usable for visualizing multidimensional data, with the categorical attributes being displayed when the data item is hovered by a mouse.

Although scatter plot still features overlap and hides important local patterns, approaches like density, distortion and animation have been proposed to enhance scatter plot as a visualization technique (Maciejewski et al., 2010; Keim et al., 2009; Willems et al., 2009; Aris & Sheneiderman, 2007). This study proposes line chart, bar chart and scatter plot as its visualization techniques because of their support for the time-changing data relationship, comparison of data quantities and x/y representation of multidimensional data, respectively. Figure 2.2 presents examples of 2D/3D visual display techniques.

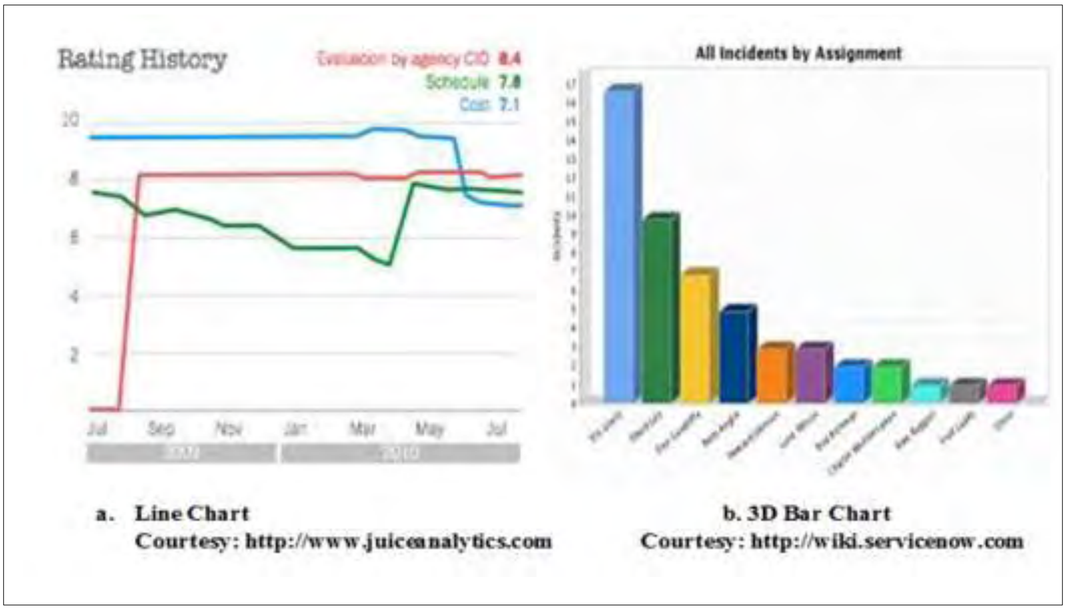


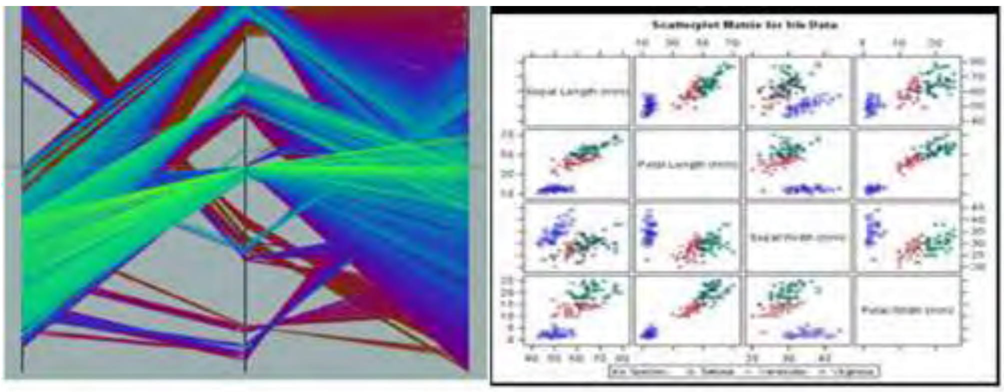
Figure 2.2. 2D/3D Visual Display Techniques

2.3.1.2 Geometrically-Transformed Display

Geometrically-transformed display aims at transforming multidimensional datasets so as to find hidden pattern through exploratory statistics. Examples of geometrically-transformed display are scatter plot matrices, pro-section view, hyperslice and parallel coordinate (Keim, 2002). The most commonly used among this class of visualization techniques are scatter plot matrices and parallel coordinate (Inselberg & Dimsdale, 1990).

Scatter plot matrices is an extension form of scatter plots. It gives an overview of possible configuration of the datasets in the multidimensional space (Elmqvist et al., 2008). The parallel coordinate technique is used by mapping k-dimensional space with two display dimension with the aid of k equidistant axes. The data dimensions to be represented with the parallel coordinate are scaled linearly, ranging from the minimum to the maximum. They correspond to the k equidistant axes. Each of the data item uses a polygonal line as representation, with an intersection with each axes at points that correspond with the value of the dimension considered (Keim, 2002).

This study did not consider the suitability of both scatter plot matrices and parallel coordinate techniques. Although both are suitable for multidimensional data visualization, scatter plot matrices present uncontrollable plots that can equally cause cognitive load. It also allows minimal integration with interaction techniques (Spence, 2007). Also, students' data does not necessarily need grid equidistant axes of parallel coordinate. Figure 2.3 shows examples of geometrically-transform display.

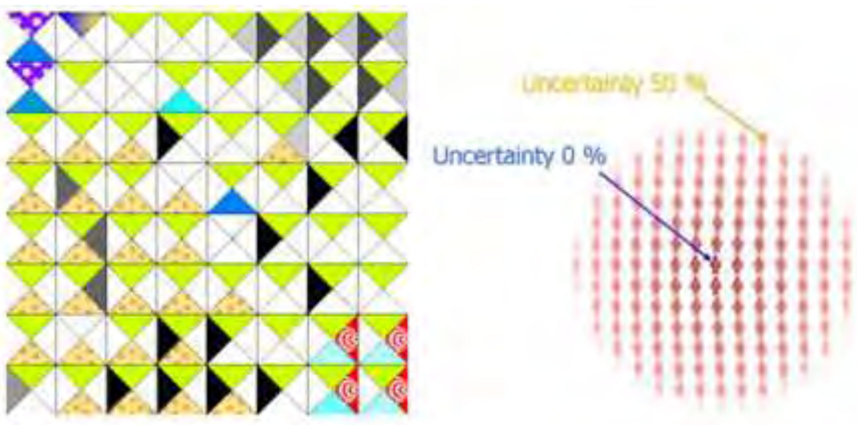


Parallel Coordinate Visualization ©IEEE **Scatter Plot Matrices,**

Figure 2.3. Geometrically-Transform Display

2.3.1.3 Iconic Display

Iconic Display is another type of visualization techniques. It maps the data attributes items to features of an icon. Icons can be of little faces, colour icons, star icons, stick figure, or any arbitrary definition made by the visualization designer (Keim, 2002). Considering the stick figure as an iconic display technique, two dimensions of the data item are mapped to the display dimension. Other dimensions can then be mapped to the angles of the stick figure. This pattern always varies in accordance with the data characteristics as recognised by preattentive perception. This study does not consider iconic display as a befitting visualization technique because it does not conform to the explicit knowledge preferences of the HEI's decision makers. It cannot clearly show relationship between two or more students' data variables as it would be done by scatter plot or bar chart. Figure 2.4 depicts example of iconic display.



Courtesy: https://www.usenix.org Courtesy: http://www.vrlab.buffalo.edu

Figure 2.4. Iconic Display Technique

2.3.1.4 Dense Pixel Display

Dense Pixel Display is characterised by mapping of the data dimension value to a coloured pixel, and supported with a grouping of the pixel that displays one pixel per data value. This technique allows visualization of large amount of data of about 10^6 data values. Its capability of displaying large amount of data is attracting mainstream researches in data visualization (Oelke et al., 2011).

The arrangement of the pixel in an appropriate way aids well represented local correlations, dependencies and hot spots (Keim, 2002). Recursive pattern technique and circle segment technique are popular examples of dense pixel display. Recursive pattern technique is mostly used for time series data with its working based on recursive back-and-forth arrangement of the pixels. The circle segments technique represents data in a circle using its segments as a representation for each data attribute. Hence, the attribute is visualized by a single coloured pixel. Pixel-based visualization is effective in visualizing non-random data distribution, and a set of data items that are not sparse (Keim et al., 1995; Ankerst et al., 1996). This study did not consider dense pixel visualization as a suitable visualization technique because of its incompatibility with random data and sparse data items like the students' data. Figure 2.5 depicts visualizations in dense pixel display techniques.

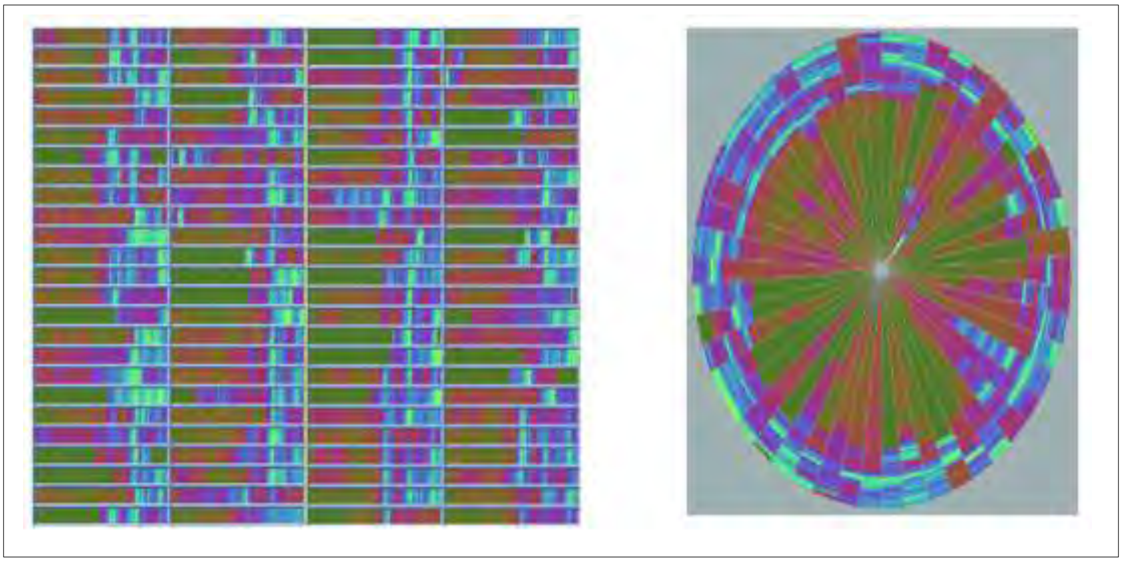


Figure 2.5. Recursive Pattern and Circle Segment Techniques: Dense Pixel Display Techniques

(Source: Keim, 2002)

2.3.1.5 Stacked Display

Stacked display is a visualization technique designed to represent data in a hierarchical pattern (Keim, 2002). Dimensional stacking is a popular type of stack display visualization technique. Examples of this are network diagram, cone tree and Treemap. Stacked display is ensured by embedding one coordinate system into another, with two attributes forming the outer coordinate system. The iterative cycle gives a resulting visualization that largely depends on the distribution of the data in the outer coordinates (LeBlanc et al., 1990). Due to this, the choice of the dimensions must be carefully done, with the most important dimensions given due priority.

One of the prominent methods of visualization that can be categorised under stacked display is radial method. Draper et al. (2009) stated that radial method is displaying data in circular or elliptical pattern. It is an appropriate visualization method for descriptive numerical data, hierarchical pattern, ranking of search results and serial periodic data. Pie charts (Playfair, 1801), Star plots (Elmqvist et al., 2007) and Sociograms (Merriam-Webster, 2007) are evolved display methods from radial method, with their respective design patterns (such as polar plot, space filling and ring) which can be adapted to represent a befitting visualization display. The space filling design pattern is to solve the non-uniform distribution of nodes observed in the tree and star. This birthed concentric, spiral and euler as design patterns, they are named radial space filling. The concentric and euler are best used for viewing hierarchical structures like trees, while spiral is best for viewing serial periodic data.

The Ring-based is of connected and disconnected patterns, but both are suitable for viewing relationship among disparate entities (Draper et al., 2009). This study did consider pie chart (but later suggested during the heuristic evaluation phase, and reported in Chapter 5 of this thesis) as stacked display visualization techniques Figure 2.6 shows examples of stacked display visualization technique.

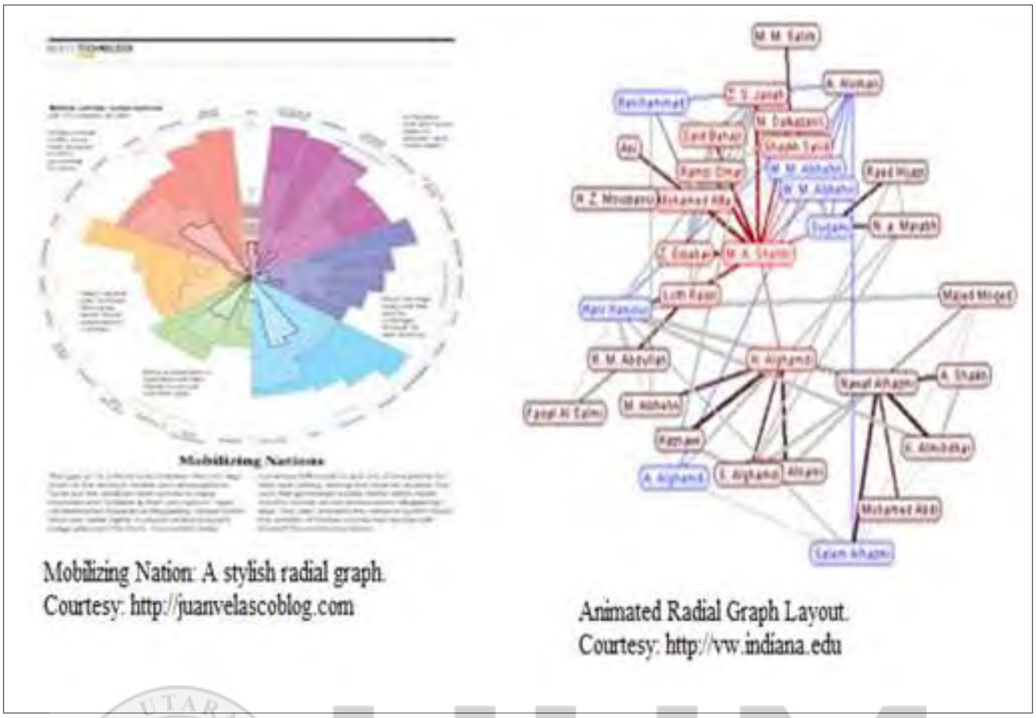


Figure 2.6. Radial Graphs: Stacked Display Techniques

2.3.2 Interaction and Distortion Techniques

The interaction and distortion techniques enable users' control. The last two steps of the Sheneiderman's information seeking mantra (Sheneiderman, 1996), zoom and filter, and detail on demand, are only realizable through the actions of interaction and distortion techniques. The interaction techniques aid direct users' interactions with the visualizations (graphical data representation). The visualizations, interactions and distortion techniques are combined according to the data exploratory purposes of the domain users (Schmidt et al., 2005; White, 2008).

The distortion techniques provide the means of focusing on the detail, without affecting the data overview. It shows a particular portion of the dataset with high degree of detail, and represents an important data presentation paradigm in InfoVis (Keim, 2002). This study presents its review of interaction and distortion techniques using the Spence's (2007) categories which are dynamic projection, interactive filtering, interactive zooming, interactive distortion and interactive brushing and linking

2.3.2.1 Dynamic Projection

The dynamic projection, as an interaction technique, centres on dynamic change of the multidimensional datasets. It helps in showing all the projections randomly, manually and precomputedly (Keim, 2002). Zhao and Zhou (2008) conceptualize projection as navigation. In their work that applied InfoVis framework in e-commerce websites, they explained that navigation will aid e-commerce website users' movement from one information landscape to the other. Zooming in and out is also classified as examples of navigation mechanism, and also is changing visual representation (e.g. changing pie chart to bar chart).

Yi et al. (2007) named dynamic projection "*Encode*". It stated that these are interaction techniques that alter visual appearance by dynamically projecting different view of the dataset. Also, colour, size and shape are core attributes of the data representation that can be manipulated to show different representation. It further assists the users in the understanding of the relationship and the distribution of the selected data items.

In this regard, this study proposes zooming, colour coding and pliant hinting response (push button) as suitable examples of dynamic projection interaction techniques. The zooming will aid magnification and reduction of data view, the colour coding will present distinct representation using colour difference, and push button will allow the user to change data view (Spence, 2007; Cooper et al., 2014). These functions are argued to be necessary users' control mechanisms in handling students' data multidimensionality.

2.3.2.2 Interactive Filtering

Browsing and querying are two typical methods of interactively filtering datasets. They are used in segmenting the data sets into subsets, either through direct selection of the desired set (browsing), or specifying the attributes of the desired set (filtering). Due to the inherent limitations of browsing and querying, studies have been conducted to improve these interactive filtering techniques (Fishkin et al., 1995; Bier et al., 1993) using magnifying and filtering lens.

Visual querying mechanism is also an example of interactive filtering. It specifies the parameters of the data sets which users can visually use to query the large data set, and the interested view will be displayed (Zhao & Zhou, 2008). Fekete and Plaisant (2002) posited that dynamic query as an example of interactive filtering does increase the effectiveness of visualization techniques because users can selectively interact with their desired subset within the large datasets for exploration (Yi et al., 2007).

These techniques allow users to change the data items being viewed by specifying a condition or range. These are necessary mechanisms of an InfoVis system because the screen space cannot present the voluminous data sets at a glance, and inability to handle data multidimensionality causes cognitive load. In this study, drop-down menu (placed at the x and y axes of the visualization span) and check boxes are proposed as dynamic filtering techniques. These techniques are simple in usability and equally address the users' goals preferences.

2.3.2.3 Interactive Zooming

Zooming, a well-known technique deals with both larger display of data objects, and its detailed presentation with higher zooming (Keim, 2002). When dealing with large amount of data, the data overview is often needed to be presented in a compressed form and allows display of data variable using different resolution. Yi et al. (2007) captures this technique under *Abstract/Elaborate*. Abstract/Elaborate techniques are the techniques that aid the users' ability to adjust the degree of abstraction or/and detail of a represented data items. It enhances the display of the data overview and the layers of details between them. Zoom in (view of smaller dataset) and zoom out (view of larger dataset) are the most common example of this techniques. The essential difference between interactive zooming and interactive distortion is that the latter does not alter the represented data view, but the former does. This study proposes interactive zooming as one of its distortion techniques because it allows users' magnification of the data view, and thus aids presentation of an elaborate view.

2.3.2.4 Interactive Distortion

This technique is mostly applicable for data exploration process whereby the overview of the data is maintained during the drill-down operations. In this situation, a certain part of the data is shown with high level of detail, while others with low level (Sedig et al., 2012). Hyperbolic and spherical distortions are popular distortions techniques used on hierarchical and graphical visualizations (Keim, 2002). On another hand, Thomas and Cook (2005) classified zooming together with magic lenses, pop up dialogs and panning as techniques for interacting with a representation, focusing on a section with a drill down to view it in detail.

Drilling down, as explained by Sedig et al. (2012), is to bring out the information set that are not encoded. From the explanation given by Yi et al. (2007), '*Select:* mark something as interesting' is an interaction technique that provides the users with the ability to mark (distort) a data item as the interested one among the many data items presented. It avails the users the opportunity to visually distinct the interesting data item. *Reconfigure* is another form of interaction/distortion technique (Yi et al. 2007), that aligns with Keim's (2002) interactive distortion. It allows users to change (distort) the perspective of the datasets through the change of spatial arrangement of its representation. These techniques essentially reveal the hidden characteristics of data and their relationship. As used in this study, zooming and drop-down menu, are the proposed distortion techniques. Both have been equally discussed under interactive filtering.

2.3.2.5 Interactive Linking and Brushing

Linking is connecting or establishing a relationship between datasets (Sedig et al., 2012). Interactive linking and brushing is as a result of the need to complement possible shortcoming in using a single technique due to inherent weaknesses in each of the interaction and distortion techniques. It conceptualizes effective combination of techniques in a way that will be complementary and thus present more information to the viewers. In this situation, colour coding can be combined with dynamic projection, and so on (Keim, 2002). Yi et al. (2007) characterises *'Connect'* as interactive linking and brushing technique. It is extensively used to show the relationship between represented data items. It also shows the hidden items of the data sets that are relevant to the specified actions. This study primarily proposes the grid system within the visualization span to connect data objects and express their relationship across the x and y axes. The combination of interaction techniques like pliant response and dynamic hinting, as earlier proposed by this study, align with features the interactive brushing.

This review of different types of visualization, interaction and distortion techniques is to present a window-view of the necessary analysis of InfoVis techniques. This is required in view of choosing the appropriate ones for the presentation and representation of HEIs' decision makers' explicit knowledge preferences. This supports the second research question of this study, presented in section 1.5 (in Chapter 1). The details of the choices made by this study, in respect to the HEI's students' data and with supporting justifications, are further stated in Chapter 4 of this thesis.

2.4 Review of Past Studies on Visualization Models and Frameworks

Past studies on visualization models and frameworks can be categorised into (a) conceptual explanation of InfoVis features and characterisation (Sheinerderman & Plaisant, 2010; Spence, 2007; Keim, 2002; Ware, 2000; Card et al., 1999), (b) InfoVis expressive process of graphical formation and visual display (Robertson & DeFerrari, 1999; Owen, 1999), (c) data-task and task-algorithm taxonomies (Tory & Moller, 2004; Keim, 2002; Chi, 2000), (d) user centred design models (Meyer, 2012; Simon et al. 2011; Koh et al., 2011; Robinson et al., 2005), and (e) InfoVis design studies (Wang et al., 2012; Simon et al. 2011). The studies on conceptual explanation of InfoVis features highlighted the fundamental charateristics a tool fit to be called InfoVis must possess, while studies on InfoVis expressive process of graphical formation and visual display show how InfoVis enables graphical formation and its visual display. Data-task and task-algorithm studies enumerated the compatibilities of data types, InfoVis task and the respective algorithms. In such studies, reference models, as developers' guide, in the choice of befitting InfoVis technique in respect to the nature of data to be visualized are presented. Studies on user-centred design models proposed process models of actualizing InfoVis that is user and task-centred, while past related InfoVis design studies focused on delivering domain-specific InfoVis, with or without novel process model.

A representative subset of these categories are chosen for proper contextualization in the review presented in sections 2.5.1 to 2.5.5. A summary of InfoVis characterisitics is presented. Robertson and DeFerrari (1999) and Owen (1999) which illustrated the systematic process of graphic formation in InfoVis' visual display are reviewed. Chi (2000) and Keim (2002) are reviewed on task-data taxonomy. User-centred design models that are reviewed are Meyer (2012), Koh et al. (2011), Robinson et al. (2005), and Roth et al. (2010), to provide extensive insight on InfoVis design process model, while a substantial amount of studies on domain-specific InfoVis design is reviewed to express the domain specificity of InfoVis researches and the basis of InfoVis' techniques choices.

2.4.1 Past Studies on InfoVis Conceptual Explanation and Characterization

An InfoVis must have good underlying structure that allows items that are close to one another to be inferred as being similar; self-directing to users who are unfamiliar with the content; useful to users of less understanding about the organization of the InfoVis; must be less cognitively loaded in its method of exploration; and its information must be easily recognizable and interpretative (Spence, 2007; Card et al., 1999; Ware, 2000; Lin, 1997). Information seeking mantra higlights "Overview first, zoom and filter, then details on demand" as the core tasks that InfoVis must be able to execute (Sheinerderman & Plaisant, 2010; Sheneiderman, 1996). Table 2.2 presents the characteristics of InfoVis.

Table 2.2

Summary of Characteristics of InfoVis

References	**InfoVis' Characteristics**
Card et al., 1999	InfoVis, through smart graphical inventions must serve two distinct purposes: communication and cognitive amplification. It must ideally use visual perception to solve logical problem: 'Using vision to think'.
Ware, 2000	It supports discovery, decision making and explanation. It employs science of colour and lighting of the users' interface environment: Lightness, Brightness, Contrast and Constancy, to deliver a visual appealing Information representation. Its process encompasses data gathering, pre-processing and transformation, data manipulation, data exploration, and graphics engineering
Keim, 2002	It visualizes all nature of data (text, algorithm, 2D/3D, multidimensional, software code) and could represent it either through the conventional statistical graphics (line chart, bubble chart, bar chart), or others like glyphs and icons. InfoVis can employ customized information representation graphics for due representation if necessary. It primarily combines interaction, visualization and visual data mining techniques.
Spence, 2007	It employs visual science to aid understanding of what is

	processed preattentively. It underscores all graphical representation of concepts, data and information. It must be capable of inducing insight and amplify the viewers' cognition through a formation of mental model of the represented concept or information. Its interface must permit visual exploration of data.
Sheinerderman & Plaisant, 2010	InfoVis uses interactive visual representation of abstract data within a compact and users' controllable medium. It conforms, thus aid the information seeking mantra: Overview first, zoom and filter, then details on demand. InfoVis must be able to support these seven (7) users' tasks: overview, zoom, filter, details-on-demand, relate, history, and extract.

InfoVis characteristics as highlighted above can be summarized into three distinct but intertwined features. InfoVis must (a) be able to explore (in view of detecting previously unknown information), (b) interactive (by allowing users to relate, zoom, filter), and (c) graphically communicative (through the use of visualizations). InfoVis techniques are thus implemented in InfoVis design to ensure the InfoVis tools meet the stated characteristics. The implication of this characterization, as presented by the studies, is setting a benchmark in checking any tool fit to be called InfoVis. The characterisation aids developers' understanding in the integration of features and functionalities during the InfoVis design and development process.

2.4.2 Past Studies on InfoVis Expressive Process of Graphical Formation and Visual Display

InfoVis expressive process of graphical formation and visual display are presented by Robertson and DeFerrari (1999) and Owen (1999). The studies attended to the lack of visualization reference model to aid designing InfoVis in a systematic manner and support its visualization and data exploratory functionalities. Robertson and DeFerrari's (1999) reference model is a generic model that shows the abstract description of visualization process, describes its components based on paradigms and data to represent. It describes six components of a visualization system as data model, visualization specification, visualization representation, matching procedure and interaction. The data model is expected to support multiple datasets which can either be dataset with number of variables, or variables with certain numbers of dimensions. The visualization specification must support users' control mechanism in deciding which part of the data is to be displayed, while visualization representation is the quantification of the visual display. The matching procedure ensures the data is described according to the data model, users' control, and with proper quantification. Visualization must support graphical representation, i.e. visual display, enables changing parameter without changing the visualization system, i.e. interaction. Figure 2.7 presents the Robertson and DeFerrari's (1999) integrated visualization model.

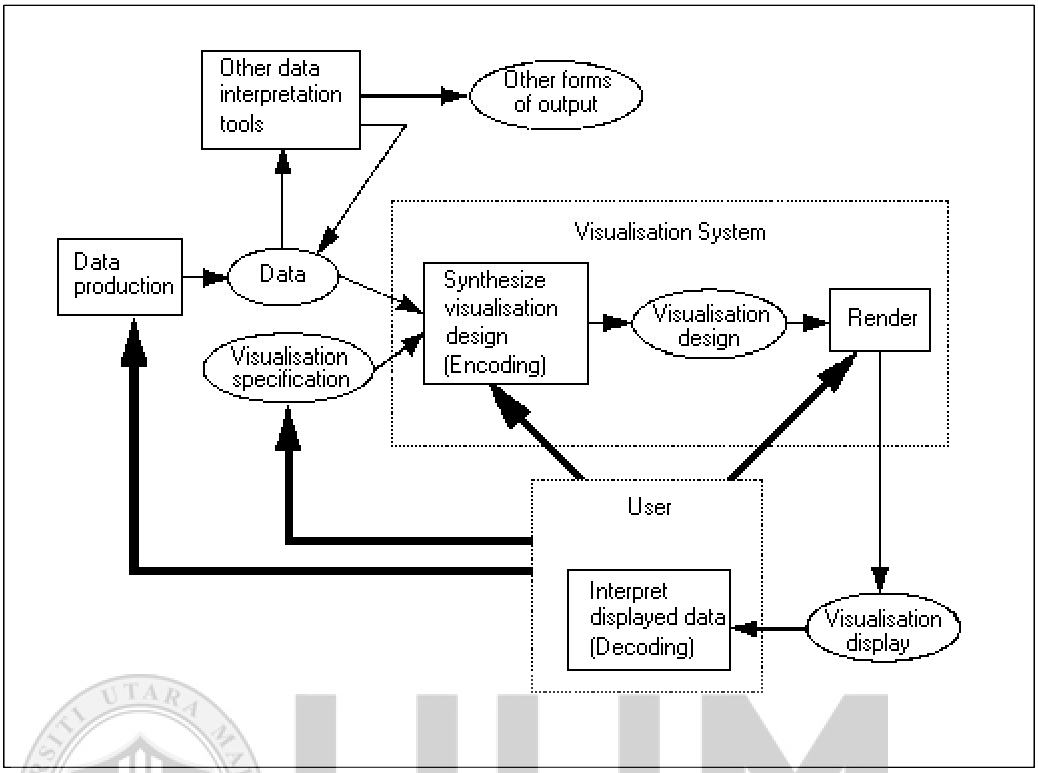


Figure 2.7. An Integrated Visualization Model

(Source: Robertson & DeFerrari, 1999)

On another hand, Owen's (1999) model of visualization suggested a linkage between the visualization systems and the mental model of the investigation and experiments under review. It posited that visualization should help in forming hypothesis during experiment and generating insights. This characterization of visualization is same as described InfoVis features of Sheinerderman and Plaisant (2010), Spence (2007), Keim (2002), Ware (2000), and Card et al. (1999). As an extension to visualization characterization, Owen's (1999) model of visualization presents a set of abstract modules and data types for visualization. The modules are (a) user module which allows the user to observe the system output and control the system; (b) the user interface module which maps the users' world with the internals of the visualization system; and (c) the control module which allows the users to control the data and transform it into their preferences. All types of data, such as text, graphics, user input (e.g. clicks), stored and retrievable data, are expected to be supported by the visualization system. The user interface of the visualization system displays the visual output and accepts users' control as an input. Figure 2.8 presents the Owen's (1999) model of visualization.

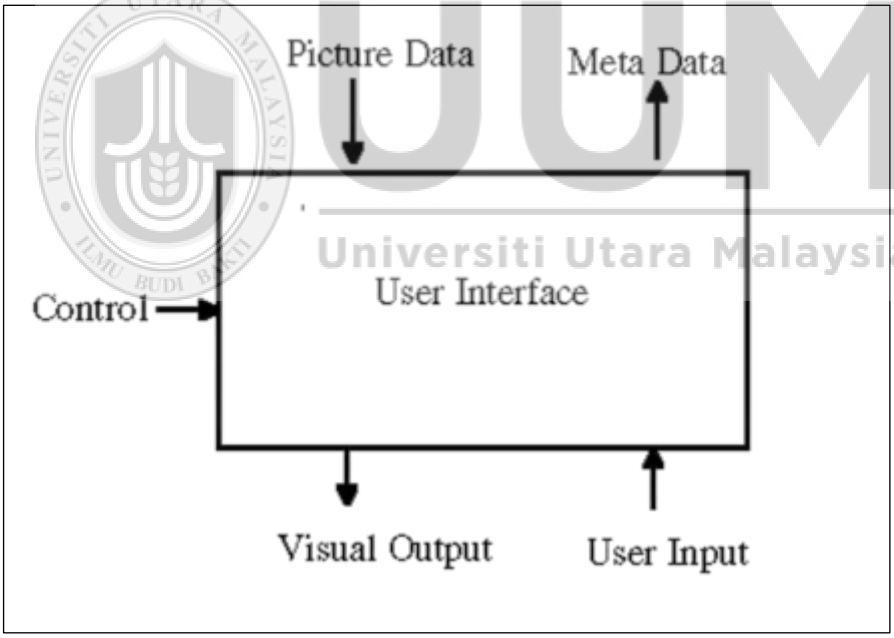


Figure 2.8. Model of Visualization

(Source: Owen, 1999)

Both Robertson and DeFerrari (1999) and Owen (1999) clearly described the process of visualization and formation of the visual display. Data representation and presentation and users' control mechanism are supported by both studies. Though, Owen (1999) diagrammatic representation of the visualization concept is largely a sub-diagram, Robertson and DeFerrari's (1999) is singly comprehensive. But, both studies showed the relationship between the data input, data transformation, graphical display and the users' ability to explore and distort the data representation through the graphical user interface. These studies gave a foundational understanding of structural components of InfoVis and InfoVis' functionalities.

2.4.3 Past Studies on InfoVis Data-Task and Data-Algorithm Taxonomies

Data-task and task-algorithm studies enumerated the compatibility of data types, and the InfoVis task with the respective InfoVis techniques and visualization algorithms. In such studies, the proposed models are developers' guides in the choice of befitting InfoVis technique in respect to the nature of data to be visualized, and in adhering to fundamental InfoVis functionalities. Keim (2002) emphasized that the three dimensions of the InfoVis techniques' classification are assumed to be orthogonal. This means that any of the visualization techniques can be used together with any of the interaction technique, and also with the distortion technique for any type of data. It is the domain specific need that will further determine the choice and the combination of the visualization and interaction techniques. Figure 2.9 shows the classification of information visualization techniques, illustrating the data to be visualized, interaction and distortion technique, and the visualization techniques.

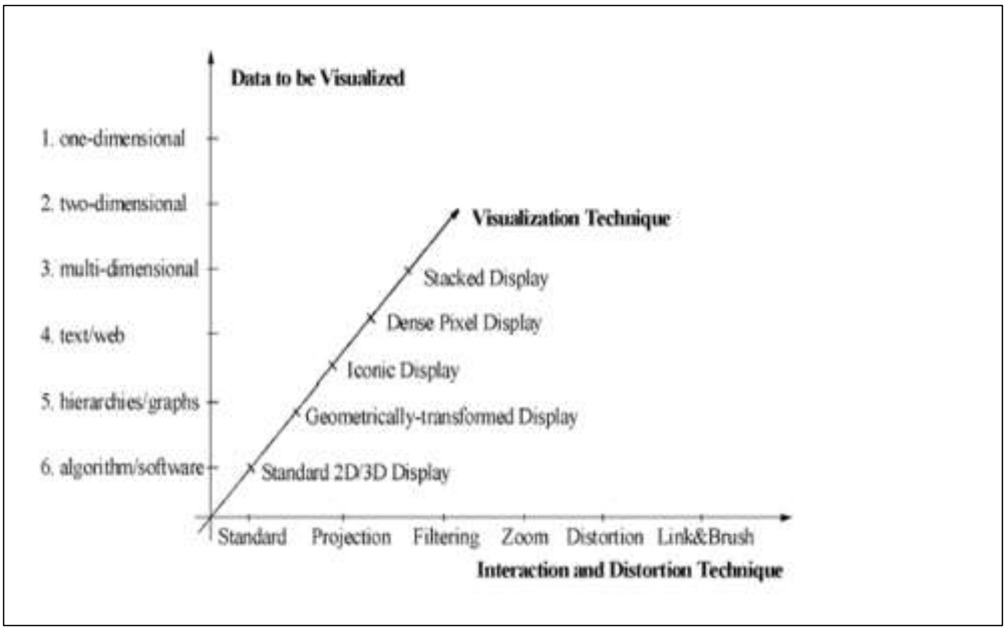


Figure 2.9. Classification of Information Visualization Techniques

(Source: Keim, 2002)

Chi (2000) presented a detailed analysis of visualization techniques through its Data State Model. The model shows the linkage in the nature of the data and the appropriate visualization techniques for a required design space. The taxonomy also included the operational steps in the InfoVis technique abstraction and design. It described the data domains in respect to the possible operations and how this can be applied to popular visualization techniques. It grouped techniques into data domains, and for each analysed visualization techniques, it classified the implementation process for different operations in designing visualization systems. The Chi's (2000) Data State model is a valuable tool in the analysis of visualization design space because of its explicit illustration of the dependencies between the visualization modules and the similarities and differences among the InfoVis techniques. The InfoVis Data State Reference Model, presented in Figure 2.10, is the illustration of the operational process which complements the spreadsheet reference of data-task taxonomy.

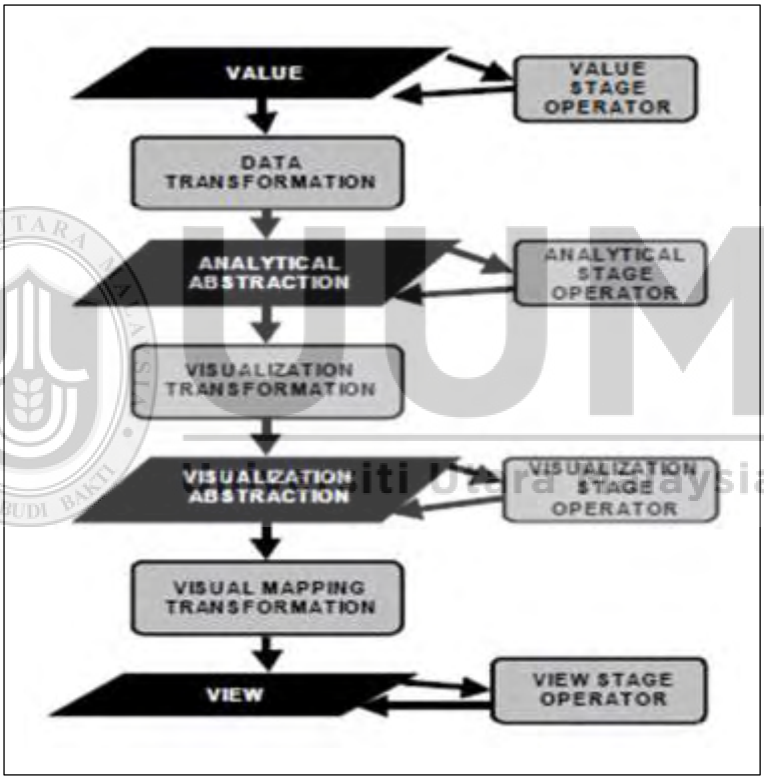


Figure 2.10. Information Visualization Data State Reference Model

(Source: Chi, 2000)

Both studies (Keim, 2002; Chi, 2000) presented taxomonies of visualization techniques, but with different biases and varying depths. Keim (2002) establishes the orthogonal framework which suggests that each of the interaction techniques can be used with any of the visualization and/or distortion techniques. Chi (2000), however, presented a thorough domain data reference model which suggests that data type is an important factor in the choice of the visualization technique. It further presented a detailed analysis that shows the guide in the visualization modules and matching of the InfoVis techniques.

2.4.4 Past Studies on InfoVis User Centred Design Models

InfoVis user-centred design (UCD) studies, with proposed UCD models, are Meyer (2012), Koh et al. (2011), Robinson et al. (2005), and Roth et al. (2010). These studies proposed phases and stages of developing InfoVis to meet domain users' needs and attend to its required tasks. The central value proposition of the models proposed by these studies, though with varying flows, is that target users must be involved in the development lifecycle of the InfoVis tools. This is consistent with the norm of user-centered design generally (Sheneiderman & Plaisant, 2010; Nielsen, 1993). The new probing attention is because InfoVis' must be able to externalize data through graphical display, and internalize same through creation of mental models in the users (Ware, 2000).

Meyer (2012), who principally worked on visualization of biological data, presented an iterative three-phase and nine-stage design framework for designing visualization.

The first phase called Precondition is of three stages. These are: learn, winnow and cast. The second phase called Core consists discovers, design, implement and deploy as the stages involved. The last stage; Analysis, is of two stages. These are: Reflects and Writes. The InfoVis design framework acknowledges the problem-driven and technique-driven nature of InfoVis research studies, the need for pre-design and collaboration with domain experts. The Precondition phase of the design framework is the stage that informs the involvement of collaborators. The collaborators are the experts in the field of study where the InfoVis is sought. Fig. 2.11 represents Meyer's (2012) Nine-stage design study methodology framework.

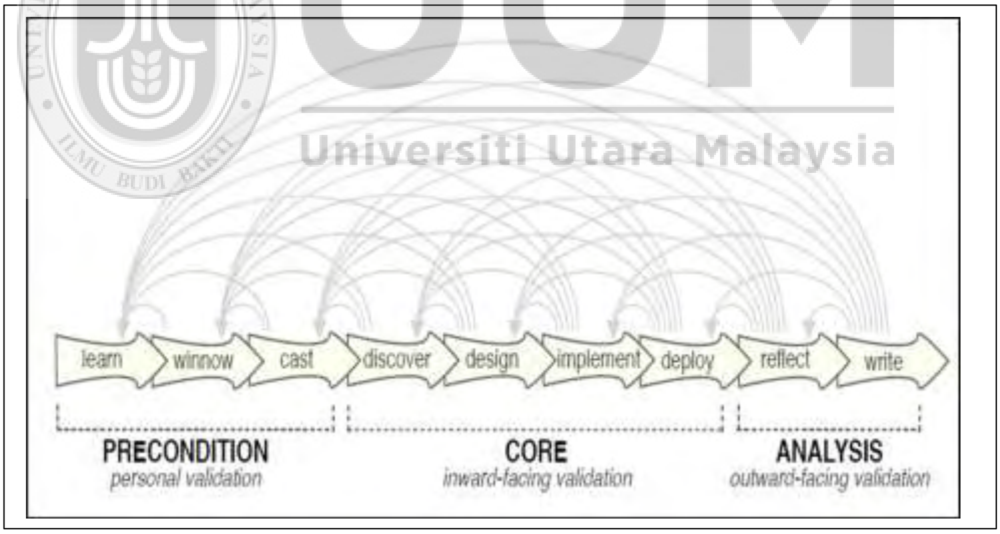


Figure 2.11. Nine-stage Design Study Methodology Framework

(Source: Meyer, 2012)

Robinson et al. (2005) addresses the problem of designing usable InfoVis for the researchers in the field of geovisualization. It is stated that techniques like protocol analysis and in-depth case study must be included in the geovisualization design methodology for a wide range of geographical issues. Interaction design techniques and usability methods are also adapted with extract from the existing design guide to realize the geovisualization-centered InfoVis design guide. The important note of the work is that the diverse application of InfoVis necessitated a domain-specific design guide. The domain specificity is an important step towards the actualization of an InfoVis that will be usable by the domain experts. The domain users are also included in the design process model proposed by the study. Fig. 2.12 presents Robinson et al.'s (2005) user-centered InfoVis design process. An improvement of this model was later made by Roth et al. (2010).

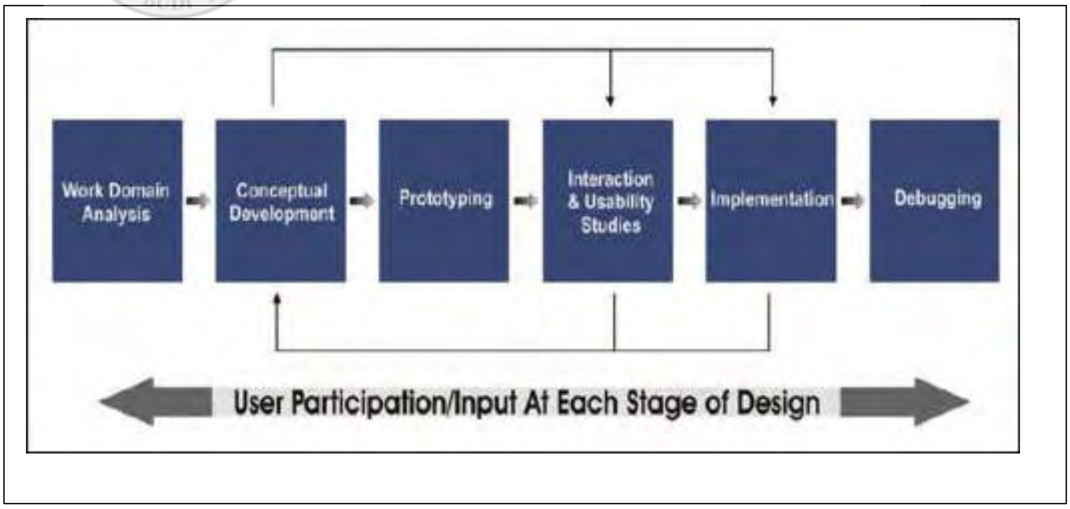


Figure 2.12. InfoVis User-Centred Design Process

(Source: Robinson et al., 2005)

Roth et al. (2010), in improving Robinson et al.'s (2005) InfoVis UCD model, suggested that work domain analysis should not be in isolation. In their improved model, Roth et al. (2010) delivered a user-centered approach for designing and developing spatiotemporal crime analysis InfoVis tool. The work, just as Robinson et al.'s (2005), is geospatial research. However, Roth et al. (2010) suggested that "work domain analysis" stage should be part of the iterative chain of processes. Also, prototyping was placed as a starting point of the design process for the purpose of exposing the users to the InfoVis prototype as the early stage. The prototyping stage as the first phase of the design process is duly integrated into the process model, while debugging remains the last and isolated stage. Fig. 2.13 represents Roth et al. (2010) InfoVis user-centered design approach.

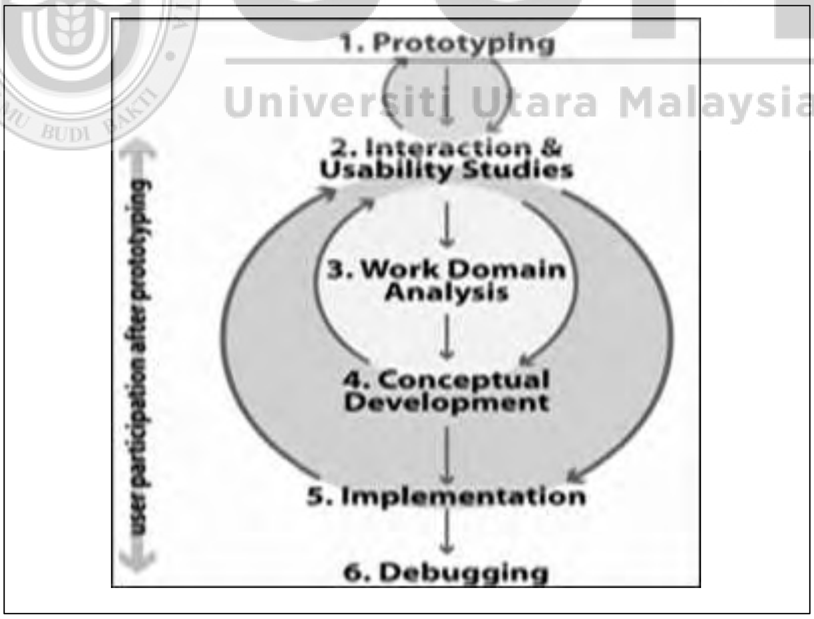


Figure 2.13. User-Centred Design Approach for InfoVis

(Source: Roth et al., 2010)

Koh et al. (2011) in their work on health care-based InfoVis tool also proposed a user-centered model that targets a larger users' community. Their design model features visualization awareness and domain visualization as stages to expose the novices among the prospective users to InfoVis. It maintains the procedural sequence in Robinson et al. (2005), but added two new stages to its start, namely: Visualization awareness and Domain visualization. The introduction of these two stages aligns with the suggestion of Roth et al. (2010). The visualization awareness is a workshop-like session designed to introduce the general concepts of InfoVis to the users, and domain visualization expresses the need to showcase paper prototypes to the domain experts and prospective users. Fig. 4 presents Koh et al.'s (2011) InfoVis users-centered design model.

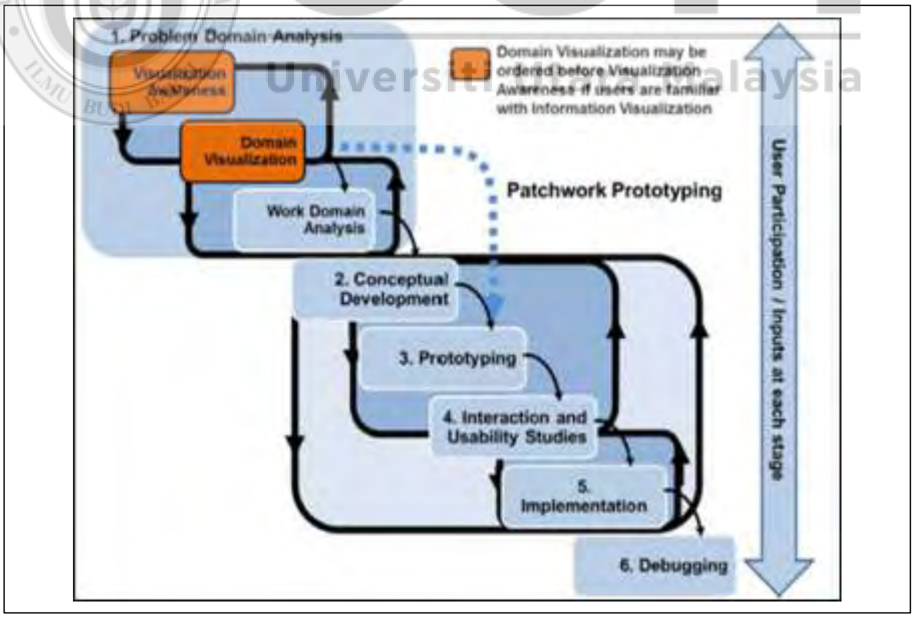


Figure 2.14. InfoVis User-centred Design Model

(Source: Koh et al., 2011)

This study observes that none of the previous studies on user-centred design models for InfoVis elaborated the process of prototyping itself. This limits the ability of developers, especially neophytes in InfoVis design, to translate conceptual design framework into physical artefacts. This study therefore suggests a stage-by-stage approach of translating the InfoVis conceptual design models to physical artifact. The further details of this study's proposed process model, which is an integral part of the proposed conceptual design framework, is presented in Chapter 4, section 4.5.

2.4.5 Past Related InfoVis Design Studies

There are rare studies that comprehensively conceptualized their proposed InfoVis techniques as design frameworks, even though the content of their works partly aligned with InfoVis design framework as presented in this study. The past studies related to designing InfoVis framework to support multidimensional datasets, with application in HEIs' students' data, can be discussed in two folds. These are past studies on diverse applications and adoptions of InfoVis and past studies on InfoVis framework design. The continuous need to transform multidimensional datasets for decision making effectiveness, and presentation of generic or specified design process are responsible for this experience (Rogers et al., 2011; Sheneiderman & Plaisant, 2010; Spence, 2007).

InfoVis had been applied to biological and genomic data (Meyer, 2012; Simon et al. 2011), where Meyer proposed a novel visualization technique for the exploration of biological genes, and Simon et al. was on visual analysis of bacterial genomes. Both works presented novel visualization techniques that support biologists' decision making process, with difference on the users' preferences. Meyer (2012) was specifically to study how same set of genes facilitate different types of cells-based functions in related species, while Simon et al.'s was to detect overlapping genes in bacterial genomes. This supports the need for domain users' preferences in designing InfoVis.

On product lifecycle management, Wang et al. (2012) designed InfoVis for aircraft production monitoring, while Zheng et al.'s (2011) InfoVis is to support aircraft engineers' decision making process. Lirong et al. (2011) and Tekir et al. (2011) both worked on web information clustering for credit information retrieval and document clustering, respectively. Other recorded InfoVis studies are on policy modelling (Kohlhammer et al., 2010), election prediction (Wanner et al., 2009), customer feedback data analytics (Oelke et al. 2009; Hao et al. 2010), news story analytics (Christina, Julian, & Micheal, 2016; Krstajic, 2012), and story telling (Boo et al., 2015). Keim et al. (2010) and Mansmann et al. (2009) designed novel visualization techniques that support monitoring network traffic and detect network intrusion respectively. Quigley (2002) worked on abstraction of software code, Stoffel et al. (2012) presented a web-based visual query analysis system for analysis and simulation of micro grid energy mix, while Harrison et al.'s (2010) work in InfoVis is in visualizing chemical structure.

The financial market, investment analysis and stock market (Ziegler et al., 2008; Schaefer et al., 2011), community informatics and social networking (Faisal el., 2008; Lee et al., 2005; Bowen, 2013), and geographic and health (Sabri, Hela, & Mounir, 2015; Robinson et al., 2005; Sopa et al., 2012; Johansson et al., 2010; Wu & Ren, 2010; Koh et al., 2011; Ning et al., 2012; Yu et al., 2012) are other domains where InfoVis had been employed. Zhong et al. (2012) and Tory and Staub-French (2008) also designed an interactive dymaxion map and web-based InfoVis interface respectively in conveying research output to professional in the building industries. Also, library and information science (Stoffel et al., 2010), document summarization and analysis (Oelke et al., 2008; Strobelt et al., 2009), institutional participation, ranking and participation monitoring (Pinto et al., 2012) are other areas where InfoVis has been applied. Table 2.3 presents the summary of the past related studies on InfoVis design.

Table 2.3

Summary of Past Related Studies on InfoVis Design

Ref.	**Domain**	**Nature of Data**	**Function of InfoVis**	**Interaction techniques**	**Visualization techniques**	**Theory used**
Robinson et al., (2005)	Geographical Information Science	Multidimensional	To explore geographic health data	Brushing, Selection, classification, and colour scheme	Scatter plot, bivariate map, time series plot, and parallel coordinate plot.	None
			To explore	Linear and	Scatter plots,	None

			geographical data on criminal activity	Composite animation functions, temporal legend, zooming	interactive geographical map	
Sopa et al. (2012)		Geospatial data	To convey research information in an interactive manner	Explore, Selection, View	Interactive geographical map,	None
Johansson et al. (2010)	Climatology	Spatial and Abstract data	To view climatic information for disaster prevention decision	Filtering, Classification, Selection,	Geographical pictorial map	None
Keim et al. (2010); Mansmann et al. (2009)	Network Security	Hierarchical/ Graph	To analyze network traffic and detect intrusion events	Filtering, classification, selection, Colour Scheme, Drag and Drop, Threshold	TreeMap layout, Line chart, grouped line-wise pixel plots	None
Tory and Staub-French (2008)	Museum & Gallery	Temporal data, Discourse	It shows the multi-faceted discourse about a Canadian artist, and	Zooming, Navigating, Touches	TreeMap	None

			allows access to temporal data about the Museum's history.			
Meyer (2012)	Biology/ Bacteriology	Graphs, Table, and Tree	To explore molecular biological data	Selection, Filtering	Spatial encoding, Line chart, Curve map	None
Simon et al. (2011)			To detect overlapping genes in bacterial genomes	Selection, Filtering, Zooming	Dense Pixel Display	None
Sabri, Hela, & Mounir (2015); Koh et al. (2011); Ning et al. (2012)	Public Health Informatics		To view the health history of the patients	Selection, Filtering.	Line graph	None
Faisal el. (2008); Lee et al. (2005);		Historical data	To view the history of community epidemiology and	Sorting, Selection, Translating, Scoping. Filtering	Bar graph, Bubble chart	None

Bowen (2013)			predict			
Sopa et al. (2012)	Telemedicine	Regional data of the common diseases in real time data	To be used for learning about the patients' conditions by the doctors	Selection, Colour Scheming	Bubble Chart	
Pinto et al. (2012)	Higher Education Institutions	Multidimensional	It is used as a classification tool of Higher Education Institutions.	Filtering, Selection, Explore, Boolean Selection, Drag and Drop	Sunburst layout, Multiset bar chart, Ring bar chart	None
			A global multi-ranking tool for HEIs	Filtering, Selection, Explore, Boolean Selection, Drag and Drop	Sunburst layout, Stacked Bar charts, Pie Chart	None
			To compare the Higher Education systems	Explore	Sunburst layout, Multiset bar chart, Ring bar chart	None
Wang et al. (2012)	Aircraft Engineering	Table, Tree, Text	To manage information regarding	Filtering, Selection, Translation,	Stacked TreeMap	None

			Aircraft Production	Zooming		
Lirong et al. (2011)	Commerce: Credit Information	Text	To crawl and analyze credit related web pages	Browse, Zooming, Drag and Drop, Colour Scheming	TreeMap, Dense Pixel display of Radial Graph	None
Oelke et al. (2009)	Online Shop, Ecommerce		To support the analysis of products' review	Selection	Dense Pixel Display. Colour Scheming	None
			To display the users' products query and relate their inter-relationship	Zooming, Hyperlinks, Navigation, Brushing	Radial display, Icons, Tree structure	None
Schaefer et al. (2011)	Banking	Multidimensional	To detect fraud in mortgage accounts	Selection, Filtering	Line Chart, Colour Scheme	None
Hao et al. (2010)	Marketing: Customers' feedback	Text: Geo-temporal	To understand customers' feedback on products' survey	Navigation, Selection	Pixel Display, Geo-Map	None
Kohlhammer	Governance and Policy	Text	To analyze opinions	Filtering, Selection,	TreeMap, Hierarchical	None

et al. (2010)	Modelling		and interactively simulate policy decisions	Analysis		
Wanner et al., 2009			To analyze public sentiment on any public issue like election, opinion polls etc.	Filtering, Zooming, Similarity Search	Shape, Colouring, Opacity	None
Faisal el. (2008)	Community of Practice: Academics	Text, Numerals	To visualize InfoVis conference proceedings	Filtering, Selection, Colour Scheming	Dense Pixel, Bar chart	None
Christina, Julian, &	News Reporting	Text	To analyze sentiment in News feed	Selection, Filtering	Bar chart, Colour Scheming	None
Micheal (2016); Boo et al (2015);			To analyze sentiment in News feed	Zooming, Details on demand, Similarity search, Filtering	Colour Scheming, Geometrical shape, Lines	None
Krstajic (2012)			To show the temporal	Selection, Filtering, Zooming	Geometrical object display	None

			characteristics of news story			
Tekir et al. (2011)	Library and Archives	Text, Image	To retrieve and visualize books, images, and archive	Selection	Colour Scheming	None
Ziegler et al. (2008); Schaefer et al. (2011)	Financial Market	Multidimensional Time-Series Data	To visualize financial data and support analysts	Zooming, Selection, Filtering	Dense Pixel Display, Colour mapping, Line	None
Stoffel et al. (2012)	Grid Energy	Time series Data	To analyze and visualize energy simulation result.	Selection, Filtering, Searching	Line Chart, Pixel display, Area chart	None
Quigley (2002)	Software Development	Software code in Hierarchical compound graph	To visualize relational information representing large software views.	Perspective zooming, Explore, Multi-level views	Graph model, Glyphs	None

This review (as presented in Table 2.3) shows the diverse application of InfoVis and the essence of its domain-specific design approach. In each of these studies where InfoVis is designed to support the decision making process of the respective domain, the users' explicit knowledge preferences and the nature of the domain data are exclusive determinants for the choice of the visualization, interaction and distortion techniques. Apart from the domain users' explicit knowledge preferences –which are usually distinct, review of these past related studies show that a pattern of usage in the visualization techniques. For instance, geographical map are commonly used by studies in the domains of climatology, community health and geographical information science (Sopa et al., 2012; Johansson et al., 2010; Robinson et al., 2005). Line chart for times series domains like finances, sales and marketing (Ziegler et al., 2008; Schaefer et al., 2011), dense pixel for biological data (Meyer, 2012), parallel coordinate for software code and network security (Quigley, 2002), among others. In these instances, the nature of the data (temporal, spatial, text, code) determines the appropriate visualization. The interaction and distortion techniques are chosen considering the domain tasks and the designers' discretion to achieve users' experience. The review also showed that user-centred design models for InfoVis are not being linked with theories to build underlying theoretical understanding that enriches scholarly researches.

2.5 Designing InfoVis from Theories

Designing InfoVis from theories has been one of the challenges of InfoVis researches, especially applying theoretical dictates in the evaluation of InfoVis tools

(Purchase et al., 2008). InfoVis design and research is growing rapidly and widely, but can still be comparatively described as nascent (Liu & Stasko, 2010). Therefore, research works are encouraged in developing and contributing to its theoretical base.

Previous works that explained InfoVis from theoretical perspective and contributed to its theoretical base are Purchase et al. (2008), Liu and Stasko (2010), Crapo et al. (2000), Yi (2010), Liu et al. (2007), Teets et al. (2010), Li et al. (2011), Benoit (2012), and Graham (2005). Benoit's (2012) explanation of image-driven explorative information system used Relevance theory, while Li et al. (2011) explored behavioural theories in InfoVis.

Also, Teets et al. (2010) evaluated InfoVis' effectiveness in decision making using cognitive fit theory. Others are Liu et al.'s (2007) proposal of distributed cognition as a theoretical framework for InfoVis' evaluation, Yi's (2010) proposition of building InfoVis theory by adopting a two-stage development from the historical background of decision science, and Crapo et al.'s (2000) perception of using cognitive theories to assess computational modelling of visualization. The generic works that endeavoured to explain how InfoVis' theories can be developed or adapted are Purchase et al. (2008), Graham (2005), and Liu and Stasko (2010). The essence of this is to create a platform through which future researches contribute to InfoVis' theoretical base.

In the light of the above, this study proposes that InfoVis design should encompass the three characteristics of theories found in the past related literatures. These past related literatures are works on perceptual theories (Crapo et al., 2000), cognitive theories (Crapo et al., 2000; Teets et al., 2010; Liu et al., 2007), and decision making theories (Yi, 2010). The perceptual and cognitive theories describe how perception and cognition can be achieved through visualization and interaction, respectively. The decision making theories, traditionally from decision science, describe how quantification of data can affect decision making. Yi (2010) proposed the adaptation of decision science theories into InfoVis design based on the argument that an InfoVis is a graphical display of quantified data.

The perceptual theories are preattentive processing theory (Triesman & Gormican, 1988; Triesman, 1985), similarity theory (Quinlan & Humphrey, 1987), guided search theory (Wolfe & Cave, 1989), information theory (Ware, 2000) and gestalt theory (Fekete, van Wijk, Stasko & North, 2008). The cognitive theories are distribution cognition framework (Hutchins, 1995) and cognitive fit theory (Teets, Tegarden, & Russell, 2010). And, the decision making theories are normative and descriptive theories (Yi, 2010). This study proposes a theoretical guide towards designing InfoVis using preattentive theory, cognitive fit theory, and normative and descriptive theories as basis for actualizing the perceptual, cognitive, and decision supports of InfoVis respectively.

2.5.1 Preattentive Processing Theory

Preattentive processing theory is chosen among the perceptual theories because it best describes the preattentive processing phenomenon which is primary in InfoVis' display of distinct graphics for data presentation. Its extension addresses feature integration, and Texton, thus makes it all contributing when compared to others like similarity theory, guided search theory and gestalt theory. Preattentive processing theory was firstly introduced by Triesman (1985), and later extended by Triesman and Gormican (1988). It basically explains the working mechanism of the human visual system, and shows how visual properties are preattentively processed.

By preattentive processing, it means that features of visual design are easily identified and processed without any conscious guide. For instance, in the process of detecting a filled circle among empty circles (Figure 2.15), or a red circle among blue circles (Figure 2.18), preattentive processing is enabled because the target visual feature is distinct. In each of these cases, a viewer can, at glance, note if the target is present or absent.

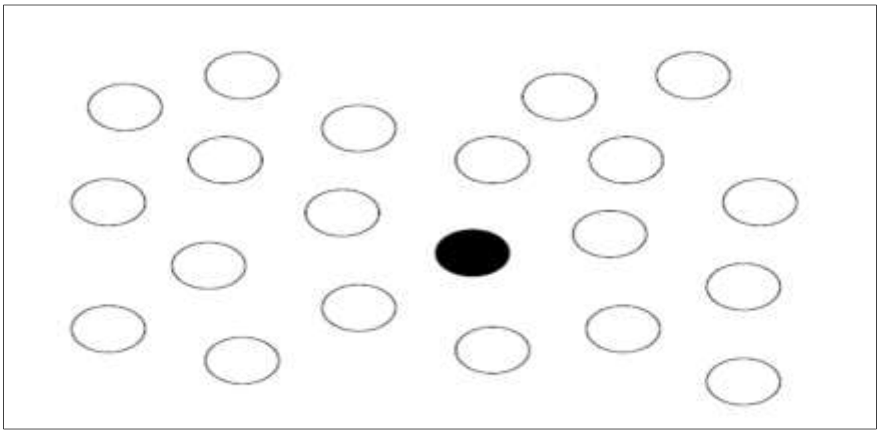


Figure 2.15. Filled circle preattentively processed because of its unique feature

(Source: Healey, 1996)

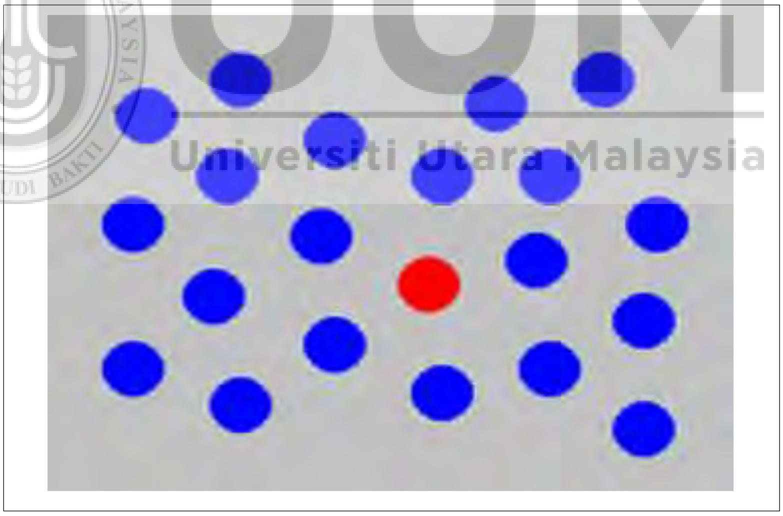


Figure 2.16. Red circle preattentively processed because of its unique feature

(Source: Fekete et al., 2008)

Feature integration theory and texton theory are additional theories introduced by Triesman and Gormican (1988) to strengthen preattentive processing theory. Feature integration theory states that when a target is of unique feature, it prompts further users' interaction. This shows the value of integrated features. Texton theory discusses texture segregation that involves preattentive location of groups of similar objects with the boundaries that separate them. Figure 2.17 shows example of similar textons. In (a), the two textons appear different in isolation, but in (b), these same two textons cannot be differentiated in a randomly oriented texture environment.

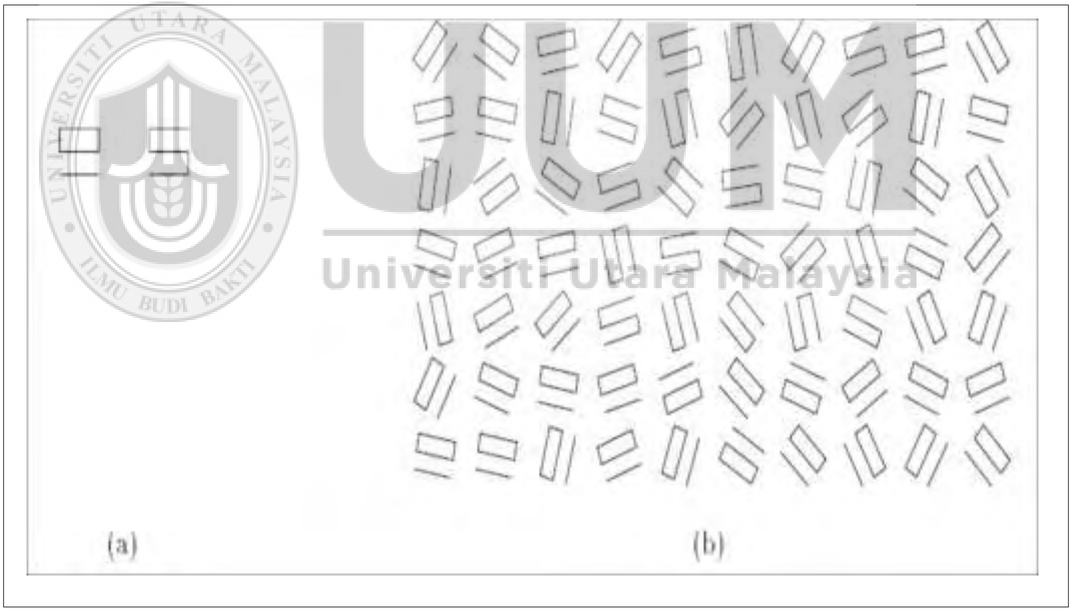


Figure 2.17. Examples of similar textons

(Source: Healey, 1996)

The preattentive processing theory illustrates the expected quality of any visual representation that will aid perceptual support as expected of an InfoVis. It is at such high processing level, when perceptual sense is made of InfoVis, that cognition is driven.

2.5.2 Cognitive Fit Theory

Few studies, as observed by Liu and Stasko (2010b), have examined the interplay between the external representation of data using visualizations and its corresponding internal representation of mental models within the human's cognitive faculties. Cognitive fit theory adequately describes this phenomenon. It is one of the problem-solving theories that suggested that interaction between the problem solving tasks (users' interaction and tasks) and the problem presentation (visualization) results in amplification of cognition (Teets et al., 2010). In essence, cognitive fit theory suggests that mental representation of a problem will only be realised when the problem representation fits the problem-solving task.

Speed and accuracy of the problem solving process can be achieved by the compatibility of the users' interaction and subsequent internalization of the data displayed through the infographics. Cognitive fit theory also expands extensively to decision support from its tasks classification – symbolic and spatial, and exploration to reduce users' cognitive load (Teets et al., 2010).

2.5.3 Normative and Descriptive Theories

Yi (2010) suggested a two stage developmental model from the history of decision science with normative and descriptive theories as constituents. Normative theory is a body of theories containing several axioms helpful in formulating mathematical modelling (e.g. dimensional relationship in x and y axes) in understanding a particular domain dataset. Decision support systems are generally referred to as normative systems. Theories that are applicable in decision science are probability theory, game theory, fuzzy logic and utility function (Yi, 2010; Druzdel & Flynn, 2002). Decision theory is a conjunctions of axioms of rational decision making, expresses uncertainty using probabilities and utilities' preferences. The descriptive theories are products of close observation of human needs and abilities. They are mostly designed to augment the social unrealistic formalization of the normative.

The graphical display of discrete data value and the modelling relationship of data objects in x and y axes of the visualization span are applications of normative theory. The relationship of the data objects are expression of the domain's explicit knowledge preferences which are actual observation of the domain users' needs. This describes how descriptive theory in decision science helps in formalization of the normative to attend to the domain specific needs.

2.5.4 Implication of the Theoretical Background for this Study

The implication of this theoretical explanation and deductions is both in the design and evaluation of the InfoVis prototype delivered by this study. This approach of theoretical application is supported by some of the past related studies (Yi, 2000; Graham, 2005; Benoit, 2012; Li et al., 2011; Crapo et al., 2000; Teets et al., 2010). Each of the theories discussed in this study characterizes a particular support from the three supports that InfoVis must be designed for.

From the evaluation perspective, Yi (2000), Teets et al. (2010) and Benoit (2012) have suggested the need for further studies on theory-based evaluation of InfoVis. The theoretical explanation shows that there is an interconnection between cognition and perception as rendered by InfoVis. Also, from the Yi's (2000) explanation of the essence of displaying data values and expressing relationship between data objects to attend to domain users' needs, InfoVis actualizes decision support through perception and cognition. The perception is through data value displayed (i.e. visualizations) in a manner that its information can be preattentively processed, in view of the dictates of the preattentive processing theory. The users' ability to control the external formation of this data, through interaction and distortion techniques, aids gaining of the insight (i.e. cognition). Lastly, since the data explored and visually presented addresses the domain needs, it can be safely used for decision making purposes.

This study therefore argues that since InfoVis is expected to support these intertwined and interdependent processes; perceptual, cognitive and decision-making, a functional InfoVis must support its processes. The process of designing such functional InfoVis is ensuring that the dictates of the theories reflect in the design features of the InfoVis. The theoretical guide, as presented in Table 2.4, highlights the design features that correspond to the InfoVis' techniques, and the attending theories. Its conceptual model is further detailed in chapter 4 of this thesis.

Table 2.4

Theoretical Guide to designing InfoVis Techniques

References	**Theory**	**Category of Support**	**Corresponding InfoVis' Techniques**	**Design Features**
Triesman (1985)	Preattentive processing theory	Perceptual	Visualization	Regular shapes for data representation; distinct color coding for distinctive information
Teets Tegarden, & Russell (2010)	Cognitive fit theory	Cognitive	Interaction and Distortion	Users' control, interactive media, communicative graphics
	Normative theory	Decision making	Visualization and Distortion	Users' control, data objects relationship, data quantification
Yi (2010)	Descriptive theory			Domain-focused data objects

From the presented review, this study attempts a two-pronged contribution. First, this study proposes a theory-based design guide that ensures that InfoVis is designed to support perception, cognition and decision making. Past InfoVis works are analysed for recurring themes and categorized based on the InfoVis' supports to suggest the

InfoVis technique(s). The design features that will drive these supports are respectively listed.

On InfoVis' evaluation, this study argues that past similar works (Benoit, 2012; Benoit, 2012; Yi, 2000) on InfoVis did not consider all the supports expected to be featured in an InfoVis which are perception, cognition and decision supports. The proposed InfoVis' evaluation framework is an extension of the adapted evaluation framework that attended to this limitation. The evaluation framework is further illustrated and explained in Chapter 3 of this thesis.

2.6 Summary of the Chapter

This chapter presents the scholarly foundation for this study. It discusses data-driven decision making process in HEIs, emphasizing the importance of data and its appropriate usage for organizational growth. In this regard, the multidimensionality of HEIs' students' data is highlighted and its associated problem of information overload. It is against the need to solve this problem by proffering appropriate tool for handling students' data multidimensionality that InfoVis, as a data exploratory and visualization tool, is proposed. This chapter thus describes InfoVis techniques which are visualization, interaction and distortion techniques. It highlighted that choice of these techniques, in view of composing the InfoVis framework must be exclusively determined by the domain users' explicit knowledge preferences, the nature of the domain data, and the users' tasks. This is evident in the highlighted limitations of past related studies on visualization models, as shown in the review.

Theories and theoretical models are equally provided to ensure that designed InfoVis tools, at large, address perceptual, cognitive and decision making supports which are expected supports from InfoVis. Finally, the review supports the earlier stated research objectives of this study. The propositions are (a) the categorization of the appropriate HEIs' students' data dimensions must be from the identified explicit knowledge preferences of the domain decision makers. This attends to the first objective stated by this study. (b) The result would subsequently determine the appropriate visualization, interaction, and distortion techniques (this accomplishes second object), and translate to the HEI domain specific conceptual design framework (this accomplished the third objective). The process of achieving these objectives is further discussed in Chapter 3 of this thesis.

CHAPTER THREE
RESEARCH METHODOLOGY

3.1 Introduction

This chapter describes the methodological phases and the research design used in this study. It discusses the research processes in holistic, and presents all the methods used in achieving the stated objectives. This study adopted a popular information system design research methodology, and each of its methods is duly explained. The corresponding steps taken to accomplish each of the phases in the research methodology, as adapted, are also explained.

3.2 Design Research

The choice of design research as the appropriate research methodology is because it aligns with the main objective of this study that is, developing an HEI-focussed InfoVis conceptual design framework. The standard practice is ensuring the chosen methodology guides the study towards the accomplishment of the set objectives. Design research paradigm is suitable for the achievement of a design framework which is the expected outcome of this study. It is a popular methodology in fields like Human Computer Interaction (HCI) (Carroll, 2000; Druin, 2002), information system (Vaishnavi & Kuechler, 2007; Purao, 2002), education research (Barab & Squire, 2004), and instructional design (Regeluth, 2008).

In HCI, several researches have employed design research. Carroll (2000) opined that a design research should be of two goals. First, it is to understand the world through design requirements gathering process. Secondly, it is to improve the world through design process, by creating tools that will solve human problems. The requirements involved in HCI are the needs of the users, and HCI researches are much concerned about the fulfilment of the users' needs (Cooper, Reimann, Cronin & Noessel, 2014). In these stances, a HCI model will be outcome, and understood as the artefact. The model is then validated by an evaluation stage in the design process. The successful evaluation of the design will determine the generalizability or otherwise of the design to a wider coverage of the HCI community.

Design research in information system is centred on problem-solving purposes with attempt to create innovations that define ideas, technical abilities and capacities (Hevner et al., 2004). Information system design research produces (i) models such as abstractions and representation, (ii) methods in terms of algorithms and practices, (iii) constructs in terms of symbols and vocabulary, (iv) instantiations, which are implemented in prototyping. Build and evaluate are two main iterative phases in information system till the final product of the artefact is achieved (Hevner et al., 2004; Vaishnavi & Kuechler, 2007). This summarizes the description of design research, and the basis of its suitability for this study.

3.3 Justifications for Using the Design Research

Design research is suitable for this study because it makes provision for all its needed phases and the expected research deliverables. The following are the justifications:

i. The design research attends to the problem of the artefact design. This study has a conceptual design framework as one of its deliverables. Developing framework/model is one of the suitable purposes for adopting design research (Hevner et al., 2004; Teegavarapu & Summers, 2007; Frankel & Racine, 2010).

ii. The domain of this study (Information system research) suits design research (Hevner et al., 2004).

iii. The design research guidelines as outlined by Hevner et al. (2004), Teegavarapu and Summers (2007) and Frankel and Racine (2010) are relevant to this study, and therefore to be practically utilized. Its structure is flexible, and can perfectly accommodate phases of InfoVis UCD methodologies.

This study is concerned with development of InfoVis conceptual design framework to solve the problem of information overload caused by the multidimensionality of the HEI students' data, and thus aid the related decision making process. The design research is adapted to produce the InfoVis conceptual design framework.

3.4 Phases in the Research Methodology

The design science research methodology conceptualized by Hevner et al. (2004) with its phase-to-phase approach outlined by Vaishnavi & Kuechler, 2007, is adopted for the accomplishment of this study's objectives. In the methodology, the primary focus is on delivering artefacts such as models, methods, or prototypes. Hence, it is mostly adopted by researchers in the information system design.

The research methodology is of five phases, namely; (a) Definition of the Problem, (b) Suggestion, (c) Development, (d) Evaluation, and (e) Conclusion. Figure 3.1 depicts the activities conducted in this study. The overall process focused mainly on the development of an InfoVis conceptual design framework for HEI, with the students' data as the domain data.

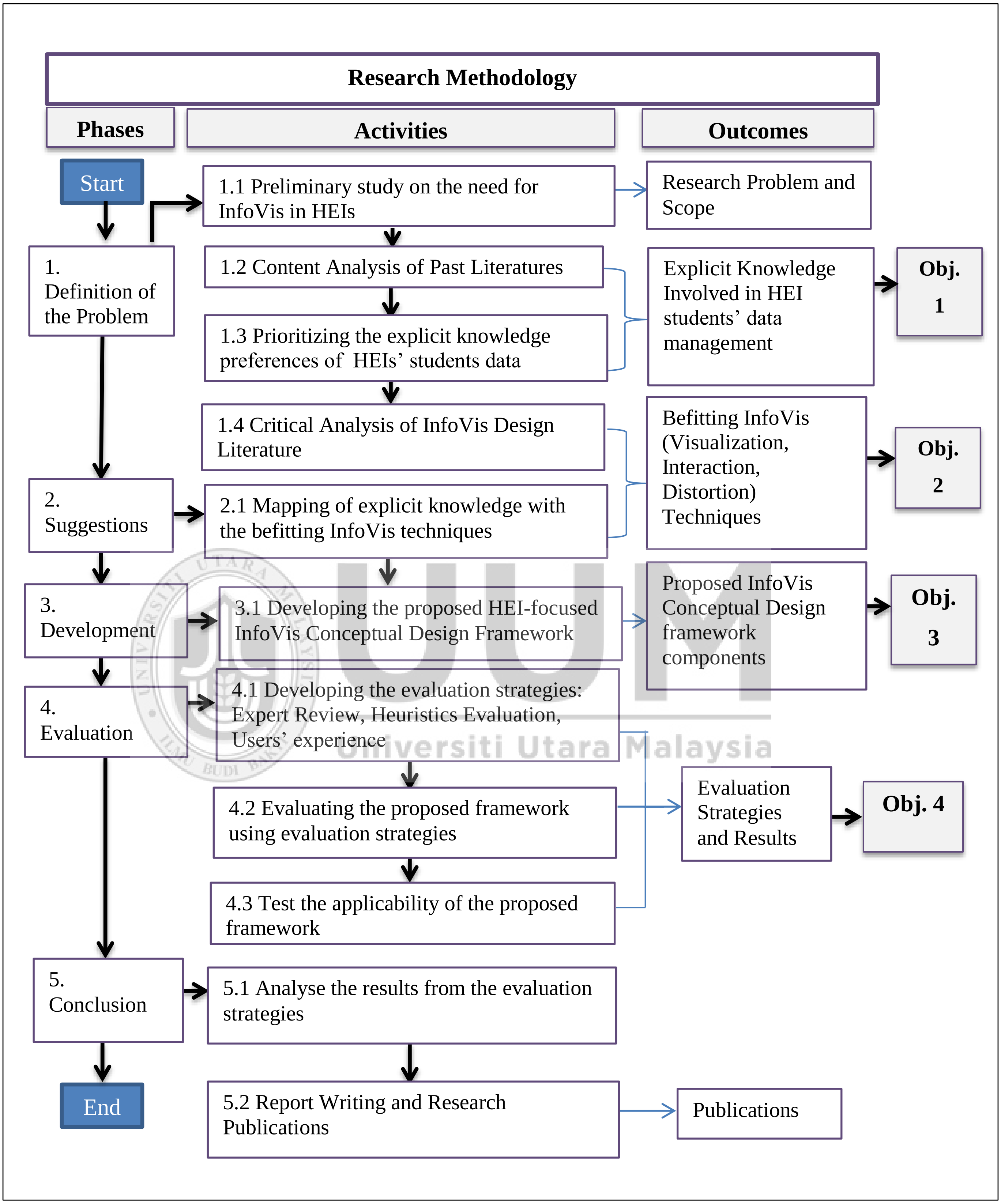


Figure 3.1. Research Phases

3.4.1 Phase 1: Definition of Problem

The first phase of this research described how the problem to be solved by this study is defined through literature review and pre-design (preliminary) study. This is stated in the problem statement. The pre-design study was conducted to validate the elicited explicit knowledge preferences, as found in the literature, and prioritize them in relationship with HEI students' data-related decision making process. The findings are presented in Chapter 4.

3.4.1.1 Content Analysis of Past Literatures

This study employed content analysis of past literature review, as supported by Babbie (2010) and Creswell (2009), to determine the key issues associated with data-driven decision making process in HEIs and the role of InfoVis. The review also helped in delineating the scope of this study. Past works on InfoVis characteristics and domain-specific InfoVis design are broadly and specifically reviewed, critiqued and synthesized to attend to this study's problem statement and the research objectives.

Preattentive processing theory, cognitive fit theory, and normative theory and descriptive theory were also reviewed to present a theoretical guide to the design and evaluation of InfoVis prototype. This ensures that the prototype delivered by this study supports perception, cognition and decision making.

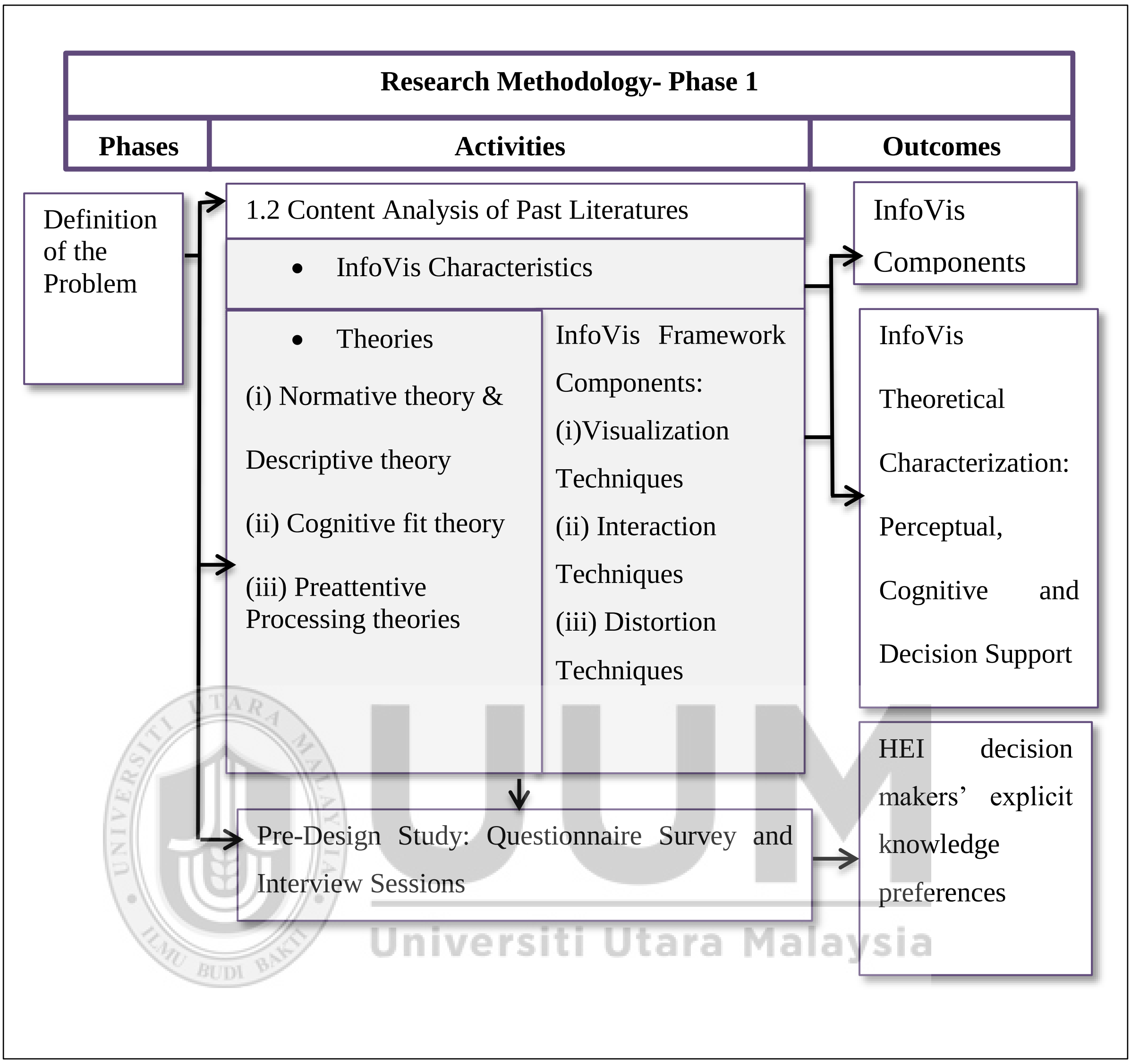


Figure 3.2. Phase 1 –Definition of Problem Phase

3.4.1.2 Preliminary (Pre-Design) Study: Survey and Interview Sessions

A pre-design study was conducted to validate the elicited explicit knowledge preferences collected from the literature review, and prioritize them accordingly. This serves as users' requirement gathering and validation phase in InfoVis design methodologies (Sedlmair, Meyer, & Munzner, 2012; Meyer, 2012). The explicit knowledge preferences are output of transformed dimensional relationship of the domain datasets (Semiu & Zulikha, 2016). It supports the HEIs' students DDDM. The pre-design study was conducted using both survey questionnaire and interview sessions. Basic descriptive analyses and mean values were used in the analysis of the data collected through survey questionnaire. The audio clips of the interview sessions were transcribed and analysed using thematic analysis (Creswell, 2009). The problem statement validation is reported in section 1.4; Chapter 1, of this thesis. The explicit knowledge preferences validation (users' requirements) is reported in Chapter 4 of this thesis.

3.5 Phase 2 & 3: Suggestion and Development

In Figures 3.1 and 3.2, the requisite steps toward suggesting the befitting InfoVis' techniques for this study were taken. These are the critical analysis of InfoVis design literatures, past related InfoVis design studies, and the pre-design study. Based on these, the InfoVis' techniques were suggested and presented in Chapter 4 of this thesis. The choice of the InfoVis techniques was made based on their appropriateness in transforming the HEI students' data to the explicit knowledge preferences of the domain policy makers. Therefore, by mapping the categorized data dimensions of the HEIs' students' and the explicit knowledge preferences of the HEI policy makers with the suitable InfoVis techniques, an InfoVis conceptual design framework was proposed. Figure 3.3 depicts the phase 2 and 3.

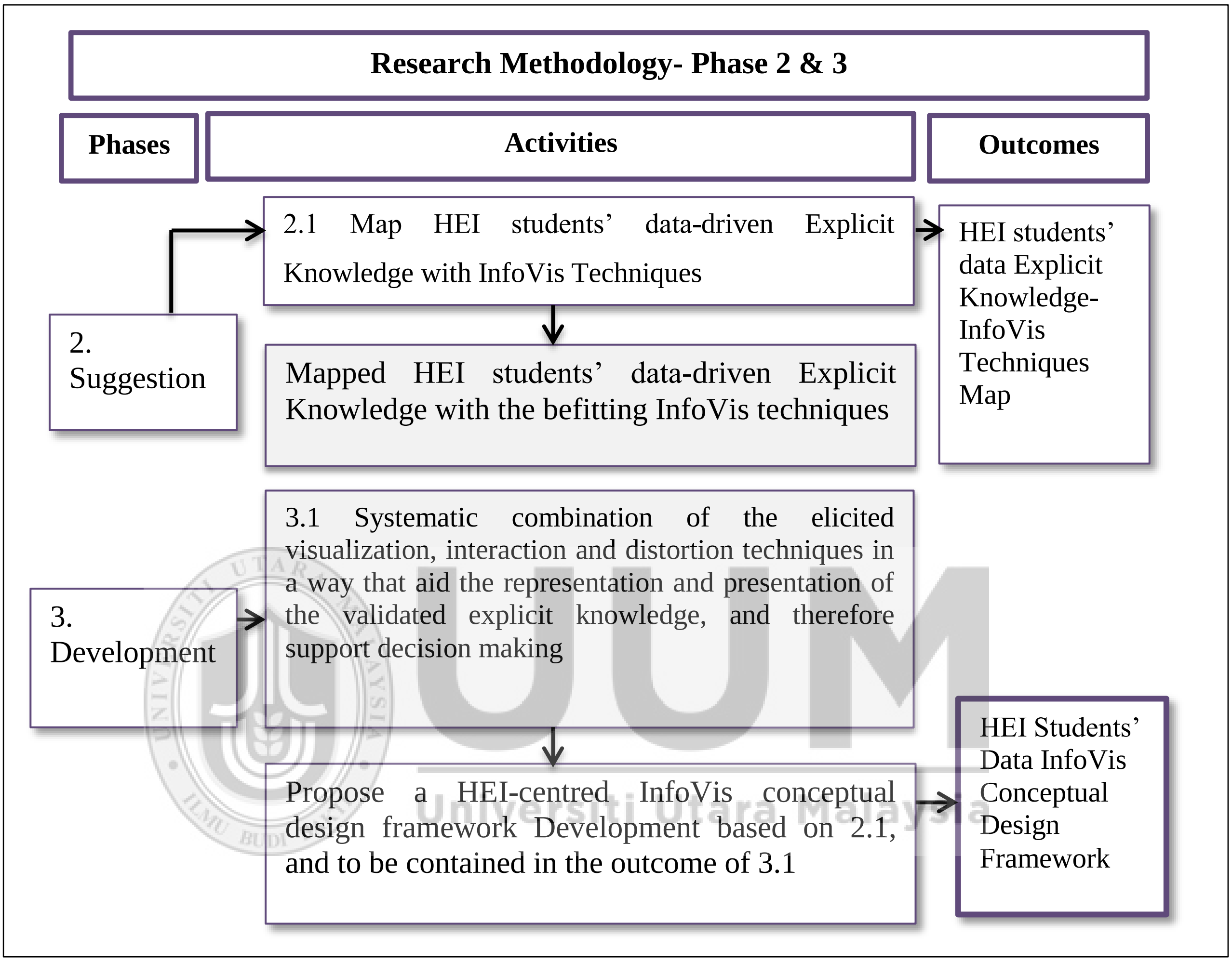


Figure 3.3. Suggestion and Development Phases

3.5.1 Mapping of HEI Students' Data-driven Explicit Knowledge with InfoVis Techniques

The InfoVis techniques are chosen based on the validated HEI students' data-driven explicit knowledge preferences, categorized data dimensions, and the respective statistical graphical displays according to the work domain. The InfoVis techniques are specified by the content delivery techniques model of the proposed InfoVis design conceptual framework. The visualization, interaction and distortion techniques, which are categories of the InfoVis techniques, are to deliver the content of the data exploration, as preferred by HEI administrators and policy makers. This ensures that the InfoVis tool supports their decision making processes.

3.5.2 HEIs' Students' Data-centred InfoVis Conceptual Design Framework Development

The proposed framework was based on the results of the previous phases. At this stage, the interconnection between the validated explicit knowledge preferences, the categorized students' data dimensions and the befitting InfoVis techniques were conceptualized and presented using tables and diagrams. The conceptual design framework is also theoretically supported based on the theoretical guide to designing InfoVis presented in Table 2.3 in Chapter 2 of this thesis. This study also presented structural and process models as supporting components to the conceptual design framework. The process model describes the stages and phases in practical development of InfoVis tool. This is followed in the InfoVis prototyping phase. The structural model highlights the generic structural components of an InfoVis, and the specificity that addresses HEI students' data. The proposed HEIs' students' data-centred InfoVis conceptual design framework is presented in Chapter 4 of this thesis.

3.5.3 Verification of the HEIs' Students-centred InfoVis Conceptual Design Framework Development Using Expert Review

The conceptual design framework was firstly verified using expert review. The expert review is recognised as a way of improving the quality of the developed conceptual design framework (Wiegers, 2002). The experts, as referred to in this study, are academics and industry-based experts in the area of HCI and InfoVis who have prior experience in developing data visualization tool since the conceptual design framework is expected to be used by InfoVis developers.

According to Neilsen (1997), a number of three (3) to five (5) experts are sufficient for expert review. This study uses five (5) experts drawn from both academics and industry, in no particular order of preference. The profile of the experts is presented in Appendix H. The expert review procedure is thus: (a) designing the review instrument based on the chosen metrics for the evaluation (detail in section 3.7.2; Appendix C), (b) conducting the review, and (c) analysing the results. The detail of the expert review session is discussed in Chapter 5.

3.6 Phase 4: Evaluation

The different types of methods usable in evaluating and validating research outcomes are experimentation, demonstration, metric usage, simulation, bench marking, logical reasoning and mathematical proofs (Vaishnavi & Kuechler, 2007; Thatcher et al., 2004). These methods are of different strength, hence, with different appropriateness. It is however emphasised that selection of evaluation method must be based on the research outcome to be evaluated.

The evaluation techniques employed in this study was in two folds –the formative and summative evaluations. Under the formative evaluation, (i) Prototyping and (ii) Heuristic evaluation, as suggested by Sherwood and Rout (1998), Rogers et al. (2010), Tory and Moller (2005), and Shneiderman and Plaisant (2010), were employed. For the summative evaluation, users' experience evaluation study was conducted (Tullis & Albert, 2008).

As indicated by Web Center for Social Research, both evaluation methods are complementary in view of achieving a befitting products or service. The formative evaluation techniques, as used in this study, provided feedback from InfoVis developers and HCI experts. The feedback was employed in improving the InfoVis framework. The summative evaluation is the final stage of evaluating the conceptual design framework through the evaluation of the InfoVis prototype.

3.6.1 Prototyping

Prototyping is the translation of the conceptual design to a physical design. In this study, it is to produce InfoVis system that uses the proposed conceptual design framework as guide. Prototyping has been acknowledged as a way of evaluating the conceptual design in software development (Rogers et al., 2010; Shneiderman & Plaisant, 2010). It aids the implementation of the design guide and development of the users' requirements. During prototyping, software designers could revise the design process to meet the practical reality. In this study, the final physical artefact achieved is a web-based InfoVis tool which can be used in the exploration and visualization of the HEIs' students' data. The successful translation of the proposed conceptual design framework to a working physical artefact validates the applicability of the framework.

The logical assumption in using prototyping as method of evaluation is that, if the physical artefact can achieve the objectives of the conceptual framework, it means the conceptual framework is valid (Rogers et al. 2010; Shneiderman & Plaisant, 2010). In this study, the ability of the prototype designed to support correct data representation, usability and decision support validates the proposed framework. Therefore, the proposed framework would be found applicable in designing InfoVis for students' data and its HEI policy makers. The details of the prototyping phase are reported in chapter 5 of this thesis.

3.6.2 Heuristic Evaluation

The heuristic evaluation was done through the experts' interaction with the prototype using their heuristic experience (Tory & Moller, 2005). The experts for this heuristics evaluation comprised of usability/HCI experts and data administrators/analysts. The choice of these disparate field experts in constituting the heuristics evaluation team is to achieve reliable findings. It is noteworthy that InfoVis functionality intersects between HCI (visualization and interaction) and data administration (decision support and data exploration). Each expert is expected to be more sensitive to his or her respective domains than the conventional users (Tory & Moller, 2005; Forsell, 2012). Nevertheless, this is not to substitute the users' experience evaluation study.

This study employed four (4) experts for the heuristics evaluation. Two were usability/HCI experts, and two were HEI data administrators/analysts. The heuristics evaluation concentrated on correct data representation and usability issues of the InfoVis prototype.

3.6.3 Users' Experience Evaluation Study

The users' experience evaluation of the InfoVis prototype was in terms of its correct data representation, usability and decision support. This is supported by Lazar, Feng and Hochheiser (2010) and Tullis and Albert (2008). It further assisted the validation of the proposed conceptual design framework. Users' experience evaluation study is a real-life and practical assessment of the developed prototype using the specified metrics (Lazar, Feng & Hochheiser, 2010; Luna-Reyes & Andersen, 2002).

The users' evaluation study was electronically conducted. Each of the participants, who are HEI decision and policy makers, is given access to the web-based InfoVis prototype through its uniform resource locator (URL) link. The briefing to guide the users is done through phone and PC call using *Skype*. Users' experience evaluation study is generally regarded as summative evaluation method in HCI and related fields (Rogers et al., 2010; Shneiderman & Plaisant, 2010; Tullis & Albert, 2008; Cooper, Reimann, Cronin & Noessel, 2014).

The option of conducting the users' experience evaluation study electronically is adopted due to the near-impossibility of having the HEI policy makers drafted into the traditional experimental laboratory. Apart from this circumstantial justification, literature also suggests that some of the traditional data collection approaches have been disrupted by internet technology (Denscomber, 2014; Carrasco, 2006). Denscomber (2014) supported virtual ethnography. Also, Carrasco et al. (2006) supported the validity of collecting social media users' data through the social networking sites for social activity research. This study suggests that electronic users' evaluation study can be explored, especially in circumstances of this study, and as justified by the experience of disruptive technology in data collection.

The (electronic) evaluation instrument for the users' evaluation study has both closed ended and open ended questions. The close ended questions are items to evaluate correct data representation, usability and decision support of the InfoVis. The responses to the open ended questions are analysed using Elo and Kyngas' (2008) content analysis approach for qualitative data. The detail of this approach is presented in Chapter 6 of this thesis (see section 6.3). The responses to the closed ended items were analysed using Statistical Package for Social Sciences (SPSS, version 19) (Pallant, 2011). Methods used are content analysis and Mann-Whitney U test respectively. The study uses a non-parametric technique (Mann-Whitney U test) for its quantitative data analysis because of its small and non-probabilistic sample size. The technique was used for the evaluation of the InfoVis prototype dimensions. The participants were grouped based on their years of experience for data compatibility purpose with the statistical technique only.

The details of the instruments designed for the expert review, heuristics evaluation and users' experience evaluation study are discussed in sections 3.7.2 to 3.7.4. The findings of the users' experience evaluation study are presented and discussed in Chapter 6 of this thesis. Figure 3.4 presents phases 4 and 5 (Evaluation and Conclusion) which are the last phases of the research design.

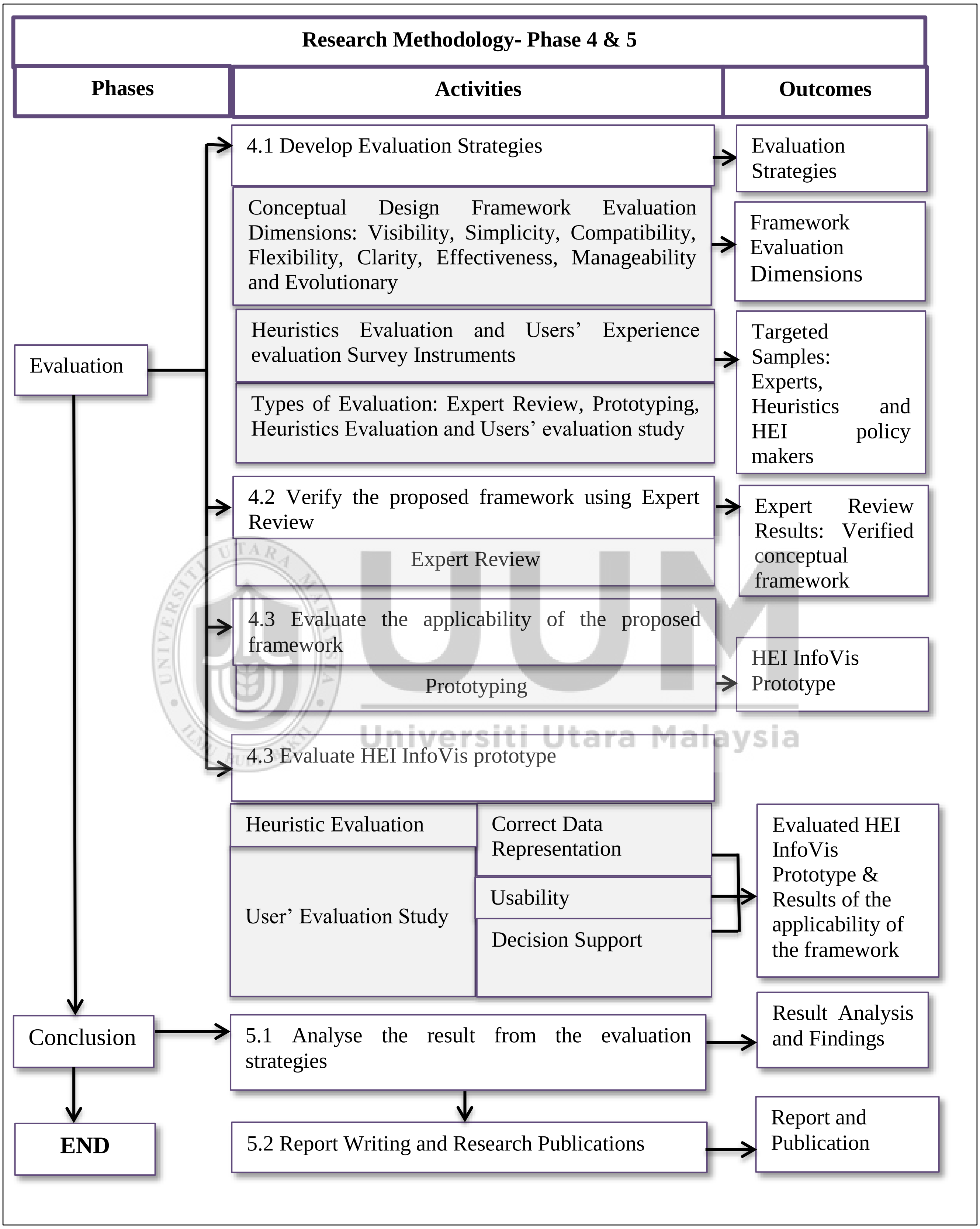


Figure 3.4. Phase 4 & 5: Evaluation and Conclusion Phases

3.7 Instruments Developed for this Study

Instruments designed for this study are:

i. Pre-Design Survey Instruments: a) Interview Sessions Guide (Appendix A) and, b) Pre-Design Questionnaire (Appendix B).

ii. Expert Review Instrument (Appendix C).

iii. Heuristic Evaluation Instrument (Appendix D)

iv. Users' Evaluation Study Instrument (Appendix E)

3.7.1 Pre-Design Survey Instrument

This is the instrument designed for the pre-design study. The pre-design study is conducted to understand the domain where the InfoVis designed (using the proposed framework) is implemented. It also investigated the explicit knowledge preferences involved in the HEI students' data-driven decision making processes, and prioritize them accordingly (refer to Appendix A and Appendix B). Section 1.4; Chapter 1, discusses part of the pre-design study. The validity and realibility of the items are also discussed in section 1.4. The findings that are related to the proposed conceptual design framework are reported in chapter 4 of this thesis.

3.7.2 Expert Review Instrument

Expert Review is employed for the verification purpose of the proposed conceptual design framework. Numerous past studies on evaluation in InfoVis border on evaluation of the InfoVis software or/and application (Toker, Conati, Steichen & Carenini, 2013; Forsell, 2012; Faisal, Craft, Cairns & Blandford, 2008).

This study adopted an expert review instrument of Syamsul (2011) which was developed by adapting the design models and methodologies proposed in general software development and multimedia studies of Hecksel (2004), Kerzner (2006), Bonner (2008), Rogers et al. (2010), and Shneiderman and Plaisant (2010). The instrument's validity and realibility were done by Syamsul (2011) and reported to be satisfactory. The specifics of InfoVis designers/developers are substituted for mobile application developer that Syamsul (2011) worked on.

The dimensions to verify the proposed InfoVis conceptual design framework are visibility, complexity, compatibility, flexibility, clarity, effectiveness, manageability and evolutionary. The expert review instrument was used by InfoVis developers and academics with prior experience in data visualization to verify the proposed conceptual design framework. Table 3.1 presents the definitions of the dimensions, the description and its source of reference as presented by Syamsul (2011) and adapted to meet the peculiarities of this study. The instrument is presented in Appendix C.

Table 3.1

Description of the dimensions for the evaluation of the proposed HEI students' data-centred InfoVis Conceptual Design Framework

Dimensions	**Description**	**Source(s)**
Visibility	It (components and process) is visible to the InfoVis designers and developers, i.e. they can judge the completeness and relevance of the InfoVis development.	Hecksel (2004); Kerzner (2006)

Simplicity	It is simple to understand and use. It is clearly understandable.	Bonner (2008); Shneiderman and Plaisant (2010)
Compatibility	It is compatible (consistent) with the existing values of software engineering. It is compatible with InfoVis development.	Kerzner (2006); Bonner (2008)
Flexibility	It provides flexible development phases with negligible rigour. It is adaptive and responsive to users' needs, then flexible for future use.	Hecksel (2004), Bonner (2008)
Clarity	Its phases with steps and activities included are easily follow-able. It gives a guide on how InfoVis can be developed.	Hecksel (2004); Shneiderman and Plaisant (2010)
Effectiveness	It will enhance productivity, effectiveness and quality of InfoVis development	Hecksel (2004), Rogers et al. (2010)
Manageability	The processes and activities in the framework/ methodology are manageable and controllable. It provides a good guide for project management.	Hecksel (2004); Rogers et al. (2010)
Evolutionary	It provides a dynamic process that evolves continuously with feedback from users. The framework/methodology has the	Hecksel (2004); Bonner (2008)

	capability for incremental change and could cope with new ideas. It provides the opportunity for communication and collaboration with end users and supports incorporation of the evolving requirements.

3.7.3 Heuristic Evaluation Instrument

The heuristic evaluation instrument was used to evaluate the prototype in terms of data representation correctness and usability using open-ended questions. Correct data representation is measured using visual appeal, simplicity, organization and understandility (Chil, Diehl, & Norman, 1988). Usability is measured using perceived usefulness, ease of use and learnability (Leventhal & Barnes, 2008; Brooke, 1996; Nielsen, 1993), (See Appendix D). This serves as a formative evaluation for the prototype. The questions are made open-ended so that the heuristics can objectively assess the prototype.

3.7.4 Users Experience Evaluation Study Instrument

This instrument contains instruction on tasks expected to be carried out by the end users while interacting with the InfoVis system. In so doing, the InfoVis techniques specified by the conceptual design framework and designed for in the prototype can be accessed. Adequate briefing on each of the task cases was given. Each of the tasks assigned has its respective open-ended question(s). The assigned tasks and attached questions are adopted from Mazza and Berre (2007).

The second part of the instrument contains close-ended questions to evaluate the InfoVis prototype in terms of its data representation correctness, usability and decision support. Participants' responses are coded using the ordinal Likert scale, as appropriate. The instrument is presented in Appendix E. Table 3.2 presents the tasks to be executed by the users, and the debriefing questions (open-ended), and the associated techniques (specified by the conceptual design framework) that are implicitly evaluated. Table 3.3 presents the items (closed-ended questions) for the users' experience evaluation instrument.

Table 3.2
Users' tasks and associated debriefing questions

Tasks	**Description**	**Debriefing Questions**	**Associated Techniques from the conceptual design framework**
Locate	The user is to find a particular information s/he knows initially, and indicate this by describing it	Can you identify particular information that you knew before the visualization?	Visualization, Interaction

Identify	The user is to find information s/he does not has its prior knowledge.	Can you identify particular information that you did not know before looking at the visualization?	Visualization, Interaction
Distinguish	The user is able to distinguish among different objects used in the visualization	Can you distinguish among different objects in the dataset?	Visualization, Interaction
Categorize	The user is able to identify divisions of items categories using visual objects	Can you describe the categories of the datasets?	Visualization, Interaction, Distortion
		Do you think this representation helped you in classifying the data?	Visualization, Interaction, Distortion
Distribute	The user characterises the distribution of attribute values over the set of data cases	Do you think that this representation gives you a clear picture of the distribution of the objects and their values?	Visualization, Interaction, Distortion

Compare	The user compares similar entities of data represented	Do you feel that this representation provides sound comparison among similar objects?	Visualization, Interaction, Distortion
		Do you feel that this representation provides sound comparison among different objects?	Visualization, Interaction, Distortion
		Do you feel that this representation provides clear picture of the relations between objects?	Visualization, Interaction, Distortion
		Do you feel that this representation underscores the most important relationship?	Visualization, Interaction, Distortion
Associate	The user forms a relationship between items displayed.	Can you identify the relationship between objects represented?	Visualization, Interaction, Distortion
		Does this representation give you the opportunity to better establish	Visualization, Interaction, Distortion

		relationship between objects?	
Correlate	The user could determine if there is a relationship between values of two attributes of dataset	Is there any correlation between values of the objects viewed?	Visualization, Interaction
		Do you think this representation gives you the opportunity of identifying new relationships?	Visualization, Interaction

Table 3.3

Closed-ended Items for the Users' Experience Evaluation Study

Dimensions	**Items**	**Source(s)**
Correct Data Representation	It highlights important information	Chil, Diehl, & Norman, 1988
	It simplifies the data representation	
	The data shown after exploration attracts my interest	
	The presented information is well organized	

	I can understand the information represented at a glance	
Usability	I would like to use this system frequently	Brooke (1996); Nielsen (1993); Leventhal & Barnes (2008)
	I found the system simple	
	The system was easy to use	
	I would need the support of a technical person to be able to use this system	
	I found the various functions in the systems well integrated	
	The system operation was too inconsistent	
	I imagine that most people would learn to use this system quickly	
	I found the system simple to use	
	I felt confident using the system	
	The system has the functions and capacities I expect it to have	Lewis (1991)
Decision Support	The screen layout is visually appealing	Shneiderman and Plaisant (2010)
	The colouration of the data objects elements is pleasant	

	The visual representation supports the data insight	Tory & Moller (2005); Lam et al. (2012)
	Exploration functions (filtering and extracting information) are consistent with needs	
	The system presents the hidden information	
	The data exploration process reveals insightful information	Lam et al. (2012)
	The system well-structures the presented information	
	The information presented meets my goal of using the system	Baker et al. (2002)
	The system supports decision making	

The items for correct data representation are adapted from Chil, Diehl, and Norman's (1988) work on web information presentation. This study also adopts Nielsen's (1993) usability metrics which are 'easy to learn', 'efficient to use', 'easy to remember', 'few errors and subjectively pleasing.' Brooke (1996) and Leventhal and Barnes (2008) are extended works on Nielsen usability metrics. The items measuring the usability are adapted from these sources, as shown in Table 3.3. The items measuring the decision support are designed based on conceptual explanation of

InfoVis functionality, as given by Lam et al. (2012), Tory and Moller (2005), and Baker et al. (2002). This is also supported by the earlier theoretical explanation on InfoVis support (presented in chapter 2, section 2.4) which posits that a combined measurement of the InfoVis' supports -perceptual, cognitive and decision making –is a valid approach of measuring the InfoVis' decision support (Semiu & Zulikha, 2015a).

3.8 Phase 5: Analysis, Findings and Conclusion

Methods of analysis are chosen based on the nature of the data collected (qualitative and quantitative) and the objective of the evaluation study. The data analysis was broadly using both qualitative and quantitative approaches. The qualitative approach was done using Elo and Kyngas' (2008) content analysis method. This method has three phases which are preparation, organization and reporting. Activities involved are definition of the unit of analysis and content of the responses, open coding, categorization and grouping, formulation of higher heading and thematization. Further description and details as related to this study are presented in Chapter 6.

The quantitative data collected from the users' experience evaluation experiement was analysed using Mann-Whitney U Test which is a non-parametric technique. Non-parametric technique is suitable when sample size is not a true representation of the population distribution, for ordinal and categorical data, and with small sample size (Pallant, 2011). This study employed the non-parametric technique because its sample is purposive (non-probability), small, and also with categorical data. The interpretation of Mann-Whitney U Test is restricted to the ranking of the evaluation metrics of the prototype (i.e. correct data representation, usability and decision support) using the Mean Rank values. The findings of the users' experience evaluation study are presented in Chapter 6 of this study.

3.9 Samples

The samples used in this study constitute a) HEI policy makers, b) InfoVis or data visualization developers in industry, and academics with prior experience, c) HCI/usability experts, and d) data analysts/administrators, as elements. The HEI policy makers, exclusively from Universiti Utara Malaysia, are used for the validation of the domain explicit knowledge. They are members of the management at the university's school, college and senate level.

The InfoVis or data visualization developers in industry, and academics with prior experience, are used as experts for the verification of the conceptual design framework. The HCI/usability experts and data analysts/administrators are used for the heuristics evaluation. Lastly, the HEI policy makers are finally used for the summative evaluation of the InfoVis prototype.

The pre-design instrument was used to collect users' requirement from HEI policy makers and administrators at the pre-design stage. The policy makers are employees of HEIs with at least 3 years in administrative posts. The sampe elements found Assistant Registrar, Registrar, Dean, Deputy Dean and Deputy Vice Chancellor. The required number of years of administrative experience was randomly chosen since there is no rigid rule. Three (3) years, as required years of experience, was chosen considering a rule of thumb which suggests that an employee with a particular rank after 3 years is likely eligible for promotion (MACPS, 2009). This suggests that, they would have acquired considerable experience that fit into the purpose of this study. A sample size of forty (40) was purposively selected because it can adequately provide responses to questions asked and sufficient for the statistical analyses. However, 32 responses were received which is still satisfactory. Purposive sampling emphasises the ability of the respondents selected and their number based on the purpose of the study with no or less emphasis on the sample size ratio to the population (Tongco, 2007).

For the expert review, i.e. verification of the conceptual design framework, five persons; three (3) academes and two (2) industrial practitioners (in no particular order) were employed. All the expert reviewers have experience in both data visualization projects and HCI. Three of them have experience in InfoVis research design framework. Nielsen (1993; 1997) supported the choice of three to five experts for expert verification.

The heuristics evaluation involved heuristics of two categories and four (4) members. These are two usability/HCI experts and two data administrator/ analysts. The profile of the expert reviewers and heuristics evaluators is presented in Appendix H. Lastly; the users' experience evaluation study involved HEI policy makers. The HEI decision makers from higher insititutions for the users' experience evaluation study are from Nigeria, Soduth Africa, Malaysia and United States of America. The total numbers of the participants are fourteen (14). This is acceptable for users' experience evaluation because there are no definite participants' size (Tullis & Albert, 2008).

3.10 Summary of the Chapter

This chapter discusses the research approach adopted by this study in answering its research questions. It explains design science research methodology with specifications illustrating each phase of the research methodology in view of accomplishing the stated objectives of this study. From these objectives, this chapter shows that a pre-design study using questionnaire, critical analysis of the InfoVis design literatures and past design frameworks are used in proposing the InfoVis conceptual design framework. The proposed conceptual design framework is verified using expert review method.

The framework is to be further used as guide in developing an InfoVis tool as a proof of concept. The evaluation strategies are into formative and summative sessions involving heuristic evaluation and users' experience evaluation study of the InfoVis prototype, respectively. The instruments of data collection are presented, and items adopted and adapted are adequately justified. Lastly, the sample size and their sampling techniques, as applicable, are described and highlighted.

CHAPTER FOUR
THE PROPOSED FRAMEWORK: HEI STUDENTS' DATA INFOVIS CONCEPTUAL DESIGN FRAMEWORK

4.1 Introduction

This chapter explicitly outlines the proposed HEI students' data InfoVis conceptual design framework. The framework entails phases, components, activities and deliverables. The behaviour and form of the InfoVis developed based on the specificity of HEI students' data for related decision support are also highlighted. The conceptual design framework is developed to guide InfoVis designers and developers, specifically in developing students' data-focused InfoVis systems. The development of the framework attends to the problem identified in Chapter 1, based on the review of its characteristics and components shown in Chapter 2, and by following the methodology outlined in Chapter 3.

HEI students' InfoVis conceptual design framework was developed after certain activities were conducted. The phases, components and activities of the framework are premised by these activities (shown in Table 4.1). The HEI students' data InfoVis (HSDI) conceptual design framework has three distinctive but interdependent components. These are process and structural models, and content delivery techniques. Phases of the process model are separate but interconnected with respective stages that are to be performed in a sequential order. The structural model depicts the core structures of an InfoVis system. The content delivery techniques highlight the chosen InfoVis techniques that are principally responsible for the delivery of the behaviour and form of the InfoVis content. Also, with special attention to the contributing theories, a theoretical guide is developed, and the conceptual design framework of HSDI was proposed. Table 4.1 presents the summary of the activities performed prior to the proposal of the framework.

Table 4.1

Activities performed prior to proposing the Framework

Elements	**Activities**	**Chapters**
Process Phases	Pre-Design Study	Chapter 1 & 4
	Critical Review of InfoVis Literature	Chapter 2
Content	Pre-Design Study	Chapter 1 & 4
Components	Content Analysis of past InfoVis studies	Chapter 2
Structure	Content Analysis of past InfoVis studies	Chapter 2
Flow	Pre-Design Study	Chapter 1 & 4
	Comparative Study of Information systems development model	Chapter 3
Activities	Pre-Design Study	Chapter 1 & 4
	Content Analysis of InfoVis Design Literature Review	Chapter 4

All these activities collectively achieved the conceptual design framework of HSDI. The next section discusses the pre-design study and its influence on the development of the conceptual design framework.

4.2 Pre-Design Study

In Information System (IS) research and development generally, users' requirement study is essentially suggested for the purpose of realising a user-centred information system (Rogers et al., 2011; Sheinerderman & Plaisant, 2010). In the same vein, InfoVis practitioners and researchers highly recommend pre-design study when developing InfoVis framework or system (Lam et al., 2012; Koh et al., 2011). InfoVis pre-design study is to understand the environment and the work practices of the domain where the InfoVis will be applied. This, arguably, will also aid the actualization of an InfoVis tool that is both task-centred and user-centred (Lam et al., 2012; Robinson et al., 2005; Greenberg, 2002).

Based on this InfoVis pre-design recommendation (Lam et al., 2012; Robinson et al., 2005; Koh et al., 2011; Roth et al., 2010; Zhong et al., 2012, and Meyer, 2012), this pre-design study was conducted to (i) identify the work practices of the HEI data administrators in terms of the nature of the data management infrastructure, and the associated administrative activities, (ii) identify the data analysis activities currently done by the HEI data analysts, and (iii) identify the explicit knowledge preferences of the HEI decision makers, i.e. the preferred output of the transformed data that will support their decision making process.

To achieve the above three (3) stated objectives, the pre-design study employed interview sessions for (i) and (ii), and administration of survey questionnaire for (iii). This method (Quantitative and Qualitative) is an addition to the few studies of same method in InfoVis study. Majority of previous InfoVis pre-design studies (Isenberg et al., 2008; Jung et al., 2011; Hinrichs et al., 2008; Huang et al., 2010; Zhong et al., 2012; Mansmann et al., 2009; Roth et al., 2010; Lam et al., 2012; Robinson et al., 2005; Grammel et al., 2010; Koh et al., 2011) extensively employed Qualitative method for users' studies, while Tory et al. (2008) used mixed methods. Therefore, this study –by utilizing a combination of both methods – is confident that reliable findings of HEI InfoVis users' study are presented.

4.2.1 Interview of HEI Data Administrators and Data Analysts: Methods and Findings

Due to the practical and task-based nature of HEI data analysts and administrators' work, interview (Rogers et al., 2011; Sheinerderman & Plaisant, 2010) was employed to precisely understand the context of the HEI data infrastructure and its enabling features for an InfoVis system. Also, the data analytical methods that are generally employed by HEI data analysts were identified. The interview session with a lead data administrator of an HEI was conducted on 23rd February, 2014, and that of the data analyst on 8th April, 2014. As suggested by Rogers et al. (2011), the interview with the HEI data administrators and data analysts was conducted with a guide (see Appendix A). After the interview sessions, a content analysis approach was employed. This involves transcriptions (see Appendix F), and thematisation (Creswell, 2009).

The findings from the interview conducted with the data administrators and data analysts showed that a) HEI institutions mainly use data warehouse technologies for data administration, b) data needed for corporate planning are demanded on a case-by-case basis by the corporate planning unit. The nature of the data usage determines the scope and varieties of the data dimensions/attributes, c) Microsoft Access, Excel and Power-pivot are used for data analysis to produce the pie chart, bar chart and line chart, depending on the information to be presented.

4.2.2 Survey of HEIs' Decision Makers Explicit Knowledge Preferences: Methods and Findings

Although there is yet to be an InfoVis framework or InfoVis system centred on HEI students' data, substantial works (Delavari & Phon-Amnuaisuk, 2008; Microsoft Inc., 2004; Lange et al., 2012; Protheroe, 2009; Ewell, 2012; Mandinach et al., 2012; El-Fattah, 2012; Wagner and Ice, 2012; Bresfelean et al., 2009; Sarker et al., 2010) have been done on data-driven decision making in schools generally and HEI specifically.

It is established that decision support is only feasible when decision makers discover a previously-hidden and unknown explicit knowledge. These knowledges are products of the transformation process of the domain data sets (Delavari & Phon-Amnuaisuk, 2008). Therefore, an InfoVis designed to produce the explicit knowledge preferences of HEI decision makers will automatically support their decision making processes. On this note, this study identified the explicit knowledge preferences of the HEIs' policy makers that support the students' data-driven decision making processes.

Survey questionnaires (see Appendix B) were administered to a total number of purposively sampled forty (40) respondents among HEI decision makers, and thirty-two (32) responses were received. This conforms to provisions of Rogers et al. (2011), Sheinerderman & Plaisant (2010) and Lazar et al. (2010). Based on the review of HEI related literatures on data-driven decision process (presented in Table

4.2), a twenty-nine (29) items suggesting the explicit knowledge of HEI decision makers were designed. The administration of the survey instrument is to practically elicit the prioritized explicit knowledge preferences

Table 4.2

Knowledge Discovery Goals and Explicit Knowledge from HEIs' students' data

	Sources	HEI students' data-centred knowledge discovery goals	Explicit Knowledge
1.	Microsoft Inc. (2004)	To identify the relationship between students' racial background and gender with the students' performance	Correlation between the students' demographics (gender and race/ nationality) and the students' performance
		To ensure an equitable admission quota across the country's diverse racial composition	Descriptive information of the admitted students based on their racial background.
2.	Delavari & Phon-Amnuaisuk (2008)	Discovering the students' topologies	The pattern of various groups of students
		To predict Alumni pledge	The pattern of previous graduates contributions to the university
		To discover the relationship between students' gender and race and their academic performance	The historical relationship of the students' score with their gender, race and attendance
		To suggest the admission preference based on gender and race.	
		To discover students' learning outcome based on different clusters.	Performance pattern of various students' groups.

To determine the most favourable mode of entry to the students' performance	Success history of the students. Relating their performance to their mode of entry.
	The success pattern of students that transferred course credits on admission
To assess the students' performances based on their mode of entry (in specific here is: students that were admitted on credit transfer).	
To predict overall students' performance	Classified pattern of previous students on the basis of their final grade (in terms of their CGPA)
To improve the quality of the graduate students from the degree of employability.	The pattern the students' majors (like BSc IT (Artificial Intelligence) and their employability.
To know the association between students' performance with the time and venue of lectures.	Students' performance relationship with Lectures venue and time
To identify the relationship in the students' performance and their level (number of semester used).	Students' performance relationship with the Students' level
To identify the pattern of students who were in a particular program.	Pattern of previous students with specificity to a particular program

		To identify the pattern of students suitable for a particular majoring.	Students' classification to courses
		To identify the pattern of the students' performance based on the courses combined	Using courses combination for students' performance history
		To identify the association of the students' health history with the students' performance	Association of Students' health history with students' performance pattern.
		Knowing the type of disability or ailment that could negatively affect the students' performance	
3.	Protheroe (2009)	To identify the specific needs of instructional upgrade based on the students' assessment report	Students' assessment data for instructional improvement, Students' academic progress report
		To design intervention strategy to arrest academic risk-prone students	
4.	Bresfelean et al. (2009)	To predict students' attrition, and identify its cause like study break (deferment), courses combination and majoring, financial aid to indigent students	Students' academic records, students' learning data and skill acquisition history (certificate courses, workshop and symposia attended), students' rating of the IT-supported courses, Student socio-economic status, Courses enrolment rate, IT infrastructure utilization index.

		To identify courses that attract the students’ enrolment	
		To measure the employability rate of the institutions’ graduate and identify the causative factors	
		To evaluate the impact of course enrolment combination on the students’ academic performance	
		To identify the impact of internship and other skill acquisition program on the students’ academic performance	
5.	Sarker et al. (2010)	To design the school curriculum in alignment with the employment reality, and to attract the international students with the globally ranked instructional module.	Students’ After Graduation waiting period, Alumni Feedback, Students’ academic records (potential and skill profile, academic attainment, pre-university academic records, GPA), Students retention, enrolment and employability rate
		To gauge the institutions’ resources (human, capital and infrastructure) to determine the number of fresh students to be admitted	
		To identify the criteria in admitting fresh students into the higher education institutions	

6.	Lange et al. (2012)	To identify the students’ learning achievement pattern based on the students’ learning data, students’ demographics data and the school demographics data	Students’ learning achievement in relationship with Student learning data (norm-referenced tests, state assessments, grade distributions, Grade point averages and benchmark tests), Students’ demographic data (students’ ethnicity, socioeconomic status, language proficiency), and School demographic data (overall attendance rates, retention rates, student discipline infractions, drop-out rates, socioeconomic factors, and political factors).
7.	Ewell (2012)	To assist in designing programs for accreditation and improved university ranking	Description and pattern of the Enrolments /degrees awarded graduation and retention rates, student-faculty ratios
8.	El-Fattah (2012)	To determine the admission rate of the fresh students based on: The institutions’ Lecturer-Students ratio The available infrastructural resources available in the institutions. To decide on any of the emerging criteria to be considered when admitting new students.	Students enrolment rate, Students- Lecturer ratio, Institutions’ infrastructure info
9.	Wagner and Ice (2012)	To identify factors affecting students’ retention and progression	Students’ performance pattern, Students’ achievement data and Students’ demographics data,

			Institutions ranking position and infrastructural index
		To determine the factors responsible for progressing distance learning students	
		To identify if students' demographics and institution peculiarity influence the students' performance pattern	
10.	Mandinach et al. (2012)	To identify the factors to develop formal coursework ware,	Courses enrolment rate, Students' employability rate, Students' academic profile (potential and skill profile, academic attainment), Program and course rating
		To design in-school skills acquisition and professional development programs for the students to path a more successful employability rate.	
		To design students-friendly academic programs for the continuing education (distance) students.	

The HEI students' data-centred explicit knowledge preferences (shown in Table 4.2) are designed as survey questions. The responses are codified with 1 = No, 2 = I don't know, and 3= Yes. Table 4.3 presents the result of the survey conducted to validate the explicit knowledge preferences of HEI policy makers.

Table 4.3

Explicit Knowledge Involved in Students Data-driven decision making process in HEIs

	Item: At my institution, we seek to know…	No =1	Don't know =2	Yes =3	Identified Data Dimensions
1.	The pattern of various groups of students based on their race, religion, sex and nationalities.	2	5	25	Race, Religion, Sex and Nationality
2.	The pattern of previous graduates based on their race, religion, sex and nationalities.	5	2	25	Year of graduation Race, Religion, Sex and Nationality.
3.	The pattern (course of study, CGPA and nationalities) of the Alumni who are responsive to the university's invitation.	11	3	18	Course of study, CGPA, Nationality, Status of Studentship
4.	The relationship between the students' performance and their gender, race and nationalities.	8	4	20	Students' performance (CGPA), Gender, Race and Nationality
5.	The trend of the students enrolment based on their nationalities and sex.	2	5	25	Year of enrolment, Nationality, Sex
6.	The trend of the students enrolment based on their mode of admission and gender.	2	5	25	Year of enrolment, Mode of admission and Gender
7.	The relationship between students' mode of entry (foundation, matriculation, credit transfer), criteria set for admission, and their academic performance.	4	6	22	Mode of entry, CGPA
8.	The pattern of the students' economic status (self/family sponsored and scholarship) per enrolment.	8	6	18	Mode of sponsorship, Year of enrolment.

9.	The performance of students with disabilities.	8	9	15	Disability
10.	The performance of students with unfavourable health history.	7	10	15	Health weakness
11.	The trend and pattern of the university's graduating students based on gender, race and nationalities.	4	8	20	Year of graduation, gender, race and Nationality.
12.	The relationship between the graduating students' performance with their gender, race and nationalities.	7	4	21	Year of graduation, CGPA, gender, race and Nationality.
13.	The relationship between students' course of study and their employability.	5	5	22	Course of study and Employability
14.	The relationship between students' majoring courses and their employability.	6	2	24	
15.	The relationship between students' performance and their lecture venue and time.	11	5	16	CGPA, Lecture time and Venue
16.	The relationship between students' majoring courses and their performance.	7	4	21	Course of study and CGPA
17.	The trend of the students' performance in relationship with number of semester used.	8	0	24	CGPA and Number of Semester used
18.	The relationship between the students' performance and their program of study (degree & graduate).	6	0	26	CGPA and Program of study
19.	The relationship between the students' performance and their mode of study (full time & part time).	2	7	23	CGPA and Mode of Study

20.	The pattern of students in a particular department, school or college based on race, gender and nationalities.	8	2	22	Department, school/College, Race, Gender and Nationality.
21.	The courses that are majorly or otherwise enrolled by the students.	5	2	25	
22.	The relationship between the majorly (or otherwise) enrolled course and the students' performance in the said course.	5	5	22	Majoring course, CGPA
23.	The relationship between the student performance and their English language proficiency.	6	3	24	CGPA and English Language proficiency
24.	The pattern of the drop-out students.	6	6	20	Year of enrolment, Status of studentship
25.	The relationship between drop-out students and their economic status (self/family sponsored and scholarship) and health history.	7	11	14	Health weakness, Mode of sponsorship
26.	The pattern of the drop-out students based on their gender, nationality and marital status.	8	5	19	Gender, Nationality and Marital status
27.	If there is impact of soft skill acquisition on the students' academic performance.	6	6	20	Soft skill quality, CGPA
28.	The significance of internship on the students' employability.	5	5	22	Internship length and employability
29.	The students' deferment history and its relationship with students' performance.	4	7	21	Status of studentship, CGPA

The explicit knowledge preferences, as elicited from the responses of the HEI policy makers, were then prioritized using the number value of the "Yes" recorded. Table 4.4 presents the prioritized explicit knowledge preferences of HEI policy makers.

Table 4.4.

Explicit Knowledge Preferences of the HEIs Policy Makers

	Students' data-driven Explicit Knowledge preferences	Yes
1.	The relationship between the students' performance and their program of study (degree & graduate).	26
2.	The pattern of various groups of students based on their race, religion, sex and nationalities. The pattern of previous graduates based on their race, religion, sex and nationalities. The courses that are majorly or otherwise enrolled by the students. The trend of the students enrolment based on their nationalities and sex. The trend of the students enrolment based on their mode of admission and gender.	25
3.	The relationship between students' majoring courses and their employability. The trend of the students' performance in relationship with number of semester used. The relationship between the student performance and their English language proficiency.	24
4.	The relationship between the students' performance and their mode of study (full time & part time).	23
5.	The relationship between students' mode of entry (foundation, matriculation, credit transfer), criteria set for admission, and their academic performance. The relationship between the majorly (or otherwise) enrolled course and the students' performance in the said course. The pattern of students in a particular department, school or college based on race, gender and nationalities. The relationship between students' course of study and their employability.	22

	The significance of internship on the students' employability.	
6.	The relationship between the graduating students' performance with their gender, race and nationalities. The relationship between students' majoring courses and their performance. The students' deferment history and its relationship with students' performance.	21
7	The trend and pattern of the university's graduating students based on gender, race and nationalities. If there is impact of soft skill acquisition on the students' academic performance. The pattern of the drop-out students. The relationship between the students' performance and their gender, race and nationalities.	20
8.	The pattern of the drop-out students based on their gender, nationality and marital status.	19
9.	The pattern (course of study, CGPA and nationalities) of the Alumni who are responsive to the university's invitation. The pattern (course of study, CGPA and nationalities) of the Alumni who are responsive to the university's invitation. The pattern of the students' economic status (self/family sponsored and scholarship) per enrolment.	18
10.	The relationship between students' performance and their lecture venue and time.	16
11.	The performance of students with disabilities. The performance of students with unfavourable health history.	15
12.	The relationship between drop-out students and their economic status (self/family sponsored and scholarship) and health history.	14

On another hand, lower quartile range, as a measure of variability, was used to detect the outliers among the responses (Tullis & Albert, 2008). Outliers, in this sense, are responses in the extreme bottom of the responses gathered. Therefore, items with more than eight (8) "No" responses were invalidated (Note: 8 is approximately the lower quartile (Q_1) of 32 (Total number of the respondents). It is derived with this formula: $^1/_4\,(n+1)$, where $n = 32$). From the findings presented in Table 4.3, items 3 and 15 were struck out of the explicit knowledges to be covered by HEI InfoVis. The result also showed that relationship between the students' performance and their program of study is the most preferred explicit knowledge (with 26 endorsements out of 32).

4.2.3 Implication of the Findings on this Study

The findings from the interview sessions showed that the proposal of InfoVis design framework for HEI students' data management is important, necessary and well-timed. The HEI data administrators agreed on the need to explore better tools that can be used by the domain users in understanding the domain dataset.

From the validated explicit knowledge preferences presented in Tables 4.3 and 4.4, a comprehensive data dimensions (attributes) expected of each data instance in the HEI students' dataset was elicited. The data dimensions represented the entirety of the validated domain explicit knowledge preferences. It is the process of relating these data dimensions appropriately that delivers the InfoVis content that support the HEI decision makers. Also, the nature of these data dimensions and the users' tasks determine the choice of the InfoVis techniques. The descriptions of the data dimensions follow a logical conceptualization that resonates with real life experience. Table 4.5 presents the HEI students' data dimensions.

Table 4.5

HEI Students' Data Dimensions

	Data Dimension	Class	Data Type	Description
1	Student ID	Basic	Numeric	This is the student identification number. It serves as the primary key for each of the data cases.
2	Gender	Basic	Categorical	This represents the students' sexuality, i.e. Male or Female
3	Nationality	Basic	Categorical	The citizenship identity of the students. E.g. Malaysia, India, Nigeria, etc.
4	Race	Basic	Categorical	Continent-based identity. E.g. African, Asian, American, etc.
5	Religion	Basic	Categorical	The faith identity of the student. E.g. Islam, Buddhism, Christianity, etc.
6	Course of Study	Basic	Categorical	The course that the student is offering. E.g. Information Technology (IT), Business Administration, etc.
7	Specialization	Basic	Categorical	When applicable, this describes the area in which the student is specializing. E.g. Artificial Intelligence as a specialization for IT students.
8	Year of Entry	Basic	Numeric	The year that the student entered into the university. E.g. 2010, etc.

9	Year of graduation	Basic	Numeric	The year that the student graduated (or will graduate) from the university. E.g. 2010, etc.
10	Number of Semester Used	Basic	Numeric	The number of semester used by the student. E.g. 2, 3, 4, etc.
11	Mode of Admission	Basic	Categorical	This describes the entry method into the university. E.g. Foundation, Matriculation, Diploma
12	Mode of sponsorship	Basic	Categorical	This describes how the student is sponsoring his or her education. E.g. Loan, Family, Self-Study, and Scholarship
13	Current CGPA	Basic	Real Number	The current grade of the student which is normally with fractions. E.g. 3.2, 3.7, etc.
14	Studentship Status	Basic	Categorical	This is to show the status of the student. E.g. Active, Non-Active, and Drop-Out
15	English Language Proficiency	Basic	Categorical	When applicable, this gives the English language certification record of the student. E.g. MUET, IELTS, TOEFL, etc.
16	Disability	Basic	Categorical	When applicable, the disability type of the student is stated. E.g. Physical, Hearing Impairment, etc.
17	Mode of Study	Basic	Categorical	This is to categorise students on full time and part time basis.

18	Program of Study	Basic	Categorical	The program the student is undertaking. The values here are: BSc, PGD, MSc, PhD, and DBA.
19	Nature of the Program	Basic	Categorical	This is a dichotomous variable, i.e. a variable of two values. These values are undergraduate and postgraduate.
20	Internship length	Derived	Numeric	The total number of months the students had used on internship.
21	Soft skill quality	Derived	Numeric	The total number of workshop or seminar on soft skill (public speaking, leadership, time management, etc.) acquisition attended.
22	Health Weakness	Derived	Numeric	This is calculated by the addition of number of times the students visited the hospital and the number of times s/he was medicated, divided by the total number of semester used.
23	Employability	Derived	Numeric	The total number of months that graduated student used before being gainfully employed.
24	Deferment length	Derived	Numeric	The total number of semesters that the students deferred his or her registration.

As shown in Table 4.5 above, the derived data attributes are computed (derived) through addition and division operations, as applicable. However, the basic data variables are the fundamentals with no arithmetic operations. The categorization method of fundamental and derived quantities of Physics was adopted. These elicited data dimensions served as guide in the subsequent stages of the research process, and finally, in the conceptualization of the design framework.

4.3 Structural Components of HSDI

HSDI follows the generic structure of InfoVis, in terms of its primary structural components. These are visual exploration interface, data binding system, and the data base. All InfoVis tools can be dismembered into this three-tier structure. Therefore, in developing any type of InfoVis, the software developers' technical decision starts with the choice of the components of this three-tier structure.

The visual exploration interface is a graphical user interface (GUI) where the users interact with the stored datasets. It is through this interface that the users interact with the data, and data sets are graphically rendered to present the domain explicit knowledge. The components of this interface are responsible for the interaction, distortion and visualization of the datasets. The data binding system contains the implementation of program code using different and appropriate types of data visualization programming frameworks. The program code at the data binding system maps the data to visual objects to produce information. Lastly, the data base houses the data sets, in form of data table(s) in cases of structured data.

Fourteen (14) InfoVis tools that can be lively accessed, or with available documentations, to serve as guide in creating the generic template of InfoVis structure are analysed. This template, with due consideration of HEI domain specificity, guided the choice of the structural components for HSDI. Appendix G contains the details of the analysed InfoVis tools. Table 4.6 shows the structural components of the analysed InfoVis tool.

Table 4.6

Structural Components of the Analysed InfoVis Tools

Tier	**Components**	**Description**	**Freq. of usage (in %)**
Visual Exploration Interface	Data layer panel	This is the space within the GUI where the information about the InfoVis and the data attributes to be visualized are held	79
	Visualization span	This is space within the GUI that holds the statistical graphics (e.g. Line chart, Radial graph, etc.) used in conveying the explored data analytics to the user.	100
	Interaction Mechanism	This is the collection of interaction events handling programs executed at the browser. Examples are pliant response hinting, dynamic hinting, etc.	92
	Distortion tools	These are tools used in rearranging, enlarging, and reducing the information displayed (through the statistical graphics) on the interface. Examples are Zooming, Scrolling, etc.	100
Data Binding system	Colour Coding	This characterizes the explored data analytics using colour scheme, i.e. different colour to distinguish the information displayed. It is sometimes classified as an interaction technique.	100

	Data Object Modelling	This implements the data visualization with chosen framework and appropriate program code	100
Data	Hierarchies	This is the data that highlights a hierarchical organization of the data sets.	50
	Temporal	This is data about time, highlighting the valid and transactional time.	14
	Discrete	This is the data type that characterizes data based on count (ordinal) or label (categorical)	64
	Networks	This is the data type that provides about nodes in a internetworked computers, IPV4, IPV6, and Mac addresses.	7
	World Wide Web	This is information on the web which are often in tags and mark-up languages	21
	Spatial	This is the data that support Euclidean geometry. It is used in geographical value representation.	14

Table 4.6 contains the findings of the analysed InfoVis tools. It highlights their structural components. With a frequency of usage of 79% and above (among the analysed 14 InfoVis tools) for both the visual graphical interface and data binding system, InfoVis tools have relatively the same structural components, except in the database structure which is exclusively determined by the nature of the domain data. However, the constituents of the visual graphical interface (i.e. visualization, interaction and distortion techniques) and data binding system (i.e. program code,

library and framework) are always specific to the domain's explicit knowledge preferences.

The database component which is to be populated by the befitting dataset is exclusively chosen based on the nature of data (structured, semi-structured and unstructured). Therefore, it does not depend on any applicability rate as weighted percentages in Table 4.6 suggest. In the case of HSDI case, a database housing data tables(s) is proposed because of the structured nature of the dataset and its multidimensionality. In the light of the above, Table 4.7 presents the proposed structural components of HSDI, the justification and the application preferences.

Table 4.7

The proposed structural components of HSDI and the justifications

Tier	Components	Justifications
Visual Exploration Interface	Data layer panel Visualization span Interaction Mechanism Distortion tools	As shown in table 4.4, data layer panel, Visualization span, Interaction Mechanism and Distortion tools are compulsory components of InfoVis' visual exploration interface.
Data Binding system	Colour Coding	As shown in Table 4.5, colour coding is a compulsory component of data binding system
	Data Objects Modelling	This is the implementation of the data visualization code and framework. This can be implemented either at the browser or in the data base as data manipulation language (DML). This is

		guided by Ward et al. (2010)'s *"Interactive Data Visualization: Foundations, Techniques, and Applications"* Details of the programming framework is given in the process model
Database	Table	Based on the nature of the students' data, i.e. a multidimensional dataset, data table in a database will be the befitting data host.

The components are said to be compulsory when they belong to the fundamentals of what InfoVis must possess. From the proposed structural components of HSDI presented in Table 4.6, model for the structural components is obtained and illustrated by Figure 4.1 below. The details of these components –which are determined by the HEI students' data dimensions and domain users' tasks –are covered by the content delivery techniques.

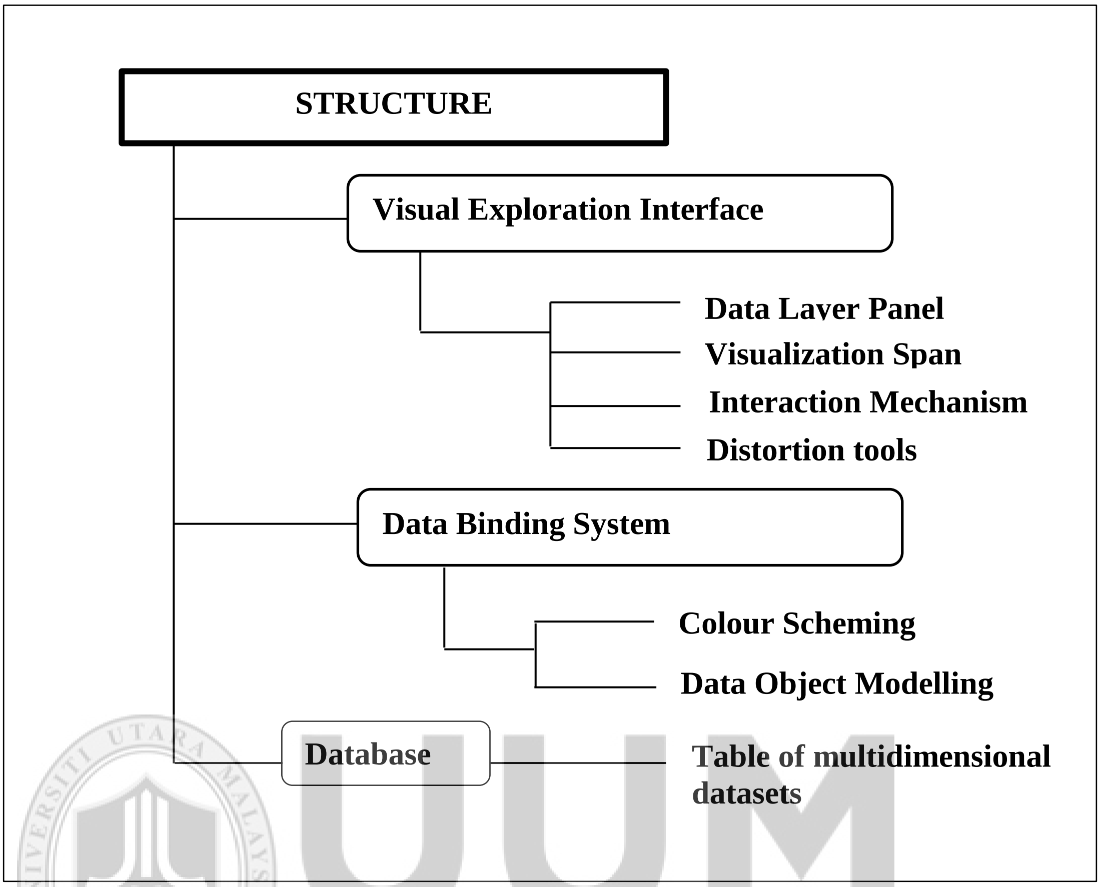


Figure 4.1. Proposed Structural Components Model of HSDI

4.4 Content Components of HSDI

Designing HSDI demands sufficient effort in ensuring that the content (output of the data exploration activities) is delivered by the appropriate content delivery techniques. These delivery techniques are the systematic combination of the InfoVis (visualization, interaction and distortion) techniques. Notably, in respect to structural components as earlier discussed, the data binding system serves as the middle tier that connects the front end (visual interactive interface) with the back end (database). The content delivery techniques present the transformed datasets as explicit knowledge to the users. The visual interactive interface serves as the medium of handling the users' tasks (Gelman & Unwin, 2012; Tan et al., 2006). A review of these content components and characteristics based on the peculiarities of the domain and the respective user's tasks are provided in Chapter 2 of this thesis.

The content components of HSDI handle HEI students-data binding, its rendering, and presentation on the interface. The delivery of the content must be in terms of the explicit knowledge preferences of the domain. The choice of the interaction, distortion, and visualization techniques' types is done based on their appropriateness (Spence, 2007; Cooper et al., 2014). Table 4.8 presents the description of the chosen interaction, distortion, visualization techniques for HSDI, and their justifications. An in-depth analysis of these techniques as provided by the reference books is further given in Appendix L. In this analysis, the functionality of each of the techniques analyses, the nature of data compatibility and the findings from the pre-design study are considered in the choices of the content delivery techniques that were finally made.

Table 4.8

The description of the chosen Content Delivery Techniques (Visualization, Interaction and Distortion) and the justifications

Technique	**Type**	**Justification**
Interaction Techniques	Mouse on Mouse on-click Drop-Down Menu	From the data dimensions defined for HSDI (shown in Table 4.5), those that are ordinal data will be placed in the axes (vertical and horizontal). The categorical data can also be explored in a data panel to support dimensional visualization beyond the 2D y and x axes. This supports filtering for discrete information spaces (Cooper, 2014, pg. 470 – 474, 477).

	Individuals of these data dimensions can only be explored through a drop-down menu. Through this, user can pick the data dimension of interest in line with the current data exploration activity.
Pliant response hinting: push button	This button is to reset or shown animation trend of the HSDI data (Spence, 2007, pg, 180). This is just to enhance users' experience.
Pliant response hinting: Check boxes	The data entity is represented by the nationality of the students. Nationality has values like Malaysian, Chinese, etc. Therefore, users of HSDI will have to individually check-in the entity values they are interested in its details (Cooper, 2014, pg. 470 – 471, addressed as point-and-click).
Dynamic cursor Hinting	This is simply a change in the colour of an area, text, or image, when it is hovered by a mouse touch. This is essentially to grasp the users' attention on the particular point s/he will be interacting with, as further clicking. It forms an essential part in the design of interactive systems generally (Cooper, 2014, pg. 470).
Mouse on-drag	This will be used on zooming tool. It would be allowed to be dragged up and down, translating to zooming-out and zooming-in respectively. This support InfoVis' essence according to the information seeking mantra (Shneiderman, 1997).

Distortion Techniques	Scrolling	This is to allow movement of the information displayed in view of showing other part of the information not shown because of the limited display screen space. There is high possibility of this, therefore is applicable, though not compulsory (Spence, 2007, pg. 98).
	Overview with Detail Exploration Zooming	Overview with details' works like a magnifying lens to provide a detail of a picked sub-region within the overall information displayed. Zooming enlarges the whole information space in view of presenting it in a bolder form. Ability to show detail on demand is one of the primary qualities of InfoVis (Keim & Ward, 2002). This is also in conformity with Sheneiderman's information seeking mantra: Overview first, zoom and filter and then detail-on-demand (Sheneiderman, 1996).
Visualization Techniques	Scatter plot Line Chart Geographical map (Map view) Bar Chart	Based on the nature of the elicited students' data dimensions as shown in Table 4.4 and the findings of the domain analysis as reported in sections 4.2.1 and 4.2.2, these visualization techniques are appropriate. The Map view is essentially to show the students' demographics based on their countries, while others –scatter plot, line chart, and bar chart – can present axes-based data display (Spence, 2007, pg. 98) as suitable for the validated HEI explicit knowledge preferences.

Based on the chosen and justified content components, as presented in Table 4.8, the content delivery techniques of HSDI is proposed and illustrated by Figure 4.2.

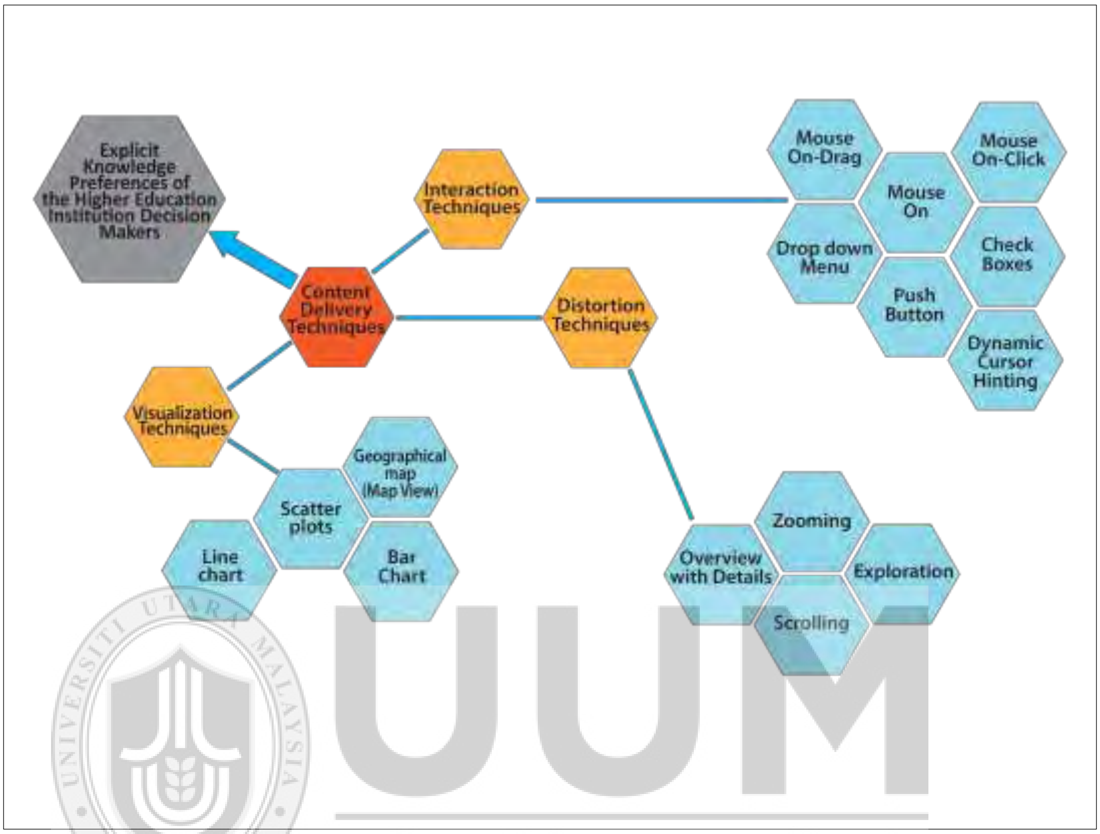


Figure 4.2. Proposed Content Delivery Techniques of HSDI (Initial)

The exploration of students' dataset, which is primary function of HSDI, was done using the content delivery techniques depicted in Figure 4.2. As shown in Table 4.8, the choice of the content delivery techniques for HSDI considered the data analysts, data administrator and policy makers as its human entities. Impliedly, these three categories of users must be able to use and relate with HSDI. The content to be delivered is the students' data focused explicit knowledge preferences, as periodically demanded during the data exploration activity. Figure 4.3 presents the content diagram of the HSDI.

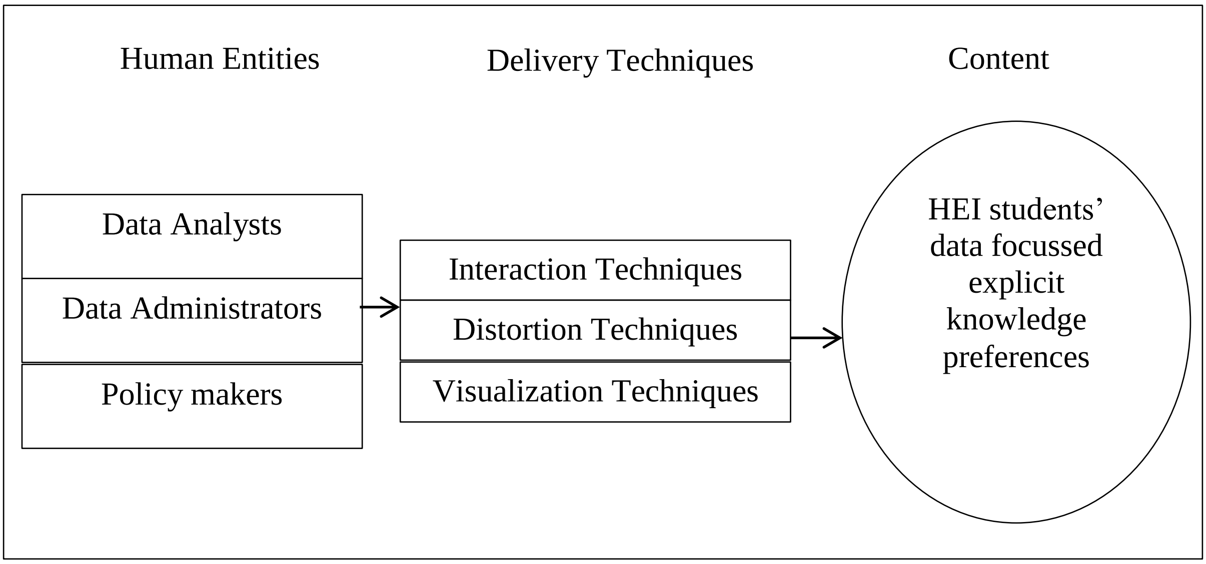


Figure 4.3. Content Delivery Diagram for HSDI

4.5 Process of Developing HSDI

The process of developing HSDI was developed after a critical analysis of the InfoVis techniques' application as related by Ward et al. (2010), visual data mining project methodology by Soukup and Davidson (2002). This is to attend to the limitation observed in previous studies on InfoVis user-centred design models, as presented in the review (see section 2.4.2, Chapter 2). As earlier highlighted, all the previous studies included prototyping as a single stage in the InfoVis development process without due elaboration and illustration. This is argued as a limitation to the needed practical guide in translating conceptual design framework to physical artefacts.

Following the typical software engineering models (Ariffin, 2009; Shneiderman & Plaisant, 2010; Rogers, Sharp & Preece, 2010; Syamsul, 2011), the process of developing HSDI is divided into three phases, namely; pre-production, production, and post-production. Each of these phases has its respective stages. The pre-development phase contains domain analysis and data collection, data preparation (which includes data cleaning and modelling), and users' tasks elicitation. The development phase contains InfoVis content development, coding and core system development, InfoVis features integration, and heuristics evaluation as formative evaluation. The last phase; post-development, has InfoVis deployment, InfoVis testing and users' experience evaluation which serves as the summative evaluation.

4.5.1 Phase 1: Pre-Development

The objective of this stage is to conceptualize the design framework for HSDI. This framework was applied in the InfoVis development which takes place in the development phase. This pre-development phase contains three different stages: Domain analysis and dataset collection, data preparation, and users' tasks elicitation.

4.5.1.1 Domain Analysis and Dataset Collection

This stage is to identify the need of the domain and the domain users so as to achieve a goal-oriented design. Also, for the development phase, the domain dataset is needed to help in the prototyping stage of the design. Domain analysis is the process of understanding the domain where the InfoVis will be applied. It encompasses identifying the explicit knowledge needed to support the decision making process in the domain work activities. Also, during the domain analysis, the InfoVis developer/researcher is expected to identify the work practices of the domain users and how InfoVis could be suitably deployed to support the decision making activities. The essence of this stage is strongly stressed by Lam et al. (2012), Robinson et al. (2005), Greenberg (2002), and Koh et al. (2011). In this study, it is named the pre-design study, as conducted. It is extensively discussed in section 4.2 (in chapter 4 of this thesis). The dataset collection characterizes the InfoVis developer/researcher's access to the domain dataset to be visualised. In cases where the datasets is cloud-based, access would be given to the cloud storage so as to experiment the visualization process. The datasets must be of the business value and conform to the operational activities of the to-be-designed InfoVis. This is done by ensuring that the data dimensions are in compliance with the identified explicit knowledge needed by the domain.

4.5.1.2 Data Preparation: Data Cleaning and Modelling

This stage is to obtain a befitting data structure that addresses the domain needs, as observed in the previous stage. Since this study is to achieve a more functional decision making support tool, it is expected that the currently obtainable structure of the data table would need cleaning, transformation and modelling –as necessary. Also, the InfoVis developer/researcher would model any missing data dimension or input missing values.

On another hand, structured datasets always demand certain preparations to meet the visualization process. This is actually due to certain unavoidable irregularity during the data collection process, human error, or data format disparity with the visualization toolkit to be employed by the developer. To solve this, data cleaning is always done. In data cleaning, the developer ensures all data dimensions are present in the dataset file or table(s) with logical connection. The data dimensions must be ensured to be within the acceptable rules. Data noises and NULL values are to be treated. This is compulsory according to interactive data visualization process methods suggested by Soukup and Davidson (2002), Ward et al. (2010), and Tan et al. (2006). For example, gender should be of two values; male and female, heights must be numeric, within a logical frame. In cases of observed data inconsistencies as suggested by the domain analysis, the developer transforms the datasets to a logical one. This characterises modelling the data to meet the domain business and operational requirements. In this study, the data cleaning and modelling is done through Microsoft Excel Spreadsheet.

4.5.1.3 Users' Tasks Elicitation

The objective of this stage is to conceptually elicit the tasks to be designed for, in the InfoVis. The users' tasks to be designed in the InfoVis allow considerable choice of the users' control mechanism. These tasks are essentially supported by the interaction and distortion techniques of the InfoVis' content delivery techniques. Tasks elicitation is an integral part of user-centred design model (Rogers et al., 2011; Sheinerderman & Plaisant, 2010). The users' tasks are earlier presented in Table 3.2.

4.5.2 Phase 2: Development

The objective of this phase is to actualize the physical product of the InfoVis conceptualised in the previous phase. This phase produced the HSDI, usable for the exploration and visualization of students' data to support the HEI policy making processes. To achieve this, four different stages are accomplished. These are: InfoVis component development, coding and core content development, InfoVis features integration, and heuristic evaluation.

4.5.2.1 InfoVis Component Development

The objective of this stage is to design the prototype, starting with a low-fidelity form, especially of its visual interactive interface (front-end). This can be done with paper sketching or prototyping software like *Justinmind* or *Balsamiq* mock-up. The InfoVis content development deals with the process of developing the content delivery techniques (as elicited in section 4.4, Table 4.7 and Figure 4.3) using the suitable programming language frameworks and libraries. This must be done to support related perceptual, cognitive and decision supports, with respective theoretical considerations. This is as earlier discussed in section 2.4, and highlighted in section 4.6.1. Developing InfoVis content entails writing the detail of the InfoVis techniques that will be developed in respect to the users' tasks.

4.5.2.2 Coding and Core System Development

The objective of the stage is to implement the software development technologies so as to bring the conceptualized design and low-fi prototype into its real state. This stage involves the core technical aspect of InfoVis development. At this stage, each of the content delivery technique is translated into rules which are implemented with their respective frameworks, libraries and codes. At this stage, and typical of an InfoVis, the backend that houses the data was designed, using appropriate database technology, or the prepared data file. The front-end is designed with much attention to the theoretical suggestion for InfoVis interface design to aid perception, cognition and decision making (Ware, 2000; Healey, 1996; Fekete et al., 2008; Teets Tegarden, & Russell, 2010; Semiu & Zulikha, 2015a).

4.5.2.3 InfoVis Features Integration

This stage is to perfect the prototyping process, by bringing all the peculiar InfoVis characteristics into reality. Developers can use any back-end technology like MySQL, and PHP library. Since HSDI will be web-based, Hyper Text Markup Language (HTML) is used for the body structure and Cascading Style Sheet (CSS) for the rendering. Data Object Manipulation (DOM) functionality and Data-Driven documents; D^3, which is also a JavaScript library for creating data visualization, are used for the data binding and visualization. Closure, a new Google-created JavaScript library is used for the interface element and features design. At this stage, the core InfoVis features –Overview, filter, demand detail, and zooming –as earlier listed is integrated, and their working is perfected. Immediately after the development phase, a review of the technical aspects of the HSDI is done to actualise a high fidelity prototype of HSDI.

4.5.2.4 Heuristic Evaluation

The objective of this stage is to receive feedback from usability/HCI and data analysts/administrators. This is the formative evaluation, as highlighted in section 3.7.3. The heuristics evaluation concentrates on data representation correctness and usability issues of the InfoVis prototype, using the designed cognitive and decision making tasks as a guide. The input from the heuristics was used to improve both the proposed conceptual design framework and the prototype developed.

4.5.3 Phase 3: Post-Development

The objective of this phase is to review the technical aspect of HSDI and ensure that its workings and operations are in order. This was done through a cross-browser assessment. Also, software engineering approach of system checking for debugging and error tolerance control are implemented. Lastly, the summative evaluation which is the users' experience evaluation experiment is conducted. The participants were the domain users; these are the HEI policy and decision makers, as earlier described.

4.5.3.1 InfoVis Deployment

The objective of this stage is to confirm the compatibility of the HSDI design with the major web browsers. At this stage, HSDI was deployed over cross-browser platform to check its workability with certain numbers of web browsers. The frontline web browsers cross-checked with are Mozilla Firefox, and Google Chrome because of their high usage statistics[1].

4.5.3.2 InfoVis Testing

The technical testing of all the core functions of an InfoVis is done at this stage, and ensured that their workings are perfect for usage. This testing stage is where each of the core functions of the InfoVis is tested to check for bugs and errors. The InfoVis core functions are taken as the users' tasks. Therefore, Overview, with details, zooming, filtering and mouse-touch interactivity, scrolling are tested and the underlying code is thoroughly debugged. This is done in conjunction with software engineers and InfoVis developers.

[1] W3S Browser Statistics, See: http://www.w3schools.com/browsers/browsers_stats.asp

4.5.3.3 Users' Experience Evaluation Study

This final evaluation of the HSDI is done at this stage. This is the users' evaluation study of the HSDI. As discussed in Chapter 3 of this thesis, at this stage, the correct data representation, usability and decision support of HSDI are evaluated by the domain users which are the HEI policy makers. In a nutshell, the overall process of developing HSDI is presented in Figure 4.4.

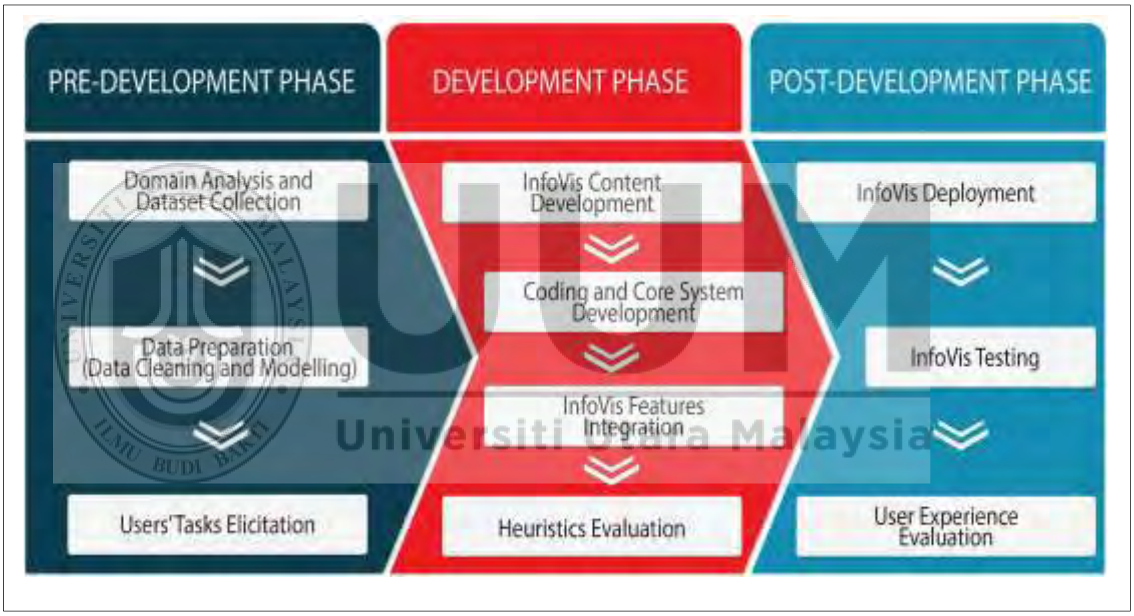


Figure 4.4. Process Model for HSDI

4.5.4 Theoretical Model for Designing HSDI

Based on a critical analysis of InfoVis related theories, presented in Table 2.3 in chapter 2 of this thesis, a theoretical model for designing HSDI was proposed. Preattentive processing theory (for perception), cognitive fit theory (for cognition) and normative and descriptive theories (for decision making) are used. These theories are chosen basically because they present more comprehensive design dictates for their respective supports.

The proposed theoretical model presents the linkage between InfoVis' theories (Preattentive processing theory, cognitive-fit theory, and normative and descriptive theories), its supports (Perceptual, Cognition and Decision support), the respective design features (as listed in Table 2.3, chapter 2) and the involved human-computer activities (visualization, data exploration and interaction). The model states that the explanations of Preattentive processing theory, cognitive-fit theory, and normative and descriptive theories must influence the design of visualization, interaction, and data exploration activities to achieve InfoVis' perceptual, cognitive and decision supports, respectively. The collectivist effect of these theories-based and human computer interaction activities is to actualize InfoVis' decision support. Figure 4.5 presents the theoretical model for designing InfoVis generally, and HSDI in particular.

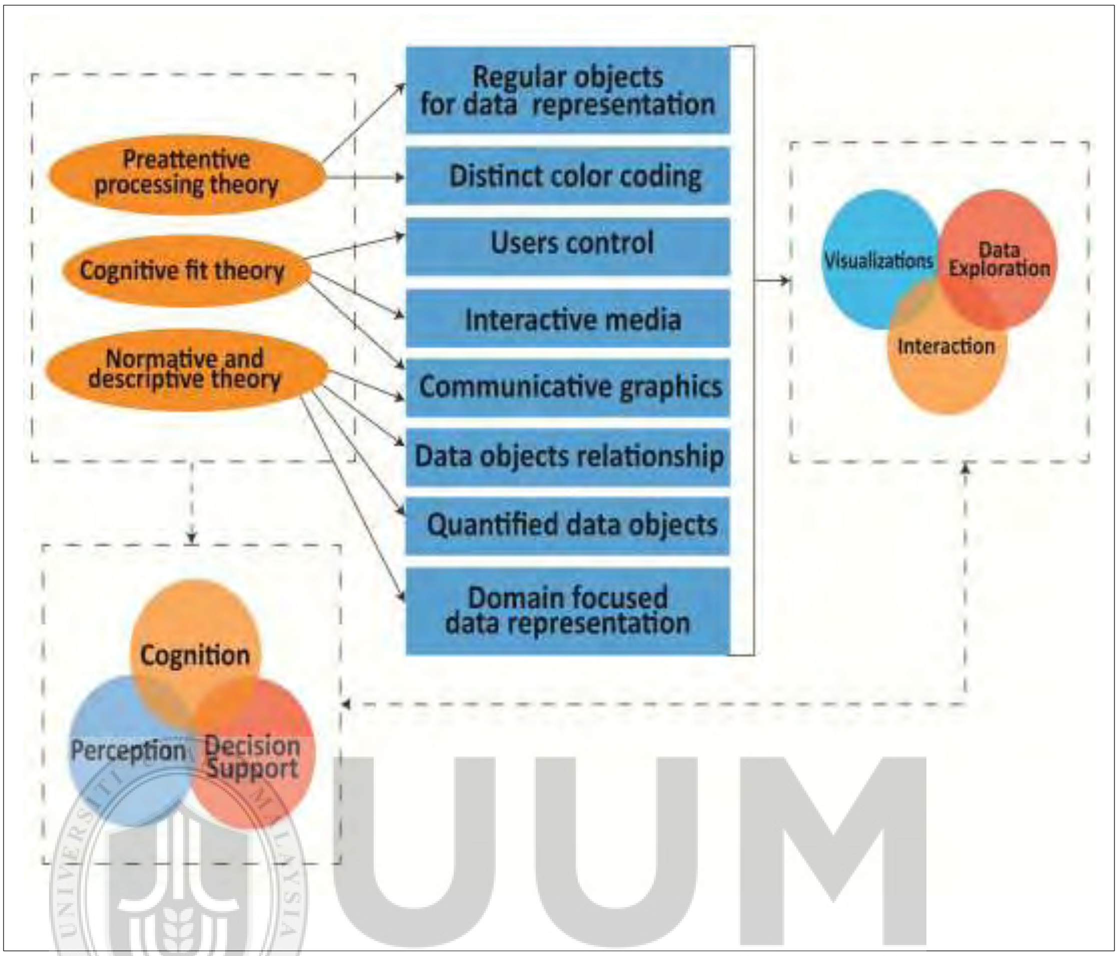


Figure 4.5: Theoretical Model for Designing InfoVis

4.6 The Proposed HSDI Conceptual Design Framework

The HSDI conceptual design framework presents the combination and interrelationship between the HSDI content delivery techniques and the theoretical design model. This conceptual design model illustrates the design guide for HSDI. The translation of this design guide to a physical artefact, i.e. the prototype, is to be done through the process model (presented in Figure 4.4). The theoretical model (earlier presented in Figure 4.5) illustrates the connections and interdependence between the design theories and the supports that InfoVis must be designed for. The theoretical provision of preattentive processing theory is to ensure perception, cognitive fit theory for cognition and normative and descriptive theory for decision support. Each of these supports is to be supported by Human-InfoVis activities which are visualization, interaction and data exploration respectively. The implications of theories on InfoVis design are earlier discussed in section 2.4.4 (in chapter 2 of this thesis). The theoretical influence, specifically in designing HSDI, is further discussed in Chapter 5 (section 5.4).

The content delivery techniques model handles the HEI students' data exploration by using both the interaction and distortion techniques. The interaction techniques are mouse-on, mouse-on click, drop down-menu, dynamic cursor hinting, mouse on-drag, check boxes, and push button. The distortion techniques are scrolling, overview with details, zooming and exploration. Their actions often overlap with the interaction techniques; hence, they are hardly separable. Spence (2007) classified both distortion and interaction techniques as presentation styles, and these techniques handle the behaviour of HSDI. Figure 4.6 presents the proposed HSDI conceptual design framework.

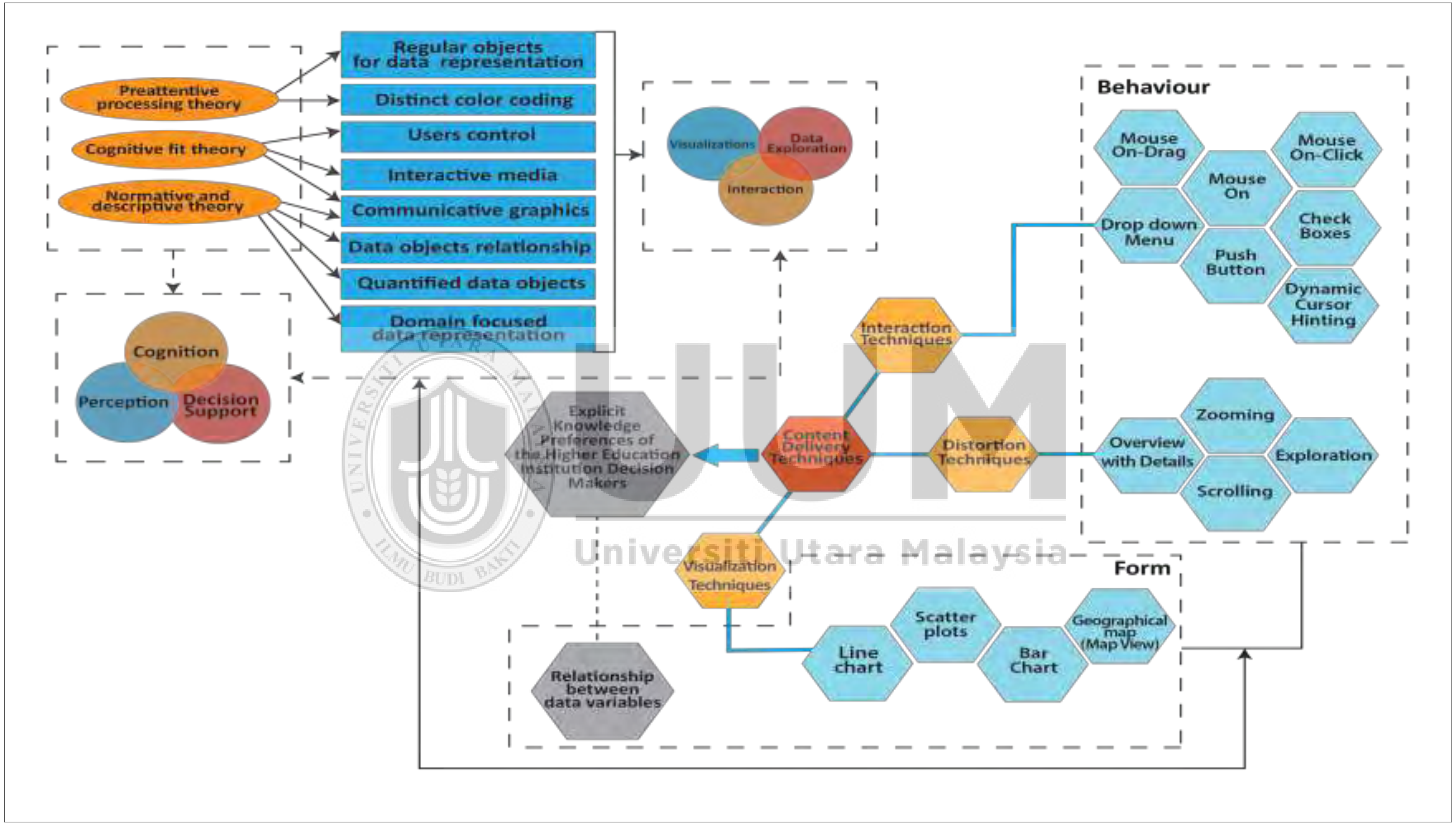


Figure 4.6. Proposed HSDI Conceptual Design Framework (Initial)

The visualization techniques are scatter plot, line chart, map view, and bar chart. These are the techniques that handle the representation of the relationship of the data dimensions. Therefore, the InfoVis techniques (interaction, distortion, and visualization) collectively handle the HEI students' data exploration in view of delivering the validated HEI domain explicit knowledge preferences.

4.7 Summary of the Chapter

This chapter firstly reports the pre-design study conducted at the domain analysis phase of this study. Essentially, the findings of the pre-design study identified the explicit knowledge preferences involved in the HEI students' data-focused decision making process. This outcome accomplishes the first objective of this research. The HSDI structural components which show the interacting tiers in InfoVis are presented. Also, the process of designing HSDI, with significant attention to software engineering principles, and illustration of its phases and stages are highlighted.

From the content delivery techniques, the visualization, interaction and distortion techniques for the representation, presentation and exploration of the domain explicit knowledge preferences are elicited. This accomplishes the second objective of this research. Lastly, the development of the HSDI conceptual design framework achieves the third objective of this study. All the components for the HEI data exploration are designed in view of solving the domain problem of information overload which is arguably caused by the multidimensionality of the students' data. This process is given a scholarly footing in the review of related literatures presented in Chapter 2.

In furtherance of the objectives set by this study in Chapter 1, expert review is employed to verify the proposed conceptual design framework, and prototyping is used as a proof of concept. Heuristics evaluation is also employed as a formative evaluation process for the HSDI. The next chapter; Chapter 5, discusses the expert review, prototyping and heuristics evaluation. The physical artefact which is the product of the prototyping stage of the HSDI conceptual design framework is finally evaluated through users' evaluation study. The findings of the users' evaluation study are presented and discussed in Chapter 6 of this thesis.

CHAPTER FIVE
EXPERT REVIEW, PROTOTYPE DEVELOPMENT AND HEURISTIC EVALUATION

5.1 Introduction

This chapter entails the formative evaluation of the proposed conceptual design framework. The expert review and their findings are reported. The prototype, which serves as a proof of concept, is also described. The prototyping process validates the praticability and applicability of the proposed framework. The proposed process model is used in developing the HSDI prototype, with the proposed conceptual design framework as the content design guide. The prototype is ensured to comply with both the structural and content components of the proposed framework. Lastly, the prototype was examined through heuristic evaluation. This is to also determine the suitability of the proposed framework in the design and development of HEI students' data-focused InfoVis, and its functionalities in terms of usability and data representation correctness. Lastly, feedback from the formative evaluation processes was applied to both the conceptual design framework and the prototype for the required modifications.

5.2 Expert Review

The expert review verified the appropriateness of the proposed HSDI conceptual design framework which is the process and content delivery techniques models. It verified the implementation process of the proposed framework and assures it can lead to the development of a befitting InfoVis tool. The expert review was done by considering the appropriateness of the proposed framework using the measurement and dimensions adapted from Syamsul (2011) (earlier described in section 3.7.2).

5.2.1 Methods and Instruments

Five experts whose areas of specialties cut across business intelligence, software engineering, and InfoVis are involved (refer to Appendix H for the profile of the expert reviewers). The adaptation process of the expert review instrument and its validity is described in Chapter 3 (Section 3.7.2).

5.2.2 Findings from Expert Review

The findings from the expert review exercise are presented in Table 5.1. As shown in the table, the mean values of the experts' findings showed they all align with "Strongly Agree." The mean values are Expert A (8.67), Expert B (8.91), Expert C (9.00), and Expert D (8.73), and Expert E (8.66).

Table 5.1

Mean Values of the Expert Review Findings of HSDI Conceptual Design Framework

Dimensions	**Expert A**	**Expert B**	**Expert C**	**Expert D**	**Expert E**	**Mean**
Compatibility	8.6	9.0	9.4	8.6	8.4	8.80
Visibility	8.5	8.7	9.0	8.5	8.7	8.75
Simplicity	9.0	8.6	9.2	9.2	8.8	8.96
Effectiveness	8.8	8.2	9.0	9.6	9.0	8.92
Evolutionary	9.2	9.6	8.7	8.4	9.0	**8.98**
Flexibility	9.0	8.5	9.5	9.0	8.8	8.96
Clarity	8.8	9.2	9.3	8.4	8.8	8.90
Manageability	7.5	9.5	8.0	8.2	7.8	**8.20**
Mean (Each Expert)	8.67	8.91	9.00	8.73	8.66	8.80

All scores for the dimensions are relatively high with at least 8.20 out of the rank of 10. The highest mean score among the dimensions is evolutionary with 8.98, and the lowest is manageability with 8.20. This implies that users' input and experts' opinion are continuously sought during the design process, and thus allows incremental changes to cope with domain realities. This also implies that the proposed design framework has weaker points in its ability of self-monitoring, self-explanation and self-instruction. It means its provision of enough details to monitor the process and stage of the developmental lifecycle is not as strong as other dimensions.

From the mean scores presented in Table 5.1, flexibility and simplicity are both ranked 8.96. This implies that the conceptual framework is can be easily amenable to fit into different domain and InfoVis design project, and also simple to understand. Effectiveness of the conceptual framework is scored 8.92, clarity is 8.90, compatibility is 8.80, and visibility is 8.75. This highlights the order of quality strength of the conceptual framework. It implies that the framework's ability to determine the completeness of the InfoVis design process is lesser than its consistency with prior experiences from InfoVis design project. The quality of the framework productivity is however seen to be stronger than the clarity of the design phases. In sum, the estimate mean of 8.80 for the totality of the dimensions suggest that the proposed design framework is satisfactory.

Also, Figure 5.1 clearly illustrates a radar graph which gives further valuable illustration to the means scores on expert-by-expert basis. It shows that none of the expert gave highest all-round score to the dimensions. Expert C gave the highest score to flexibility, clarity, simplicity, visibility, and compatibility. Expert B also gave highest scores to evolutionary and manageability where other experts' scores rank copiously lesser than others. There is no consistent pattern that could relate the experts' scores with either their core areas of specialization or years of experience.

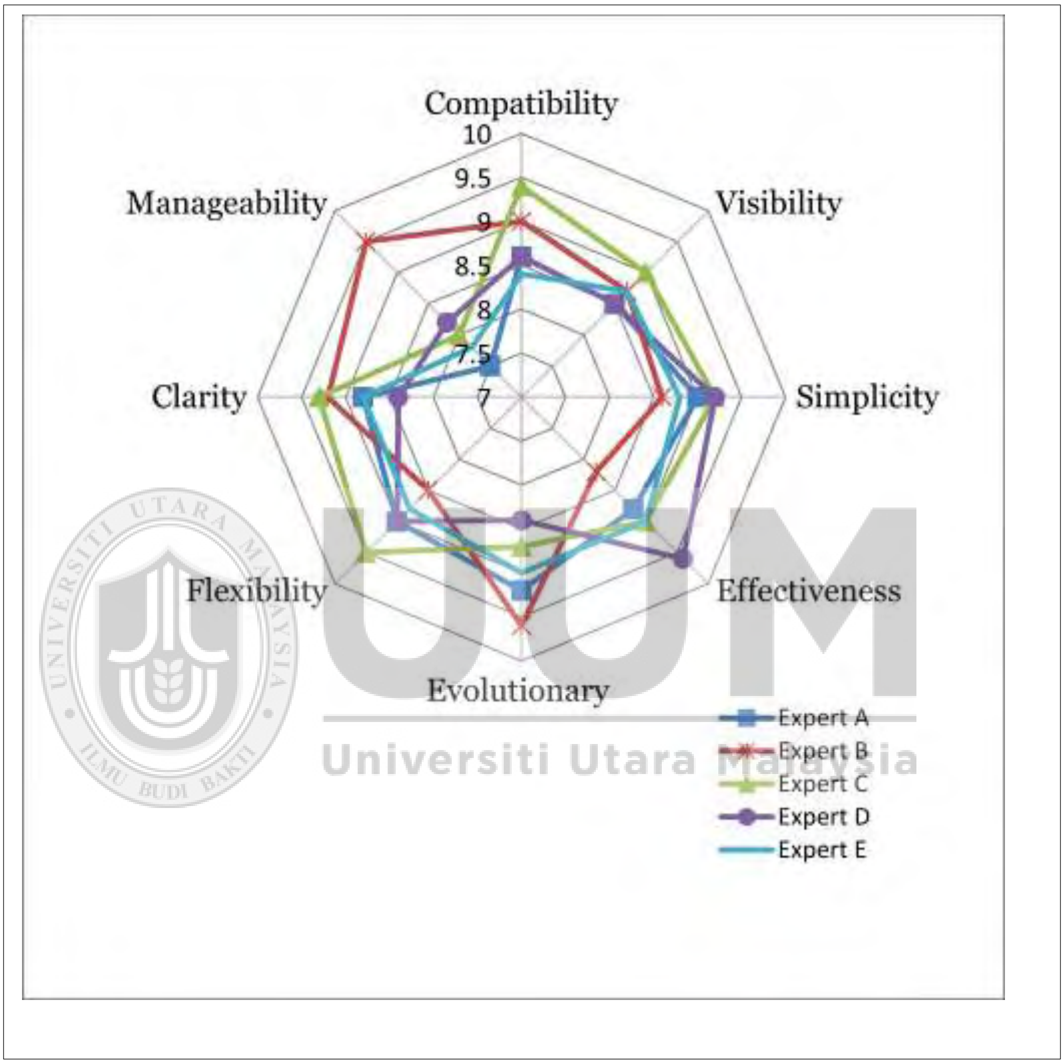


Figure 5.1. Radar Graph for the Expert Review Findings

Overall, the proposed conceptual design framework is verified to be satisfactory and certified for prototype development. The important feedback from the expert is the need to explicitly state that the data preparation stage (i.e. second stage in the first phase of the process model) is exclusively adoptable or adaptable in structured data visualization.

5.3 HSDI Prototype Design and Development

The design and development phases of the HSDI prototype employed the proposed HSDI conceptual design framework, as described in Chapter 4. The framework has structural, content delivery technique and process models which are all needed for the development of the HSDI. The HSDI system is an InfoVis system that aids the exploration and visualization of the HEI students' data, by using interactive visual interface. It is developed to capture the domain users' explicit knowledge preferences during any students' data exploration activity, and to support their data-driven decision making process.

5.3.1 Pre-Production Phase

This phase consists of three different stages. In this phase, the explicit knowledge preferences for the data-driven decision making need of the HEI policy makers were identified. Based on this, the needed business data sets was extracted from the institution operational database, and prepared through respective data cleaning, transformation and modelling. The users' tasks, which are defined as necessary activities to be carried out by the HEI policy makers while interacting with the business dataset, are elicited (as earlier presented in Chapter 3, section 3.7.4).

5.3.1.1 Domain Analysis and Data Collection

The domain analysis which was done during the pre-design study is exclusively covered in section 4.2 in Chapter 4. The most important deliverables of this stage are the explicit knowledge preferences, from which the HEI data usage framework (Access, Monitor, & Develop) and the appropriate students' data variables for a comprehensive data analysis and visualization are formed (Semiu & Zulikha, 2015c). Therefore, for the purpose of the HSDI prototype development, a file of business dataset that is close to the outlined students' data variables is requested and collected.

5.3.1.2 Data Preparation: Data Cleaning, Transformation and Modelling

The available datasets do not absolutely meet the data model needed for the development of HSDI in terms of the comprehensive list of the data dimensions, therefore data preparation is done. This involves treatment of noise and NULL values in datasets that are extracted from operational datasets. At this stage, data cleaning and transformation are done so as to achieve the befitting data model that will be suitable for the development of the HSDI prototype.

Data Cleaning

A check performed on the dataset (57368 rows, 9 columns) showed the data types (ordinal and categorical) are consistent with data need, and it represents an experimental unit. Experimental unit as an important property of exploratory data table describes the granularity of facts about the subject (Soukup & Davidson, 2002). However, the following irregularities are observed: Missing Values, Inconsistent Labelling, Insignificant columns, and Duplicate columns.

a. Missing Values: These are data cases that have no values provided. When necessary, the mean value (maximum value + minimum value/2) of the data column is used to fill in the missing cases. When this is not applicable, like categorical data value, values are created in a manner that resonates with reality and practicality (e.g. If “Race” value is missing for Nationality that is of “Ghana” value. It is practically taken to be “African”).

b. Inconsistent Labelling: This is when columns defined in the collected dataset are inconsistent with the dictates of the pre-design findings. Though little disparity is often experienced. To solve this, some new columns are created and labels are provided to be in line with pre-design findings’ data definition.

c. Insignificant columns: These are columns that are not needed as data variables (as earlier defined in Table 4.4) for the HSDI development. In this case, the students’ names column is observed. It is therefore deleted.

d. Duplicate columns: These are columns that are repeated, either explicitly or implicitly, in the dataset. In this case, the program’s code is observed. It is an implicit repetition since the program name is already included. It is therefore deleted.

Data Transformation

a. Handling NULL values: Data cases that must contain NULL values, like disability, are provided using the applicable SQL program code.

b. Splitting Columns that handled more than one variable: From the dataset collected, columns that contain more than one variable, as presently defined in the data model, are split.

c. Imputation: This is the process of imputing data to newly created data columns or missing values. This is also done using the applicable SQL program code.

5.3.1.3 Users' Tasks Elicitation

The tasks to be carried out by the HEI policy makers while interacting with the HSDI prototype are elicited for proper documentation and befitting development of the HSDI prototype. These users' tasks are: locate, identify, distinguish, categorize, distribute, compare, associate and correlate. The descriptions of these tasks have been earlier presented in Table 3.2. As an InfoVis, the HSDI prototype is developed so as to support the elicited users' tasks. The users' ability to carry out these tasks is a prerequisite to data-driven decision making support which is the utmost functionality of an InfoVis.

5.3.2 Production Phase

This phase consists of four different stages. These are InfoVis content development, coding and core system development, InfoVis features integration, and the heuristics evaluation. The physical artefact is developed in this stage, purposively to address the earlier identified problem of information overload caused by the multidimensionality of the HEI students' data, and the subsequent data-driven decision making constraint.

5.3.2.1 InfoVis Component Development

A low fidelity prototype of HSDI interface is designed. The placement and actions of the visualization, distortion, and interaction techniques as they relate with the interface elements and forms are outlined. The low-fi interface design is presented in Figures 5.2. Table 5.3 presents the guide for the coding and core content development (section 5.3.2.2) to ensure that HSDI development follows the proposed conceptual design framework and the underlying theoretical principles.

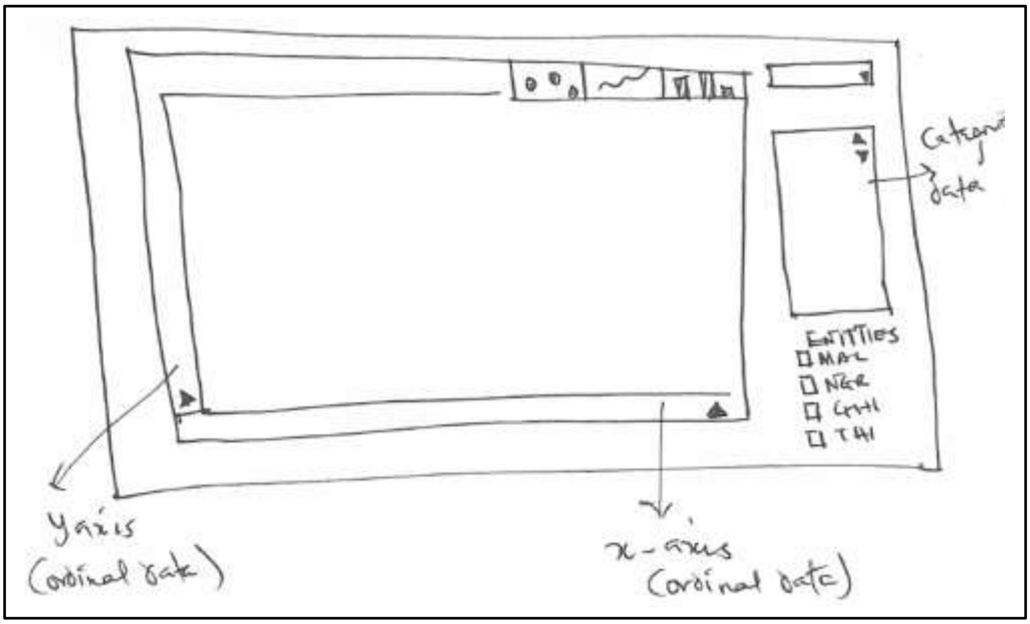


Figure 5.2. Hand-Sketched Interface for HSDI prototype

Table 5.2

Theoretical and Design Guide for HSDI Coding and Content Development

Theories and Design Steps	HSDI Characteristics and Functionalities
Preattentive processing theory	HSDI allows users' control of the colour types of its visualization entities. With this, especially when the user chooses 'unique colours', the statistical graphics (scatter plots, line chart and bar chart) are rendered in different colours and this aids pre-attention. Also, with a pliant-hinting as an interaction technique, the details of the section of the rendered data are further shown and this captures the users' visual perception. It thus aids preattentive processing of such section that is mouse-hovered by the user.
Cognitive fit theory	The users can pick their preferred data dimensions, at a specific time, during the data exploration process. This can be done through the drop-down menu and check-boxes. With this, handling the multidimensionality is simplified

	because user can pick interested manageable dimensions at a particular time, and establish their relationship through the chosen visualization technique. This reduces the cognitive load.
Normative Theory	The y and x axes relationship and the pattern of the statistical graphics allow the user to identify any relationship between ordinal data dimensions. The ability to correlate and associate the data dimensions presents the explicit knowledge. It also allows other categorical dimensions to be added, and thus allows HSDI to present more than 2D at a glance.
Descriptive theory	HSDI allows the ordinal and categorical data dimensions to be quantifiable, and this forms part of the presented explicit knowledge (through the $y - x$ axes representation).
Data Dimensions and Weighting	The data dimensions are classified into basic and derived. These are highlighted and listed in Table 4.5, and also included is the weighting of the derived dimensions. The weighting is the quantifying formulas conceptualised by the researcher for the derived data dimensions following logical and real life scenarios.

5.3.2.2 Coding and Core Content Development

The Google API framework for interactive chart and the MySQL relational database as the database are built using the appropriate libraries that are further explained in section 5.3.2.3. System architecture, which is a conceptual explanation of the structure and behavior of the system, is diagrammatically presented in Figure 5.3.

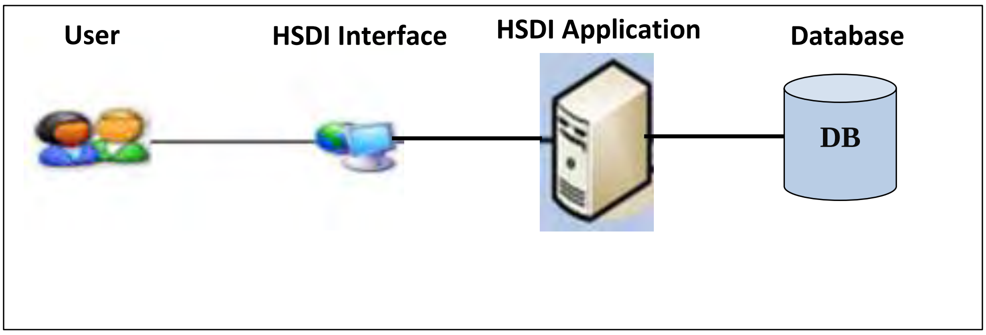


Figure 5.3. System Architecture for HSDI prototype

The system architecture shows four different interdependent elements. These are (a) the users, (b) HSDI interface, (c) HSDI application, and (d) the database. The users are primarily the HEI decision makers or anyone who is interested in exploring the students' data for insights. The HSDI interface is the visual-exploratory and graphical user interface where the content delivery is done and users can distort, interact and visualize in line with their tasks' preferences. The HSDI application is the data object mapping and binding application that matches the data to be visualized with the appropriate visualization object and rendering mechanism for the interface to deliver. Lastly, the database is the storage medium of the data where data is retrieved during the data exploration exercise. A related description of this is earlier made in the structural model of HSDI in Figure 4.1.

5.3.2.3 InfoVis Features Integration

All the InfoVis features, essentially the interactivity and ability to render the data in a way that support perception, are integrated. HTML, CSS, JavaScript are the major programming language, and PHP for needed database connection. The screen shots of HSDI and their respective explanations are provided in section 5.4.

5.3.2.4 Heuristics Evaluation

Experts in usability and data analysis/administration are selected to evaluate the prototype developed. The evaluation is specifically on the usability of the prototype and its data representation correctness. The detail of this stage is further presented in section 5.5.

5.3.3 Post-Production Phase

The main activities of this phase are deployment of the prototype designed over two different browsers, overall debugging and the user evaluation study. At this phase, the InfoVis deployment, InfoVis testing, and users' experience evaluation is performed. At the InfoVis deployment stage, the InfoVis prototype is deployed across Mozilla Firefox and Google Chrome. It works better with Mozilla Firefox than Google Chrome. This is because Adobe Plug-in, which is an adds-on extension, works better in Firefox than Google chrome

Then, the InfoVis was tested and checked for bugs and errors. Its core functions which are overview, with details, zooming, filtering and mouse-touch-and-click interactivity, scrolling are equally tested. At the final stage, the users' experience evaluation, the end-users which were HEI policy/decision makers are the participants of the users' evaluation study. The details of this sample are described in section 3.7.4, and summative evaluation and its findings are presented in chapter 6 of this study.

5.4 Screenshots of HSDI Prototype

The following screen shots in Figure 5.4 to Figure 5.14 are the user interfaces of HSDI (given a persona name *StudentViz*). When the web-based system is accessed through its URL, the landing page (shown in Figure 5.3) shows the introductory message about the system. It informs the users of the value proposition of the system, its functions, and contains navigation links to the visualization pages.

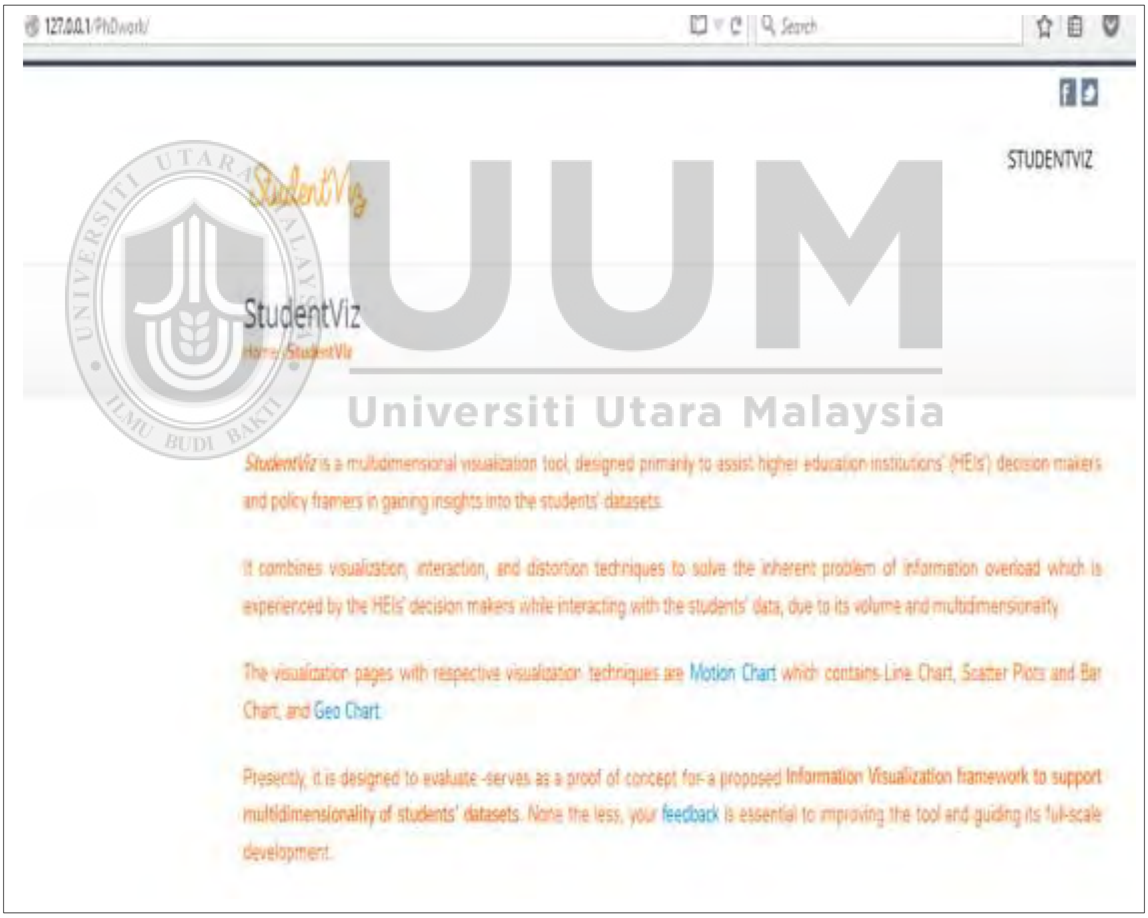


Figure 5.4. Landing Page for HSDI Prototype

The visualization pages –as shown on the landing page – are Motion chart and Geo Chart. The Geo chart (shown in Figure 5.5) is the name given to the Map view, and the motion chart contains three different visualization techniques, namely; Scatter plot, Bar chart and Line Chart. This makes a total of four visualization techniques as proposed by the conceptual design framework.

Figure 5.5. Geochart (Map View) of HSDI Regions that Exhibit Highest Population Growth

Geo chart gives a population distribution of the students based on their nationalities. The geo chart (visualization technique) is used with dynamic cursor hinting and colour scheming (interaction techniques) to display the HEI policy makers' explicit knowledge preferences in terms of the current population of the students on nationality-basis.

The nationalities are graphically represented by their respective geographic map within the conventional world map. The brightness/contrast of the colour of each of maps is proportional to the nationality's students' population. The user needs to dynamically hint (with mouse-on event) the map of the country of interest for its details (the country name and the students' population size). The display of the population size as quantification of the colour radiance of the dynamically hinted area is influenced by the descriptive theory. At a glance, before the dynamic hinting of an interested area to reveal the detail, the difference in the colour brightness gives the user a clue of the difference in their population size. This colour difference concept is influenced by preattentive processing theory.

The other visualization technique used by this study is the motion chart which contains scatter plot, bar chart, and line chart. At the upper part (as shown in Figure 5.5) of the motion chart interface, an information panel about the data dimensions is embedded. The data dimensions' definitions and weighting are provided to inform users of the measurement used for each of the data dimensions. Figure 5.7 shows the larger part of the information panel which will be shown to the user when scrolled up.

Figure 5.6. Highlight of the Data Dimensions' Definitions Page

Motion Chart/Data Dimension Definitions

Motion Chart showcases Scatter plots, Line chart and Bar Chart with the ability to dynamically change from one to another, at user's click.

While users are assumed to know the definition of data dimensions like gender, disability and Cumulative Gross Point Average (CGPA), the newly-conceptualized and specifically-defined dimensions are listed for clarity sake.

These are:

Year of Entry (Time): This is given and labelled as Time on the axes. It marks the student's year of enrolment.

Employability: This is defined as the number of months that the student waited before being gainfully employed.

English Language Proficiency: The extent at which the student can proficiently speak and write English Language. This is measured by English Language proficiency examination taken. The students that are exempted from these examinations are perceived to be the best.

Deferment Length: The number of semester that the student has deferred his or her study.

Figure 5.7. Data Dimensions' Definitions Page

Scrolling as a distortion technique is used in accessing full information about the data dimensions. Figure 5.8 depicts some interaction techniques included in the development of HSDI and their respective distortion techniques (shown in Figures 5.9, 5.10 and 5.11). Figure 5.12 and 5.13 respectively show other visualization techniques, i.e. Line chart and Bar Chart.

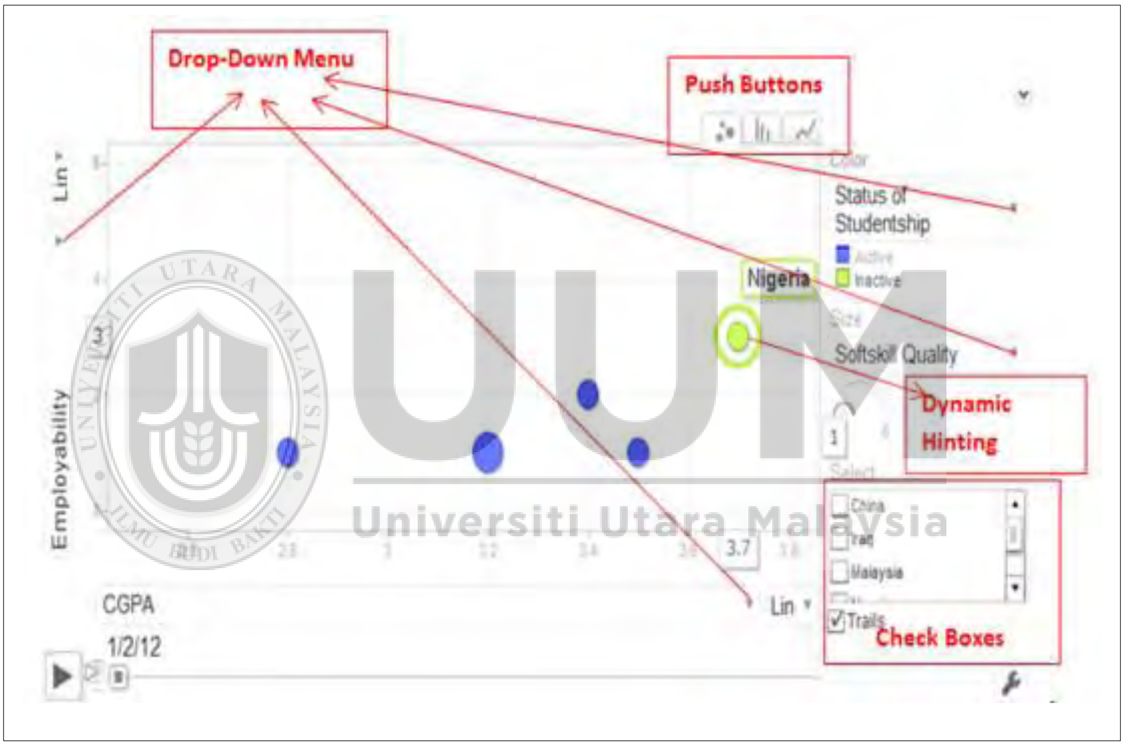


Figure 5.8. Interaction Techniques in HSDI

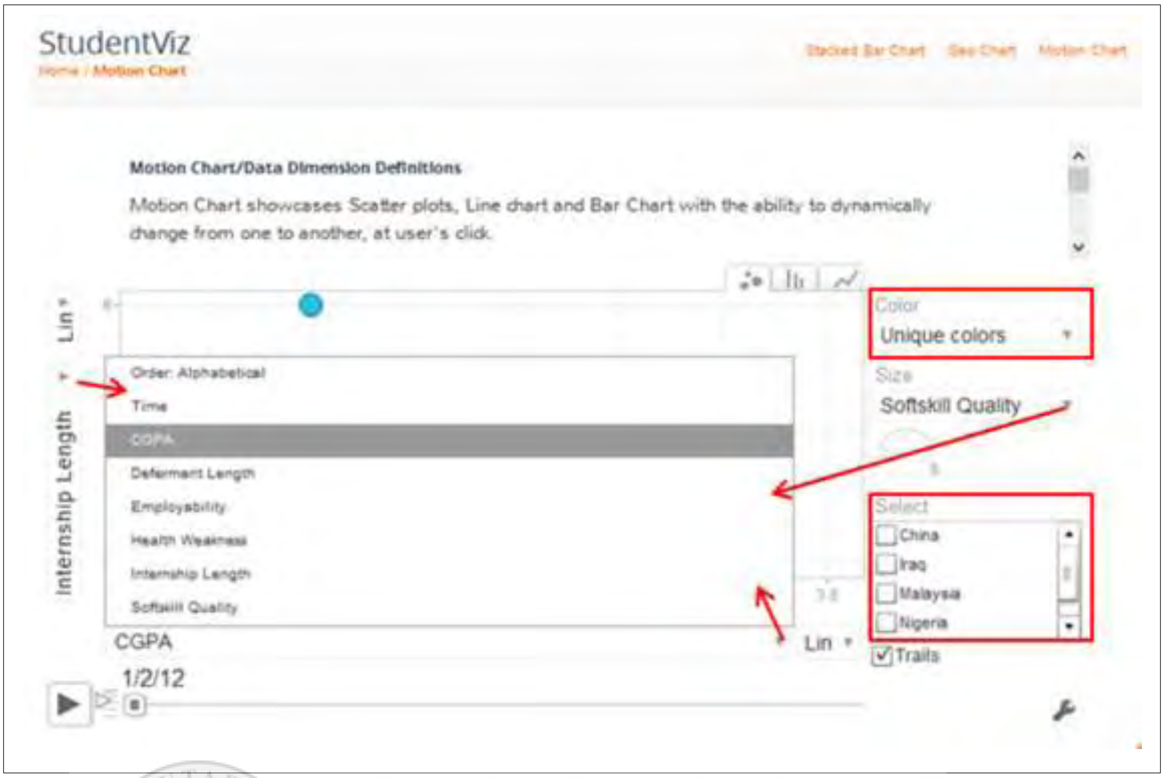


Figure 5.9. HSDI Interface (with Drop-down menu for exploration) shows the Data Variables

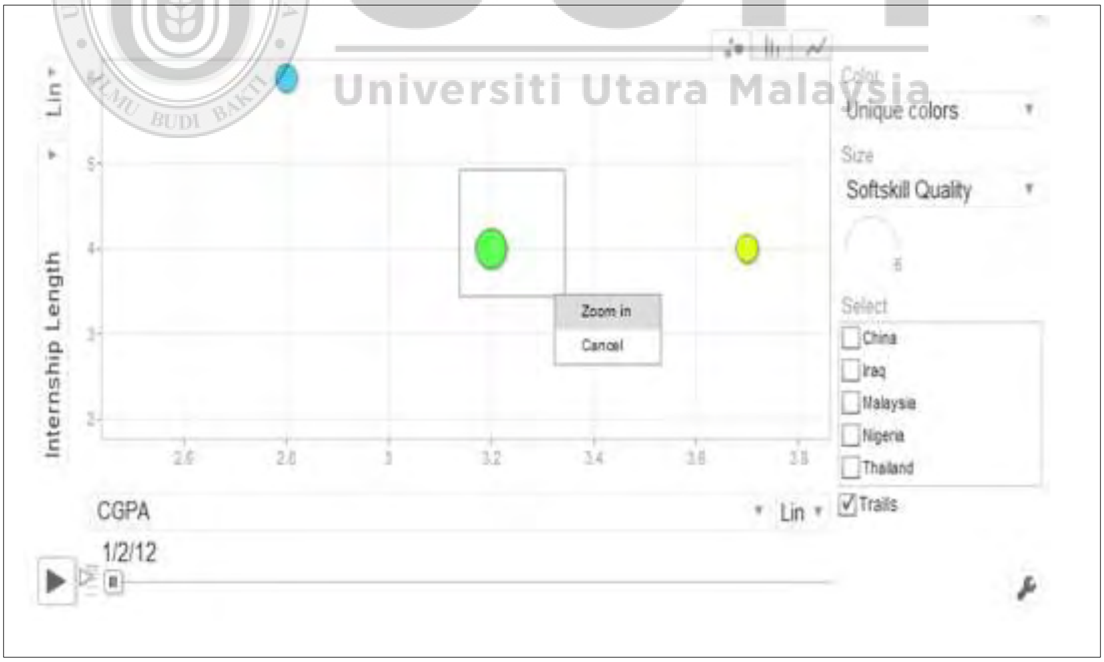


Figure 5.10. HSDI Interface (with Zooming Functions) activated by Mouse-on Drag Technique

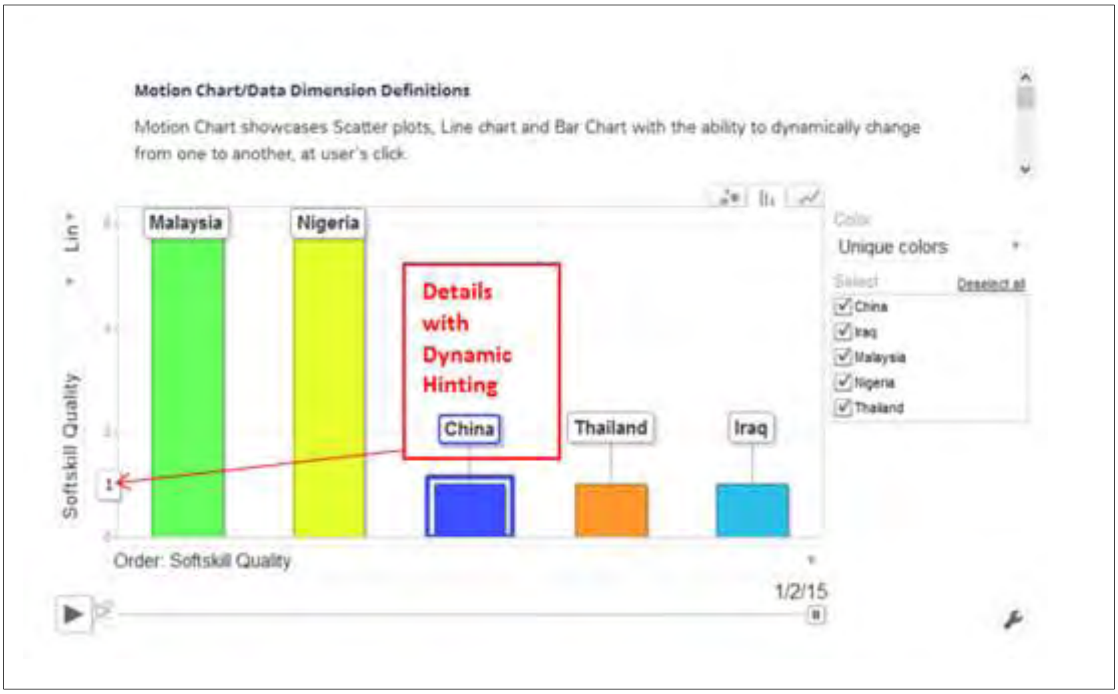


Figure 5.11. HSDI Interface showing Overview (Details with Dynamic Hinting) of the selected data

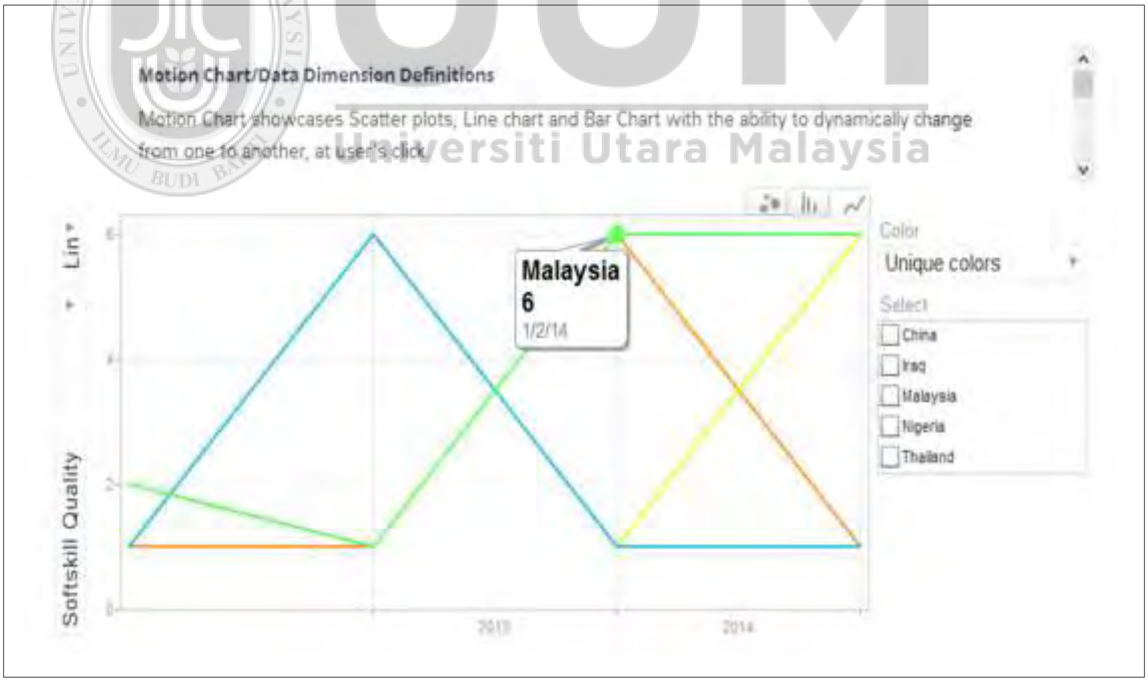


Figure 5.12. HSDI Interface showing Line Chart as a Visualization Technique

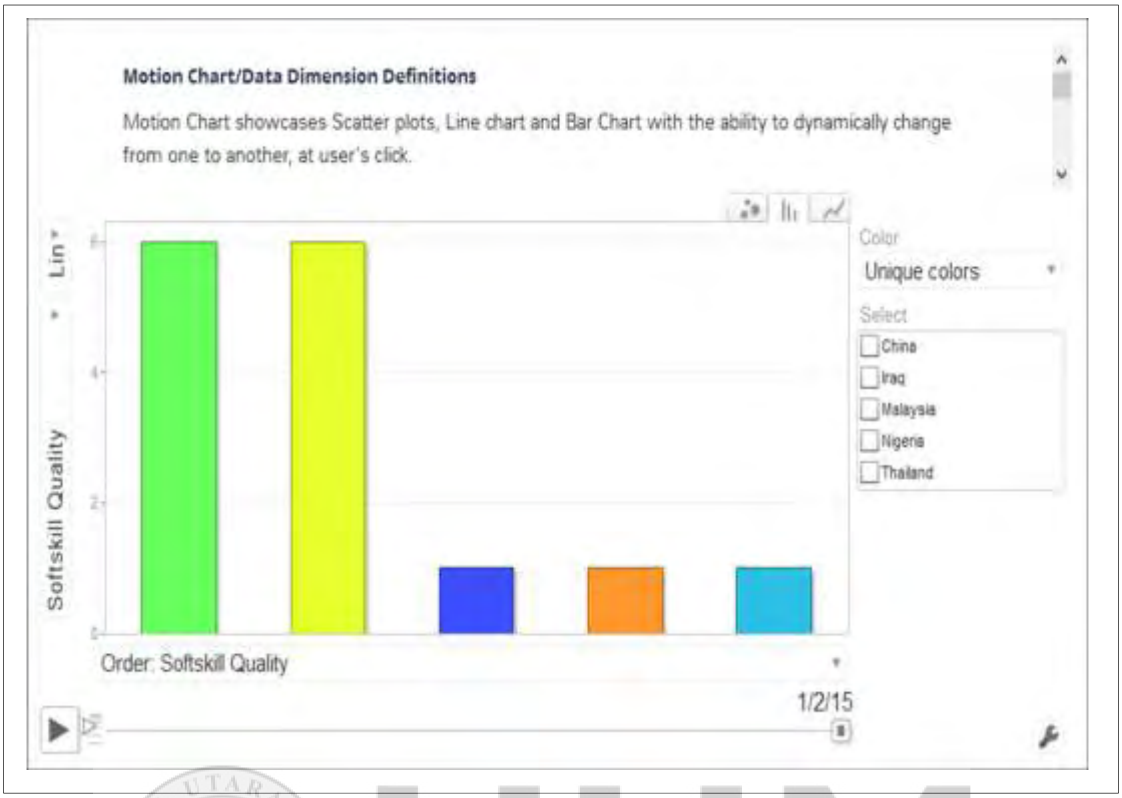


Figure 5.13. HSDI Interface showing Bar Chart as a Visualization Technique

As normal of any InfoVis, the user takes the control of the data exploration and the choice of the visualization depending on the users' interest. However, a knowledge-based choice will achieve better users' experience and task fulfilment. The interaction and distortion techniques (as earlier graphically illustrated and discussed) are employed with the chosen visualization to convey the preferred explicit knowledge preferences to the users. Therefore, the major human-InfoVis activities during usage are visualizations, interaction and data exploration. As shown in Figure 5.14, as an illustration, the user picks 'Internship Length' as the data dimension of interest on the y-axis, and CGPA on the x-axis. The y and x axes contain all the ordinal data dimensions listed in Table 4.4, the ordinal and categorical data dimensions can be picked using the left-side panel. In the example in Figure 5.14, the user picks soft skill quality, an ordinal data dimension.

Therefore, through data exploration, the usual cognitive load caused by multidimensionality of the students' data is reduced because the user can pick the dimensions of interest while disengaging others. This follows the cognitive fit theory. HSDI provides the users' control mechanism for the conveyance of the explicit knowledge. The ability to connect a data dimension with the other is suggested by normative theory. HSDI also allows the third (or more) data dimension to be represented using size, as suggested by descriptive theory (an example with Soft skill quality in Figure 5.14). The different colours are used for the representation of the entities and the reaction to dynamic hinting support preattentive processing theory. In sum, all the expected InfoVis' supports are designed for in accordance to the dictates of their respective underlying theories. The proposed conceptual design framework is applicable in the development of InfoVis and for students' data specifically, but adaptable to related data dimension structure.

Figure 5.14. HSDI Interface showing Rendered Scatter plots as a Visualization Technique

5.5 Heuristics Evaluation of HSDI

The heuristics evaluates the HSDI prototype in terms of its usability and data representation correctness. As earlier mentioned, each of these dimensions evaluates its respective heuristics: Usability experts for usability, and data administrator/analyst for data representation correctness. This verifies the listed qualities of HSDI from an expertise perspective. The heuristics evaluation is done through an evaluation framework and its applicable measurement and dimensions (Tory & Moller, 2008). These dimensions (usability and data representation correctness) are discussed in sections 3.6.3 and 3.7.3. The evaluation sessions were held on 17 and 24, August, 2015 for the data analyst/administrator and usability experts respectively.

5.5.1 Methods and Instruments

The heuristic evaluation instrument contains semi-structured questions bordering on the usability and data representation correctness (see Appendix D).

5.5.2 Findings

The findings from the heuristic evaluation are presented in Table 5.3.

Table 5.3

Findings from Heuristic Evaluation

Questions	**Specialty**	**Comments (summarized and paraphrased)**
Does the InfoVis highlight important information? Any further suggestion?	Data specialist 1	*"Yes, it does [highlight important information]. I would suggest it has detailed explanation on the info for new users"*
Does the InfoVis simplify information representation? Any further suggestion?		*"Yes, the information is well represented. Information with percentages and values should be included"*
		"Yes, it does support data exploration. The data definition is supportive, also."
Does the InfoVis support data exploration that attracts interested analysis?		*"Yes, the tool is interesting for presentation and it gives detail"*
	Data Specialist 2	*"Yes, it does [highlight important information]. I suggest sentences "about the InfoVis", what to do [affordances]"*
Is the information presented by the InfoVis well organized? Any further suggestion?		"*Yes, it [the InfoVis] is simple.it gives 3 dimensions of data at once and allows explorations. I suggest brief explanation on what the explicit knowledge could imply."*

Is the information presented by the InfoVis understandable at a glance?		*"Yes, it is [well organized]. More information would be helpful"*
Can the InfoVis attract frequent users' patronage?	Usability engineer 1	*"It [the visuals} is very good for perceptual-intermediate use"*
Is the InfoVis simple in operation? Any further comments?		*"The InfoVis is simple but needs improvement in its colour grading. Disparate info on line and scatter plot should be checked"*
How will you evaluate the ease of usage of the InfoVis?		*"Need to improve statistical values, may be with pie charts. Add RVMF for information display"*
	Usability engineer 2	*"Yes [it attracts users' usage. More detail info about the usage guide will be better"*
How will you evaluate the functionality of the InfoVis and its learning process?		*"It is simple, good and functional, though with further improvement for real life adoption"*

The heuristic evaluation shows that the HSDI prototype supports data representation correctness, and with satisfactory usability quality. The essentials of the feedback are the suggested improvements. These are (a) a users' guide that explains a "how to" process for users (b) addition of pie-chart for further descriptive details of the visualized data, and (c) improved colour scheming. The suggested improvements are accordingly attended to. These are reported in section 5.6 as the implications of the formative evaluations on this study.

5.6 Implications of the Formative Evaluations on this Study

The formative evaluation presents four improvements to both the proposed design framework and the HSDI prototype. First, pie chart is now included in the proposed visualization techniques to handle the descriptive distribution of selected data dimensions. Second, based on the improved conceptual design framework, the HSDI prototype is also improved. Pie chart visualization page is therefore designed. Figure 5.15 presents the improved conceptual design framework, showing pie chart as newly-added visualization technique.

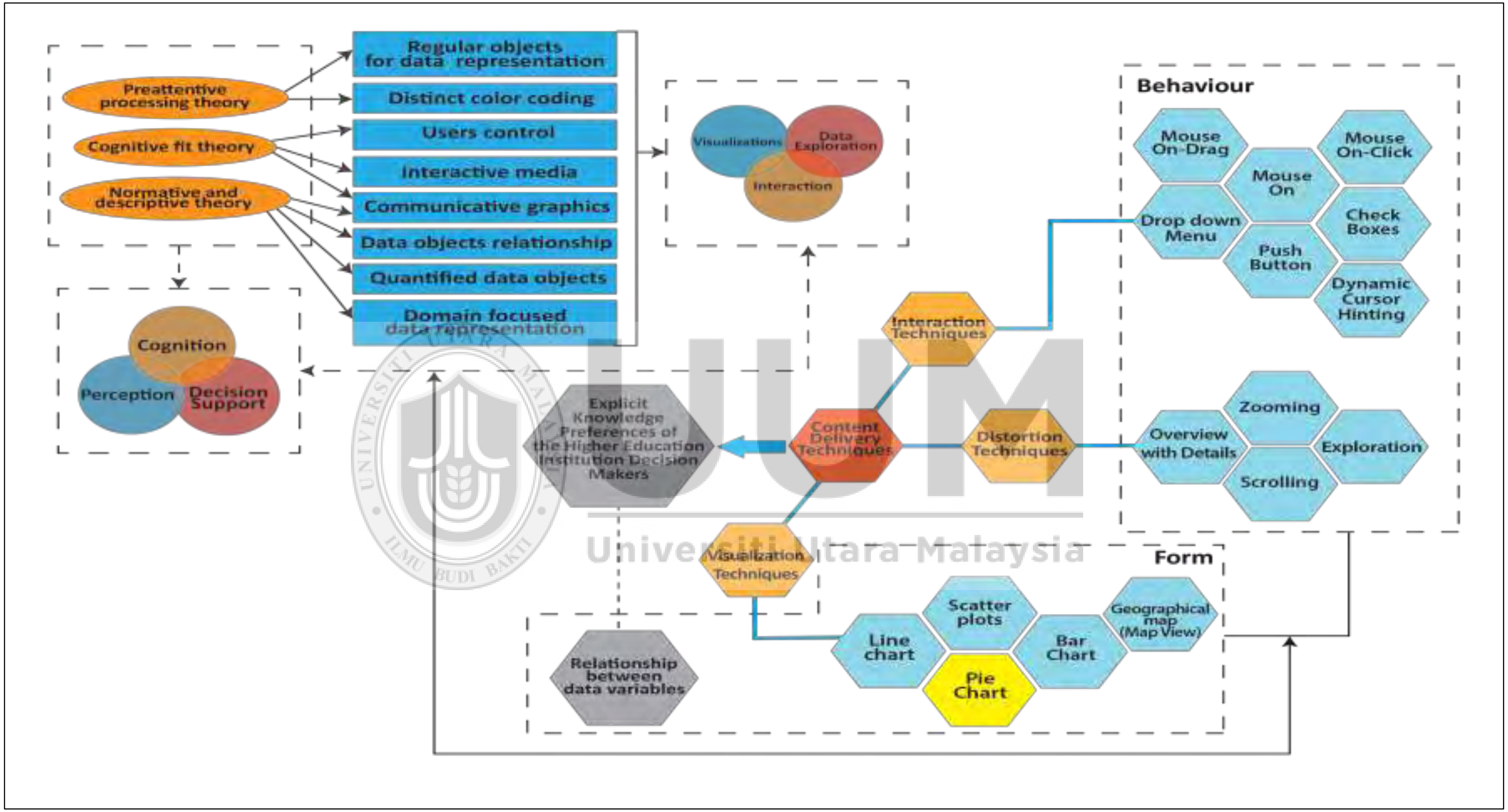


Figure 5.15. Final Conceptual Design Framework

Third, a users' guide page is now included into the HSDI prototype to explain its working, guide users and enhance their affordances. Fourth, a new colour scheming is introduced for the categorical data dimensions to enhance users' visual perception and experience. Figures 5.16 – 5.19 are screenshots depicting the improvements made to the HSDI prototype.

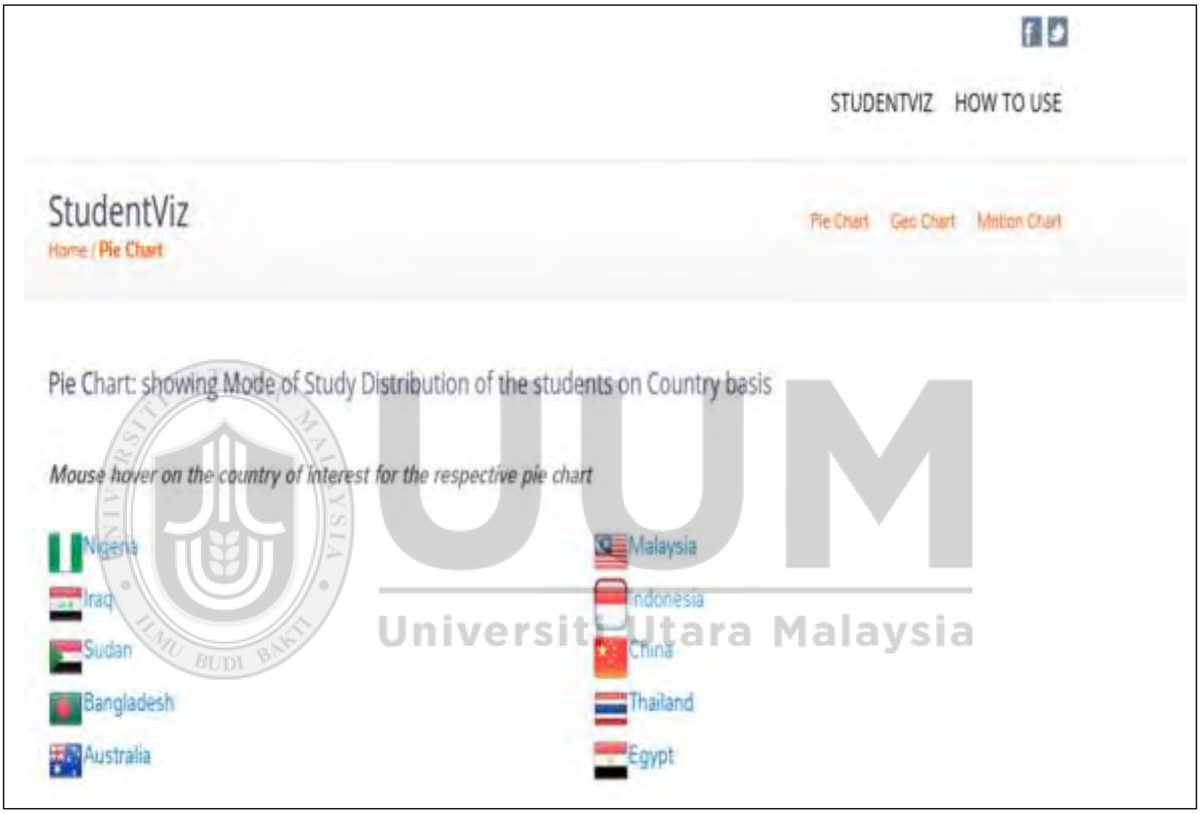


Figure 5.16. Pie Chart Visualization page (Before the Dynamic Hinting i.e. Mouse Hover)

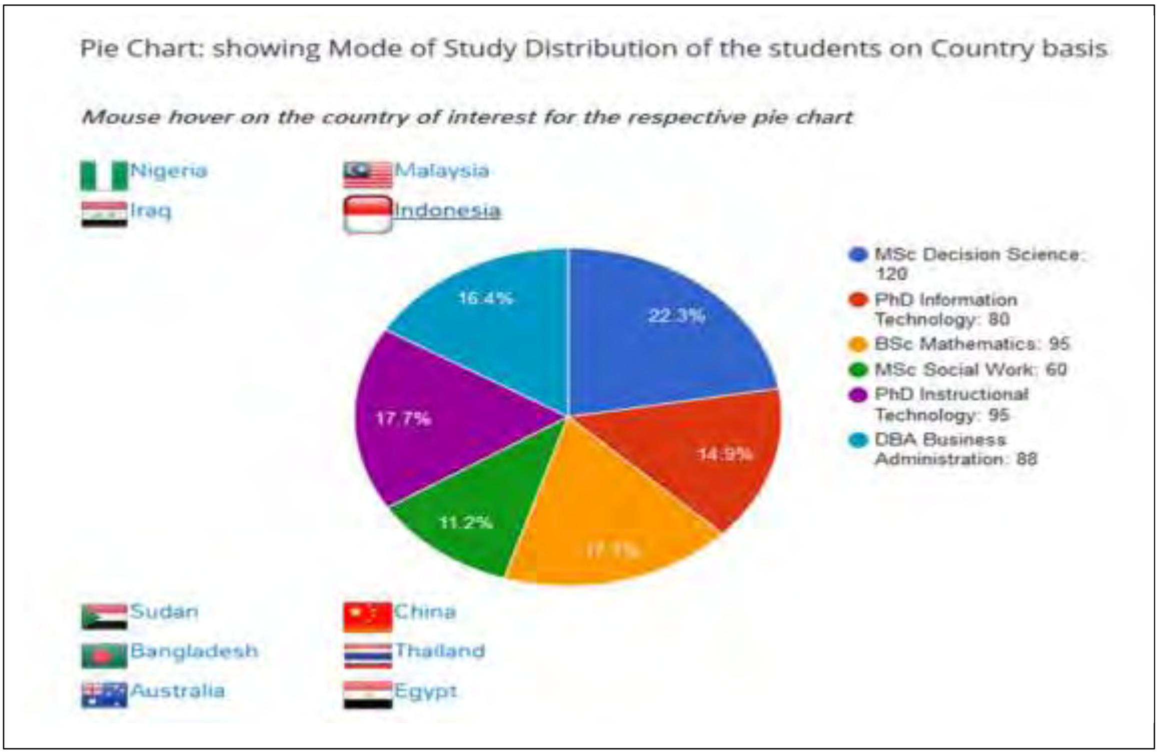


Figure 5.17. Pie Chart Visualization page (When it is mouse-hovered)

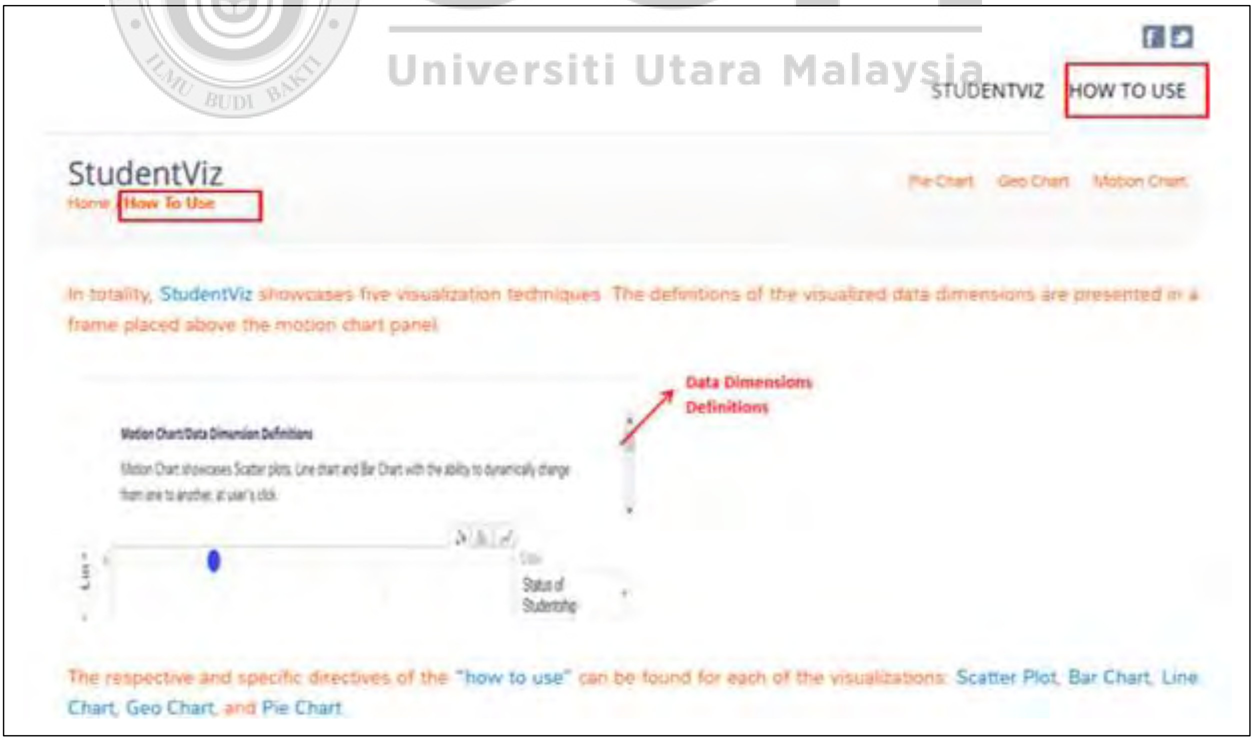


Figure 5.18. "How to page" (Detailed Information about the HSDI and its usage)

Figure 5.19. Improved Colour Scheming "How to page" (Detailed Information about the HSDI and its usage)

The HSDI now supports choosing 'unique colours" to activate rendering of the visualization objects with clearly different colours, or "same colour" for same colour rendition. In scatter plot, the visualization can support four dimensions: Two ordinal dimensions in y and x axes, colour scheme for categorical dimension, and the size of the scatter plot can also represent another ordinal data dimension. After this improvement and necessary post-development activities (described in section 4.5.3 in Chapter 4), the domain users' experience evaluation study is conducted.

5.7 Summary of the Chapter

This chapter describes the expert review, prototyping and heuristic evaluation processes of the proposed HSDI conceptual design framework. The heuristic evaluation is the formative evaluation of the prototype. The expert review was conducted by experts with expertise in software engineering, data visualization and InfoVis. The mean value of their review of the conceptual design framework, and the respective means of the dimensions, suggest a satisfactory, functional and usable framework. The prototyping stage was the process of turning the proposed conceptual design framework to a physical artifact, i.e. HSDI prototype. This achieved a web-based InfoVis system which is developed to handle the multidimensionality of the HEI students' dataset. The successful implementation of the proposed framework in achieving the prototype validates its applicability. The heuristics evaluation is on the usability and data representation correctness of the developed prototype, and done by usability and data analysis/administrator experts. The findings from the heuristics evaluation also suggested that the prototype is usable and it correctly represents the data and present previously unknown information. The suggested improvements are subsequently made on both the conceptual design framework and the prototype. At the overall, the formative evaluation findings are positive and support the viability of the proposed conceptual design framework. In the next chapter, the reports of users' experience evaluation study are presented with its findings.

CHAPTER SIX

USERS' EXPERIENCE EVALUATION STUDY

6.1 Introduction

This chapter contains the findings of the users' evaluation study. The proposed conceptual design framework is discussed in chapter 4. This framework is verified through expert review and validated through prototyping and heuristics evaluation, as reported in chapter 5. The users' evaluation study is the summative evaluation process to assess the users' experience of the HSDI prototype in terms of its correct data representation, usability, and decision support. Figure 6.1 presents the users' evaluation study process.

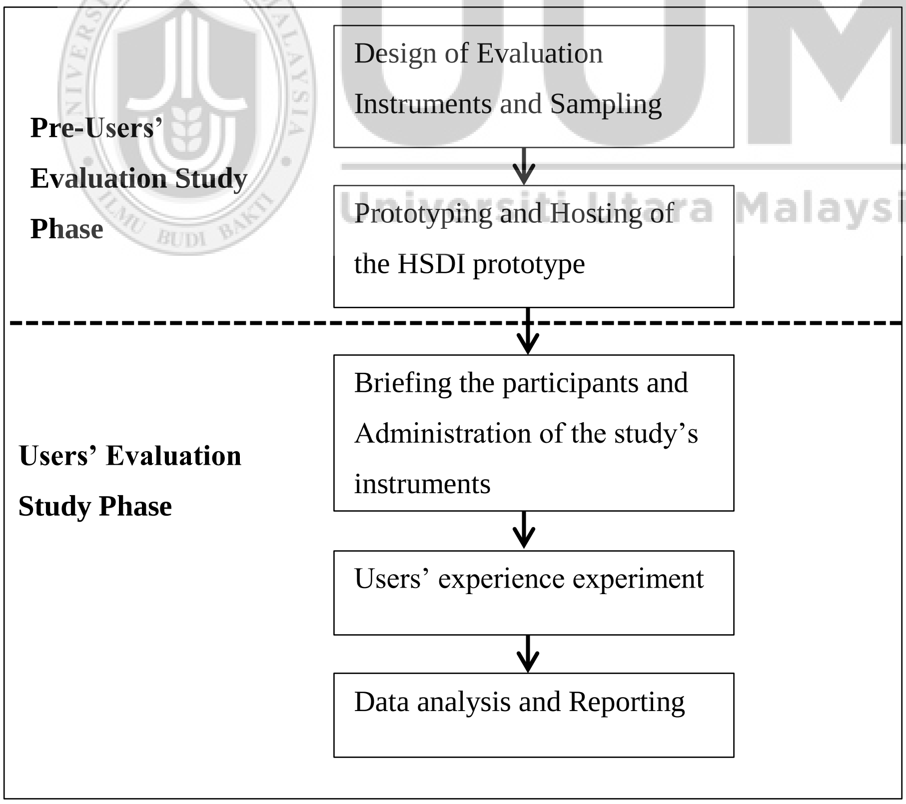


Figure 6.1. Users' Evaluation Study Process

The design of the users' experience evaluation instrument is earlier discussed in chapter 3 of this thesis (section 3.6.4 and 3.7.4), and the prototyping stage is reported in chapter 5 (section 5.3). The evaluation phase which contains briefing of the participants, administration of the instruments, data analysis and findings is discussed in the following sections.

6.2 Briefing of the Participants and Administration of Instruments

The users' evaluation study was electronically conducted. HSDI prototype, a web-based InfoVis, was hosted on 8th October, 2015. The prototype's URL was sent to invited participants. The characteristics of HEI policy makers and the justifications for the sampling method are earlier discussed in Chapter 3 (section 3.6.4).

A mail requesting for the attention of the selected HEI policy makers in the users' evaluation study was sent on 10th October, 2015. Out of the 43 HEI policy makers selected, fourteen (14) partook in the users' evaluation study. As earlier mentioned, users' evaluation study has no definite participants' size. The circumstance of the study and its objective are main determinants (Tullis & Albert, 2008). The participants are briefed through phone and *Skype* calls. An electronic survey instrument is used to collect their feedbacks. The briefing and data collection took place within 15th October to 1st December, 2015. The data collected is thereby analysed and the findings reported.

6.3 Data Analysis

The data analysis process involves data coding, qualitative data analysis, quantitative data analysis and triangulation. The data coding is labelling of the survey items with alphanumeric values to enable error-free data analysis. The survey instrument contains both semi-structured (qualitative) and closed-ended (quantitative) questions and items for collection of objective and subjective evaluation feedback respectively. Finally, the survey items assessing the HSDI prototype's decision making support, from both the qualitative and quantitative data, are triangulated. This provides further reliability and validity to the findings (Creswell, 2009).

6.3.1 Data Coding

The data coding for the closed-ended (quantitative) items are presented in Tables 6.1.

Table 6.1

Data Coding for the Quantitative Data

Code	**Items**	**Dimension**
CDR01	It highlights important information	Correct Data Representation
CDR02	It simplifies the data representation	
CDR03	The data shown after exploration attracts my interest	
CDR04	The presented information is well-organised	
CDR05	I can understand the information represented at a glance	
USB01	I would like to use this system frequently	Usability
USB02	I found the system simple	
USB03	The system was easy to use	
USB04	I can use the system without assistance from a technical person	
USB05	I found the various functions in the systems well integrated	
USB06	The system operation was consistent	
USB07	I imagine that most people would learn to use this system quickly	
USB08	I found the system simple to use	
USB09	I felt confident using the system	
USB10	The system has the functions and capacities I expect	

	it to have	
DMS01	The screen layout is visually appealing	Decision making support (Perceptual)
DMS02	The colouration of the data objects elements is pleasant	
DMS03	The visual representation supports the data insight	
DMS04	Exploration functions (filtering and extracting information) are consistent with needs	Decision making support (Cognitive)
DMS05	The system presents the hidden information	
DMS06	The data exploration process reveals insightful information	
DMS07	The system well-structures presented the information needed	Decision making support (Decision making)
DMS08	The information presented meets my goal of using the system	
DMS09	The system supports decision making	

The quantitative items cover all dimensions proposed to evaluate the HSDI prototype. Correct data representation is assessed through items CDR01 – CDR04, usability is assessed through items USD01 –USB 10 and decision support is assessed through items DMS01 – DMS09. The items for the decision making support is further divided into three (DMS01 – DMS03, DMS04 –DMS06, DMS07 –DMS09) to subjectively assess the perceptual, cognitive and decision support of HSDI prototype. This is theoretically supported and discussed in chapter 2 (section 2.4.4) and chapter 3 (section 3.7.4). The data coding for the semi-structured (qualitative) questions is presented in Table 6.2.

Table 6.2

Data Coding for the Qualitative Data

Code	Items	Dimension
QL-DMS01	Can you identify particular information that you knew before the visualization?	Decision making support (Perceptual)
QL-DMS02	Can you identify particular information that you did not know before looking at the visualization?	
QL-DMS03	Can you distinguish among different objects in the dataset?	
QL-DMS04	Can you describe the categories of the datasets?	
QL-DMS05	Do you think this representation helped you in classifying the data?	Decision making support (Cognitive)
QL-DMS06	Do you think that this representation gives you a clear picture of the distribution of the objects and their values?	
QL-DMS07	Do you feel that this representation provides sound comparison among similar objects?	
QL-DMS08	Do you feel that this representation provides sound comparison among different objects?	
QL-DMS09	Do you feel that this representation provides clear picture of the relations between objects?	
QL-DMS10	Do you feel that this representation underscores the most important relationship?	Decision making support (Decision making)
QL-DMS11	Can you identify the relationship between objects represented?	
QL-DMS12	Does this representation give you the opportunity to better establish relationship between objects?	
QL-DMS13	Is there any correlation between values of the objects viewed?	
QL-DMS14	Do you think this representation gives you the opportunity of identifying new relationships?	

The qualitative data items only assess the decision making support of the HSDI prototype. The 14-item semi-structured questions are coded QL-DMS01 to QL-DMS14. The items are also sub-divided into three: Perceptual (QL-DMS01 to QL-DMS04), cognitive (QL-DMS05 to QL-DMS09), and decision making support (QL-DMS10 to QL-DMS14).

6.3.2 Quantitative Data Analysis

The quantitative data analysis is done using SPSS version 19, with Mann-Whitney U Test as the non-parametric technique. The technique is suitable for measuring continuous variables, and also to investigate its differences between two independent groups. The continuous variables assessed in this study are correct data representation, usability and decision making support. The two independent groups are sub-divisions of the participants based on their years of experience. The researcher classified the years of experience into two, i.e. medium experience (3 – 16 years) and high experience (17 – 28 years). The classification range is based on the median value of the participants' years of experience (Maximum and minimum years of experience are 28 and 3 years, respectively). The classification into two groups is to allow the data fits into Mann-Whitney U Test. Descriptive analysis is also to explore the frequencies of the participants' assessment of the decision making support sub-variables –perceptual, cognitive and decision. The following hypotheses were therefore tested during the quantitative data analysisThe findings are discussed in section 6.4.

6.3.3 Qualitative Data Analysis

The qualitative data, i.e. responses from items presented in Table 6.2, are analysed using Elo and Kyngas' (2008) content analysis approach. This content analysis approach is presented in Figure 6.2.

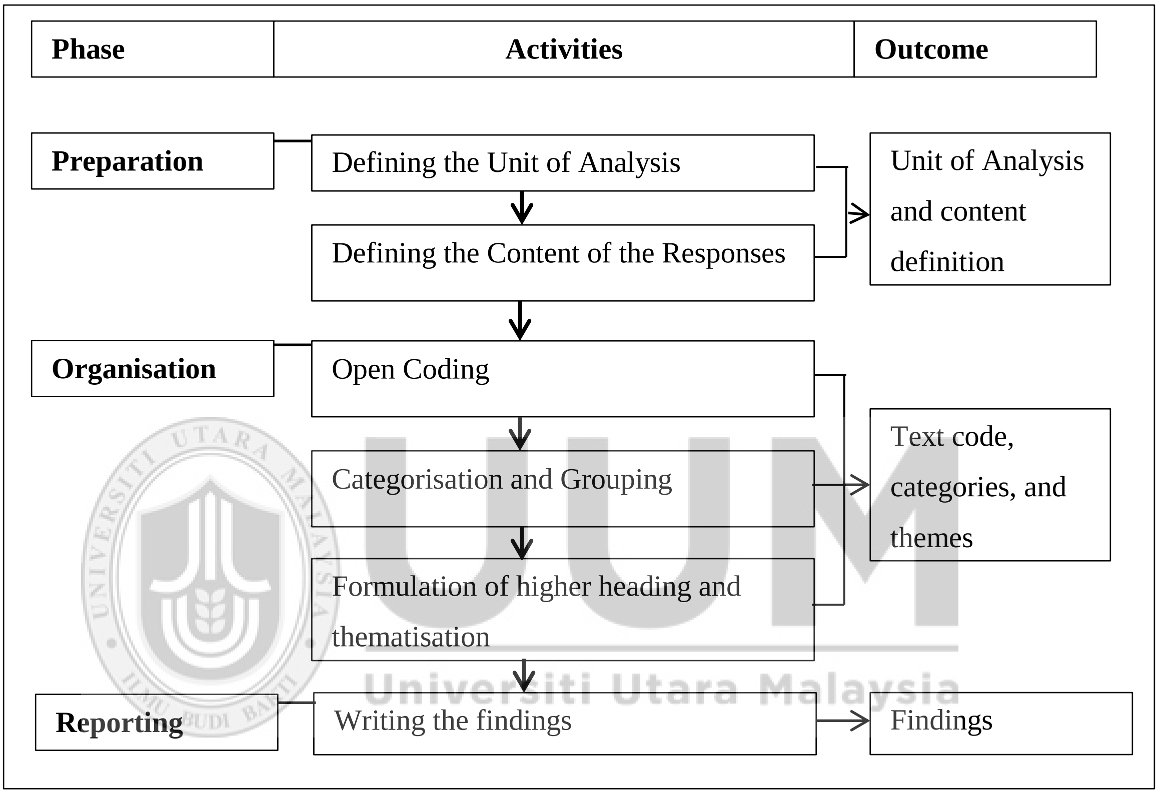


Figure 6.2. Content Analysis Approach

(Source: Elo & Kyngas, 2008)

Further details of the preparation and organisation phases are discussed in sections 6.3.3.1 and 6.3.3.2 respectively. The findings are discussed in conjunction with the quantitative data analysis findings in section 6.4.

6.3.3.1 Preparation

The unit of analysis of the data collected and its content definition are determined at this stage. The unit of analysis is the smallest indivisible entity or whole which forms the basis of the analysis (Chenail, 2012). This study takes response to item (the semi-structured question) as its unit of analysis. It allows close inference to be made since each of these items is direct assessments of the HSDI prototype which is the objective of the users' experience evaluation study. The content of the response is developed into positive and negative categories, accordingly. This is thus defined as manifest content (Elo & Kyngas, 2008).

6.3.3.2 Organisation

Open coding for the qualitative data collected is the first stage in the organisation phase. This is presented in Table 6.2. It allows error-free data documentation and analysis. The data sheet for the qualitative data analysis on the designed response-to-item basis is presented in Appendix I.

From the analysis, as presented in Appendix I, 'Yes' and 'No' responses are the categories, and headings are grouped into positive and negative statements represented in percentages (%). Information that is not covered under these generic terms is also reported as complementary.

6.3.4 Triangulation

Triangulation is the process of examining evidences from different data sources to build coherent justification for themes (Creswell, 2009). In this study, it is used for the assessment of participants' evaluation of the decision making support of the HSDI prototype. It is the only evaluation dimension that has its data sources from both the qualitative and quantitative data. The triangulation is also reported under the findings in section 6.4.

6.4 Findings

The findings from the data analysis are discussed under four headings, namely; demography, correct data representation, usability and decision making support. The useful demographic data, as regards this study, is the years of experience of the participants. The participants' assessments of the HSDI prototype, in terms of its correct data representation, usability and decision making support, are therefore reported.

6.4.1 Demography: Participants' Years of Experience in HEI Administrative Role

The years of the experience, as a demographic data of the participants, are classified into medium experience (3 – 16 years) and high experience (17 – 28 years). Table 6.3 presents the demography data of the distribution.

Table 6.3

Years of Experience Distribution of the Participants

	Frequency	Percentage
Medium Experience	9	64.3
High Experience	5	35.7
Total	14	100.0

The least value of year of experience of the participants is 3, and this is 7% of the participants' population. The rest have their years of experience ranging from 5 years to 28 years. The years of experience distribution shows that the participants are well-experienced in handling students' data related decision making in HEIs. The classification into two independent groups (medium and high experience) is done to meet the condition of Mann-Whitney U Test, and investigate the effect of years of experience in the participants' assessments of the HSDI prototype.

In the three hypotheses tested, the predictor variable (can also be called independent variable) is HEI decision makers' years of experience. The response variables (can also be called dependent variables) for hypothesis 1 – 3 respectively are correct data representation, usability, and decision making. The hypotheses tested are:

H_1: HEI decision makers with high experience find HSDI to correctly represent data than those with medium experience

H_2: HEI decision makers with high experience find HSDI to be usable than those with medium experience

H_3: HEI decision makers with high experience find HSDI to support decision making than those with medium experience

6.4.2 Correct Data Representation

The results of the participants' assessments of the HSDI prototype in terms of its correct data representation are presented in Tables 6.4 to 6.6. Table 6.4 presents the mean ranks for HSDI's correct data representation. Table 6.5 presents the Mann-Whitney U statistics test, and Table 6.6 presents median values of the groupings for H_1 (*HEI decision makers with high experience find HSDI to correctly represent data than those with medium experience*).

Table 6.4

Ranks for Correct Data Representation

	Years of Experience	N	Mean Rank	Sum of Ranks	Max	Mean
Correct Data Representation	Medium Experience	9	9.22	83.00	32	28.86
	High Experience	5	4.40	22.00		
	Total	14				

Table 6.5

Test Statistics for Correct Data Representation

Item	Correct Data Representation
Mann-Whitney U	7.000
Wilcoxon W	22.000
Z	-2.210
Asymp. Sig. (2-tailed)	.027
Exact Sig. [2*(1-tailed Sig.)]	.042

Table 6.6

Median Values for Groups' Assessment of the Correct Data Representation

Years of Experience	Correct Data Representation
Medium Experience	29.00
High Experience	27.00
Total	29.00

A mean value of 28.86, when the maximum is 32, shows that the HSDI prototype possesses correct data representation, as assessed by the participants. With a Z value of -2.2 (approximately) and a significance level (p) of 0.027, there is statistically significant difference in the assessment of the HSDI prototype's correct data representation between the medium and high experience participants (because p value < 0.05). The median values of 29 and 27 for medium and high experience participants respectively show that participants with medium experience found HSDI prototype to support correct data representation than participants with high experience. The Effect size (r)is calculated with $z/\sqrt{N}$, where N is the number of cases, i.e. 14. Therefore, r = 0.5. This is large effect size according to Cohen (1988) criteria. This means that the significance size is large.

The result can be presented as: Mean value of 28.86 revealed HSDI prototype supports correct data representation. A Mann-Whitney U test revealed significant difference in the assessments of HSDI prototype in terms of its correct data representation between the medium ($Md = 29,\ n = 9$) and high experience ($Md = 27,\ n = 5$) participants, $U = 7,\ z = -2.210, p = 0.027, r = 0.5$.

The first hypothesis; *HEI decision makers with high experience find HSDI to correctly represent data than those with medium experience,* is rejected, but in retrospect supports the opposite. The results rather showed that HEI decision makers with medium experience find HSDI to correctly represent data than those with high experience.

6.4.3 Usability

The results of the participants' assessments of the HSDI prototype in terms of its usability are presented in Tables 6.7 to 6.9. Table 6.7 presents the mean ranks for HSDI's usability, Table 6.8 presents the Mann-Whitney U statistics test, and Table

6.9 presents median values of the groupings for H_2 (*HEI decision makers with high experience find HSDI to be usable than those with medium experience*).

Table 6.7

Ranks for Usability

	Years of Experience	N	Mean Rank	Sum of Ranks	Max	Mean
Usability	Medium Experience	9	7.28	65.50	60	51.57
	High Experience	5	7.90	39.50		
	Total	14				

Table 6.8

Test Statistics for Usability

Item	Usability
Mann-Whitney U	20.500
Wilcoxon W	65.500
Z	-.268
Asymp. Sig. (2-tailed)	.788
Exact Sig. [2*(1-tailed Sig.)]	.797

Table 6.9

Median Values for Groups' Assessment of the Usability

Years of Experience	Usability
Medium Experience	53.00
High Experience	51.00
Total	52.00

A mean value of 51.57, when the maximum is 60, shows that the HSDI prototype's usability is satisfactory, as assessed by the participants. With a Z value of –0.27 (approximately) and a significance level (p) of 0.788, there is no statistically significant difference in the assessment of the HSDI prototype's usability between the medium and high experience participants (because p value > 0.05). The median values for the medium and high experience participants are 53 and 51 respectively.

The Effect size (r) (of the insignificant difference) is calculated with $z/\sqrt{N}$, where N is the number of cases, i.e. 14. Therefore, r = 0.1 (approximately). This is small effect size according to Cohen (1988) criteria. This means that the insignificance size is small.

The result can be presented as: Mean value of 51.57 revealed HSDI prototype supports usability. A Mann-Whitney U test revealed no significant difference in the assessments of HSDI prototype in terms of its usability between the medium ($Md = 53,\ n = 9$) and high experience ($Md = 51,\ n = 5$) participants, $U = 20.5$, $z = -0.27, p = 0.788, r = 0.1$.

The second hypothesis; *HEI decision makers with high experience find HSDI to be usable than those with medium experience,* is rejected. It shows that the difference in experience of the HEI decision makers has no significant effect in the usability evaluation of HSDI.

6.4.4 Decision Making Support

The results of the participants' assessments of the HSDI prototype in terms of its decision making support are presented in Tables 6.10 to 6.12. Table 6.10 presents the mean ranks for HSDI's decision making support, Table 6.11 presents the Mann-Whitney U statistics test, and Table 6.12 presents median values of the groupings for H_3 (*HEI decision makers with high experience find HSDI to support decision making than those with medium experience*).

Table 6.10

Ranks for Decision Making Support

	Years of Experience	N	Mean Rank	Sum of Ranks	Max	Mean
Decision making Support	Medium Experience	9	6.17	55.50	61	53.86
	High Experience	5	9.90	49.50		
	Total	14				

Table 6.11

Test Statistics for Decision Making Support

Item	Decision making Support
Mann-Whitney U	10.500
Wilcoxon W	55.500
Z	-1.605
Asymp. Sig. (2-tailed)	.108
Exact Sig. [2*(1-tailed Sig.)]	.112

Table 6.12

Median Values for Groups' Assessment of the Decision Making Support

Years of Experience	Decision Making Support
Medium Experience	51.00
High Experience	56.00
Total	54.00

A mean value of 53.86, when the maximum is 61, shows that the HSDI prototype supports decision making, as assessed by the participants. With a Z value of -1.6 (approximately) and a significance level (p) of 0.108, there is no statistically significant difference in the assessment of the HSDI prototype's decision making support between the medium and high experience participants (because p value > 0.05). The median values for medium and high experience participants are 51 and 56 respectively. The Effect size (r) is calculated with $z/\sqrt{N}$, where N is the number of cases, i.e. 14. Therefore, r = 0.4. This is medium effect size according to Cohen (1988) criteria. This means that the insignificance size is neither small nor large.

The result can be presented as: Mean value of 53.86 revealed HSDI prototype supports decision making. A Mann-Whitney U test revealed no significant difference in the assessments of HSDI prototype in terms of decision making support between the medium $(Md = 51,\ n = 9)$ and high experience$(Md = 56,\ n = 5)$ participants, $U = 10.5,\ z = -1.605,\ p = 0.108,\ r = 0.4$.

The third hypothesis; *HEI decision makers with high experience find HSDI to support decision making than those with medium experience*, is rejected. It shows that the difference in experience of the HEI decision makers has no significant effect in the decision support evaluation of HSDI.

On the other end, the responses to the semi-structured questions assessing the HSDI prototype's decision making support are analysed using Elo and Kyngas' (2008), as discussed in section 6.3.3 (see Appendix I). The 'Yes' and 'No' are built from the responses to connote positive and negative responses and calculated in percentages (%). Table 6.13 presents the summary of the analysis of the responses.

Table 6.13

Qualitative Analysis of the Responses assessing HSDI Decision Making Support

		Yes		No	
		Frequency	%	Frequency	%
Decision	Perceptual	41	74	14	26
Making	Cognitive	55	98	1	2
Support	Decision	63	93	5	7
	Total	159	**88**	20	**12**

The summary of the analysis presented in Table 6.13 showed that 88% of the respondents are positive about the HSDI prototype ability to support decision making. This aligns with the findings from the closed-ended items (presented in Tables 6.10 to 6.12), and thereby strengthens the conclusion. The implication of this finding is that users can identify previously unknown information when using the InfoVis tool for students' data exploration. It also means the HEI policy makers attested to the ability of the InfoVis tool to present distinctively-categorised information, ability to locate interested section of the students' data (filtering out the uninterested portion), to describe presented information with suitable statistical graphics, and to give an insightful comparison of data represented. It is the collectivist effect of the data exploratory ability, easily-recognizable data presentation, and insightful information that aid the domain users' decision making process. These are the respective human-InfoVis activities that aid perception, cognition and decision support.

Analysing the findings further, the perceptual sub-dimension of the decision making support attracted the highest negative percentage (26%). The major contributor to this is the item that inquires if participants "*can [you] identify particular information that [you] knew before the visualization?* Obviously, responses are in negative like "*No. The dataset visualized is new*" from participant 01, and "*It is a data set of no prior knowledge*" from participant 04, among others. The higher negative ratings do not discredit the perceptual quality of the prototype, but rather highlight the unfamiliarity of the data being explored by the participants.

Also, there are other important statements made by the participants that capture their experiences while interacting with the HSDI prototype. The participants relatively found visualization techniques that demand less users' control mechanism more fulfilling. Participant 04 writes *"Yes I think so [I can interpret the information readily]. Especially the pie chart and the geo chart."* Also, the participants emphasise prior experience of both the domain and technology as factors for pleasant users' experience. Participant 14 writes *"Prior experience is needed to comprehend the relationship [of the data objects]"*, and from participant 01, *"... But fairly, the user can use his discretion and prior experience to highlight the relationship."*

Lastly, even though assessment of the HSDI prototype is satisfactory and each of the dimensions is well-evaluated, participants still hope that improved design can still be achieved. Design improvement is infinite and it is actually continuous in product design (Cooper et al., 2014).

6.5 Summary of the Chapter

This study evaluated the HSDI prototype based on the three dimensions, namely; correct data representation, usability and decision making support. From the findings of the users' evaluation, the HSDI prototype supports correct data representation, usability and decision making. Findings from both the subjective (quantitative) and objective (qualitative) feedback show that the HSDI prototype is satisfactory, and by extension, the conceptual design framework is practically usable in designing InfoVis that supports the stated dimensions. This implies that the proposed framework presents the suitable content delivery techniques for handling the multidimensionality of the HEI students' data and presenting the domain explicit knowledge preferences. It shows that the implementation of the specified visualization, interaction, and distortion techniques, addresses the cognitive load which is the effect of information overload caused by the multidimensionality of the HEI students' data.

The study also found that the participants' assessment of the HSDI prototype's correct data representation significantly differs based on their year of experience. The participants with medium experience rated the prototype higher than those of high experience. This shows that a factor which is traceable to participants' years of experience is responsible for such findings. The researcher suggests that access to technology and/or knowledge of technology usage, especially data visualization tool could be responsible. However there is no statistically significant difference found in the assessment of the HSDI prototype in terms of usability and decision making support, based on their year of experience. In sum, the findings revealed that the InfoVis conceptual design framework is suitable for HEI students' data-focused InfoVis. Also, the HSDI prototype is found to satisfactorily present and represent HEI students' data, usable, and aid the HEI policy makers' decision making process. Chapter seven, the next and last chapter, concludes this thesis.

CHAPTER SEVEN
DISCUSSION, LIMITATION AND FUTURE WORK

7.1 Introduction

This study developed a domain specific HEI students' data-centred InfoVis conceptual design framework which addresses the problem of information overload inherent in HEI students' data because of its multidimensionality. The problem is addressed by proferring suitable specifications of the content delivery techniques and process models. The conceptual design framework is new as its addresses the specificity of HEI students' data and domain decision makers. It falls within the InfoVis user-centred design studies, like Christina, Julian, and Micheal (2016), Krstajic (2012) and Boo et al. (2015) on visualization of news story analytics, and Meyer (2012), Koh et al. (2010) and Roth et al. (2005) on biological and medical/health data, among others.

In constrast with previous related studies on visualization models and frameworks, this study presents extended and novel ideas, where applicable, on four different fronts. First, it extended the components of models on expressive process of visualization and visual display with a linkage to explicit knowledge preferences of the domain. Similar previous studies in this regard are Robertson and DeFerrari (1999) and Owen (1999). Both studies described the abstract description of visualization process, its components based on paradigms and data to represent. This study however emphasizes the value of understanding the explicit knowledge to be represented and presented, beyond the data. It establishes that the InfoVis' ability to amplify cognition and generate hypotheses is inherent in its functional advantage of meeting the domain users' explicit knowledge preferences.

Second, it further establishes that the domain users' explicit knowledge preferences is also a determinant in the choice of visualization, interaction and distortion techniques. This extends the earlier works of Keim (2002) and Chi (2000) which stated the orthogonal connection in the choice of InfoVis techniques and the domain data, respectively. This study posits that, aside the domain data which is usually in its raw form, as Chi (2000) observed, that the explicit knowledge (i.e. the transformed data into actionable information) must be considered in the choice of the visualization, interaction and distortion techniques. Third, this study presents a novel theory-based design model in InfoVis design. The theorectical model presents a linkage between perpetual, cognitive, and decision science theories with InfoVis techniques and their respective design features. It further adds to the much-anticipated studies on InfoVis theoretical foundations advocated and worked on by Yi (2010), Liu et al. (2007), Teets et al. (2010), Li et al. (2011), Benoit (2012), Graham (2005), and Benoit (2012). This study presents a physical abstraction of design features, based on theories, while past theory-related InfoVis studies are on evaluation. Lastly, the prototyping stage in previous user-centred design models of Meyer (2012), Koh et al. (2011), Robinson et al. (2005), and Roth et al. (2010) are disintegrated into phases and stages in the process model proposed by this study. It presents a practical guide to translating InfoVis conceptual design framework into physical artefacts.

This study presents a HEI domain-specific InfoVis content delivery techniques. This is further explained in sections 7.2 – 7.5. It has answered its earlier stated three research questions: a) What are the explicit knowledge preferences (in relationship with students' data) of the HEI policy makers?, b) What are the visualization, interaction and distortion techniques that can represent and present HEI students' dataset with due consideration of the domain policy makers' explicit knowledge preferences?, and c) Would the proposed visualization, interaction and distortion techniques and the consequent InfoVis tool assist in supporting decision making based on the explicit knowledge identified? This chapter discusses the answers proffered to these research questions, and revisits the respective research objectives. Also, it summarizes the research contributions, concludes with the discussion of its limitations, and suggests the future research works.

7.2 Research Question 1: What are the explicit knowledge preferences (in relationship with students' data) of the HEI policy makers?

The preliminary domain study (discussed in section 1.4) revealed that HEI policy makers experience information overload while interacting with students' data (Semiu & Zulikha, 2014). This information overload experience which is defined as getting lost in data which may be irrelevant to the task at hand, processed in an inappropriate way or presented in an inappropriate way (Keim, 2008; Ning et al., 2012) is caused by the inappropriate data visualization tool. Clearly, there have been no domain-specific InfoVis for handling HEI students' data, thus, the absence of the domain InfoVis conceptual design framework. To achieve this, as usual in all previous InfoVis user-centred studies (e.g. Robinson, Chen, Lengerich, Meyer, & MacEachren, 2005; Roth, Ross, Finch, Luo & MacEachren, 2010; Sedlmair, Meyer, & Munzner, 2012), identifying the HEI domain's explicit knowledge preferences (in relationship with students' data) was the first step.

This study investigated the HEI domain in a pre-design study (discussed in section 4.2). Through a content analysis of past literatures on HEI data-driven decision making processes (see Table 4.2), it elicited the related explicit knowledge preferences (see Table 4.3). The explicit knowledge preferences are then validated by purposively selected HEI policy makers. The responses from the HEI policy makers are used to prioritize the domain explicit knowledge preferences (see Table 4.4). In summary, the following (a – e) are the validated students' data-centred HEI policy makers' explicit knowledge preferences (Semiu & Zulikha, 2015c).

a. The patterns of various groups of students based on their race, religion, sex and nationalities; colleges; previous graduates based on their race, religion, sex and nationalities; course of study, CGPA and nationalities of the Alumni,

b. The relationship between the students' performance and their gender, race and nationalities; mode of entry; disabilities; health weakness; majoring courses; number of semester used; programme of study; mode of study; English language proficiency; soft skill acquisition; and deferment history,

c. The relationship between drop-out students and their economic status; gender, nationality and marital status,

d. The relationship between students' employability and internship record; majoring courses; course of study, and

e. The students' enrolment pattern based on nationalities, gender and sex; mode of admission.

Based on the validated explicit knowledge preferences, the comprehensive data dimensions for a HEI student data instance based on the related data-driven decision making process are categorized. The categories are into basic and derived as data class, and numeric, real and categorical as data types (see Table 4.5).

7.3 Research Question 2: What are the visualization, interaction and distortion techniques that can represent and present HEI students' dataset with due consideration of the domain policy makers' explicit knowledge preferences?

After the identification of the HEI domain policy makers' explicit knowledge preferences, and the final elicitation of the students' data dimensions, identifying the visualization, interaction and distortion techniques was the next step. This was to ensure that presented and represented HEI students' dataset essentially support the domain decision making process with the aid of an InfoVis.

This study attended to this question in two fronts. These are structural model (see Figure 4.1) and content delivery techniques (see Figures 4.2, 4.6, and finally 5.14). Fourteen (14) InfoVis tools that can be lively accessed, or with available documentations, were analysed (see Appendix G) to create a generic template of InfoVis structure. Then, with due attention to HEI specifics, as found in the pre-design study, the structural components of HEI InfoVis were proposed (see Table 4.6). It is found that InfoVis is generically into three tiers. These are visual interactive interface, data objects binding system and the database. The components of the visual interactive interface are data layer panel, visualization span, interaction mechanism, and distortion tools. Colour scheming and data object modelling are components of the data objects binding system; and multidimensional data table is the HEI specifics of the database (Semiu & Zulikha, 2015d). The details of the HEI students' data-centred InfoVis tiers are further described by the content delivery techniques model (see Figure 4.2).

The proposed content delivery techniques are into three categories: Visualization techniques, Interaction techniques and Distortion techniques. To identify these component techniques, this study analyses the appropriateness of all available techniques as shown in Spence's *"Information Visualization Design for Interaction"* and Cooper's et al.'s *"About Face: The Essentials of Interaction Design"* (Spence, 2007; Cooper et al., 2014), with due consideration to the findings of the pre-design study (see Table 4.8 and Appendix L).

This study finally found that mouse-on, mouse-on-click, drop down menu, push button, check boxes, and dynamic cursor hinting, and mouse-on-drag are the appropriate interaction techniques. The distortion techniques are zooming, overview with details, scrolling, and exploration. And, line chart, scatter plot, map view and bar chart are the visualization techniques. The heuristic evaluation process later suggested pie chart to be part of the visualization techniques.

Furthermore, a theoretical justification and guide was provided for the proposed content delivery techniques. It presents the linkage between InfoVis' theories (preattentive processing theory, cognitive-fit theory, and normative and descriptive theories); the design features (regular objects for data representation, distinct colour coding, users' control, interactive media, communicative graphics, relationship between data objects, quantification of data objects, and domain-focused data represent representation); and the supports (perceptual, Cognition and Decision support). The respective human-InfoVis activities (visualization, interaction, and data exploration) evolve to the InfoVis' techniques components (visualization, interaction, and distortion) (see Figure 4.5). This finally birthed the HEI students' data-focused InfoVis Conceptual Design Framework (see Figure 5.14) which is one of the contributions of this study.

7.4 Research Question 3: Would the proposed visualization, interaction and distortion techniques and consequent InfoVis tool assist decision making based on the explicit knowledge identified?

The proposed visualization, interaction and distortion techniques were evaluated through expert review (see section 5.2), prototyping (see section 5.3), and heuristics evaluation (see section 5.5). The consequent InfoVis tool, *StudentViz*, was evaluated through users' experience study (see Chapter 6).

The expert review was done to verify the appropriateness the conceptual design framework, the proposed visualization, interaction and distortion techniques, and the applicability of InfoVis design process model. The dimensions and their respective measurements for the evaluation of the framework are outlined (see sections 4.5, 5.2). A cumulative mean value of 8.80 out of 10, as the experts' ranking, showed that the proposed design framework is satisfactory. The ranking of each of the dimensions (from the lowest, manageability with 8.20, to the highest, evolutionary with 8.98) is worthy. It implies that the conceptual design framework, for InfoVis design project and InfoVis content creation for HEI students' data, is evolutionary, manageable, simple, compatible, and visible, among others. It also means that the specified visualization, interaction and distortion techniques would assist in supporting HEIs' policy makers' decision making based on the explicit knowledge identified and suitable .

Furthermore, the prototyping is the translation of the conceptual design framework to physical artefact. The realization of *StudentViz* (see section 5.4) showed that the framework is viable, practical and applicable. *StudentViz*, which is an InfoVis prototype for HEI students' data, is one of the contributions of this study. Also, the findings from heuristic evaluation (see section 5.5) showed that the HSDI prototype supports data representation correctness, and has satisfactory usability quality (see section 5.5.2). The design of the InfoVis prototype equally attends to the demand (for data tool) of the conceptual model for data-driven decision making in HEI (see Figure 2.2). This is also one of the contributions of this study.

Lastly, the findings from the users' experience evaluation study (see section 6.4), showed that the consequent InfoVis tool; *StudentViz*, supported decision making based on the explicit knowledge identified. The subjective (quantitative) feedback from the HEI policy makers implied that the prototype can be used for visually exploring students' data and support users' control. This certified the suitability of the distortion and interaction techniques. The data representation correctness is also attested to. It showed that the communicative graphics which are the visualization techniques are compatible with the data object relationship, and aligns with the domain explicit knowledge preferences.

The objective (qualitative) feedback from the participants suggested that the HSDI prototype supports decision making activity. This is evident in the HEIs' policy makers' assessment of the prototype's abilities. The HSDI prototype is assessed to be able to locate subset of data that is of interest, differentiate between data representation, categorise data objects accordingly, show different view and details of data, and reveal previously unknown information. These abilities and functionalities collectively produce actionable insights, thus support decision making process. Therefore, it showed that the conceptual design framework is practically usable in designing InfoVis that supports correct data representation, usability, and decision making.

Therefore, the HSDI prototype's abilities to "locate subset of data that is of interest" through filtering, and zooming; "to differentiate between data representation", and "categorise data objects accordingly"; "show different view and details of data"; and "reveal previously unknown information" attest that the prototype conforms with provisions of Sheinerderman and Plaisant (2010), Spence (2007), Ware, 2000), Card et al. (1999), Lin (1997), and Sheneiderman (1996). The summary of what the features of InfoVis must possess (as earlier presented in Table 2.2) are supported by HSDI prototype. It allows less cognitive load (Card et al., 1999; Lin, 1997), and duly supports the information seeking mantra; "Overview first, zoom and filter, then details on demand" (Sheinerderman & Plaisant, 2010; Sheneiderman, 1996).

7.5 Revisiting the Objectives of the Study

The main aim of this study is to propose a HEI students' data-focused InfoVis conceptual design framework. The framework is developed to guide InfoVis designers and developers in developing InfoVis prototype that attends to the specifics of HEI students' data, the domain explicit knowledge preferences, and thus, aid the related decision making activities. To achieve this main aim, four specific objectives are formulated. These are (a) to categorise the appropriate data dimensions of the HEIs' students data based on the identified and prioritized explicit knowledge preferences, (b) to elicit the analysed visualization, interaction and distortion techniques which are suitable for the representation and presentation of the identified explicit knowledge preferences, (c) to design an InfoVis conceptual framework that is suitable for the representation and presentation of the identified explicit knowledge preferences, and (d) to evaluate the InfoVis conceptual framework.

The main aim is achieved through the accomplishments of the specific objectives. A pre-design study and users' requirements validation are used to achieve the first objective. Content analysis of InfoVis design literatures in view of the findings of the pre-design study is used to achieve the second objective. The third objective is achieved by developing the findings from second objective to diagrammatic representation and showed connection with the design processes and the conceptualized specifications. The fourth objective is achieved through combination of approaches: verification of the framework, prototyping, heuristics and users' evaluation. Collectively, the proposed conceptual framework is found to be practical, viable and applicable in designing students' data-focused InfoVis that handles the data multidimensionality, reduces the cognitive load, and supports decision making.

7.6 Contributions of the Study

This study contributes generally to InfoVis design knowledge and HEI as a domain. These contributions fall under the practical, functional and the theoretical contribution.

7.6.1 InfoVis Conceptual Design Framework

This study proposed an InfoVis conceptual design framework. This domain-specifc framework is novel, and would benefit the InfoVis designers. It would be used as a design guide to deliver an InfoVis system that captures HEIs' students' data and support the associated decisions by delivering the domain's explicit knowledge preferences. The framework consists of content delivery and process models. It shows the phases involved in designing HEI students' data-focused InfoVis, and how InfoVis conceptual design can be translated into physical InfoVis tool. It also presents specific InfoVis' techniques that align with the HEI domain specificity, with theoretical design guide. It can equally be adapted to other related domains. The proposed InfoVis conceptual design framework is a theoretical contribution.

7.6.2 An InfoVis Prototype for HEI students' data

This study delivered an InfoVis prototype following the proposed InfoVis conceptual design framework. The findings from the domain's pre-design study showed the varieties of explicit knowledge related to students' data of HEIs. The prototype is developed to solve the problem of information overload in the management of HEI students' data by practically ensuring that the study's pre-design findings are utilized. It is to be used by HEI policy makers, data administrators and analysts since they interact with students' data for administrative and decision making purposes.

7.6.3 Conceptual Framework for Data-driven decision making in HEI

This is a diagrammatic representation of the data usage culture of HEI. It is developed based on review of literatures to position data-driven decision making in

HEI. It is also to guide HEI policy makers on the dynamics of data usage in HEI and the strength of data-driven decision making process.

7.7 Limitations of the Study and Recommendations for Future Works

There are few limitations encountered in this study. Even though they did not have direct bearing on the findings of this study, future researches are encouraged for further improvements. First, in the proposed structural model, multidimensional data table is suggested because as it is presently, the suggested data dimensions are decentralized in different data locations within the HEI. If decentralization for the purpose of data security remains a better argument than centralized form for easier data management, then future work is necessary. This will be to determine the specific data object binding system for necessary exploration and visualization in a decentralized but distributed data sources' environment.

Second, most of the data dimensions (e.g. soft skill quality, health weakness, employability, etc.) proposed for delivery of the domain explicit knowledge preferences do not presently exist as basic data dimensions in HEIs' students' data tables. They are rather computed by disparate units handling the data respectively, on demand. This will not definitely match student-by-student data accuracy. Due to this, real life data that captured real life scenario is unavailable for prototyping. Future researches, which are only feasible when HEIs have started documenting all these data dimensions on per student basis, are necessary for improved user-centred design studies. This would possibly suggest other elements for better decision making supports.

Third, future research on a real time users' experience experimental study will be necessary. In this, apart from evaluating the overall functionality of the InfoVis which this study did, task execution time can be studied. Also, perceptive data from users' emotional reaction, and eye gazes, while interacting with the InfoVis can be captured for further deeper insights.

Fourth, the interaction techniques that are specified in the proposed conceptual design framework are minimally chosen based on Cooper et al.'s (2014) maxim of minimalist design. The researcher opines that even though this achieves the objectives of the study, richer *IxD* and more users' control mechanism can be employed for deeper insights and enhanced users' experience.

7.8 Conclusion

This research has proposed and produced the HEI students' data-focused InfoVis conceptual design framework. It combined the specifications of the content delivery techniques with the theoretical design model. This highlights the interconnection between theories, design features, and InfoVis' techniques. The conceptual design framework describes the behaviour of the HEI students' data-focused InfoVis through its interaction and distortion techniques, and the form through the visualization techniques. Preattentive processing theory, cognitive fit theory and normative and descriptive theories are supporting theories for the design of the visualization, interaction and data exploration, respectively. Collectively, an InfoVis that supports perception, cognition and decision making is produced, and explicit knowledge preferences of the HEI domain decision makers are adequately attended to. Despite necessary future researches, this thesis demonstrates the development process of domain-specific InfoVis conceptual design framework. The findings obtained equally showed that the proposed framework is significantly satisfactory and the prototype supports correct data representation, usability and decision making.

This implies that InfoVis designers and developers now have a design guide to leverage on in the design and development of HEI students' data-focused InfoVis. The technique specifications provide the functional requirements for an InfoVis that handles students' data multidimensionality, solves the problem of information overload and support the decision making process. The process and structural models of the conceptual design framework are absolutely generalizable. The content delivery techniques are specifically for HEI students' data-focused explicit knowledge but adaptable to closely-related domains of teaching and learning institutions. The HSDI prototype can be used by anyone interested in actionable insights from students' data, especially the HEI data analyst, data administrator, and policy makers. In conclusion, this study does not only demonstrate conceptualization of InfoVis design framework, it addresses HEI domain-specific, user-centred and task-based developmental process.

REFERENCES

Alavi, M., & Joachimsthaler, E.A. (1992). Revisiting DSS implementation research: a meta-analysis of the literature and suggestions for researchers, *MIS Quarterly*, *16* (1), 95–116

Ankerst, M., Keim, D. A., & Kriegel, H.-P. (1996). Circle segments: A technique for visually exploring large multidimensional datasets, *in Proc. Visualization 96*, Hot Topic Session, San Francisco, CA, 1996.

Ariffin, M. (2009). *Conceptual Design of Reality learning Media (RLM) Model Based on Entertaining and Fun Construct.* A PhD Thesis submitted to College of Arts and Science, Universiti Utara Malaysia.

Aris, A., & Shneiderman, B. (2007). Designing semantic substrates for visual network exploration, *Journal of Information Visualization*, *6* (4), 281–300.

Arkhipenkov, S., & Goluber, D. (2001). *Oracle Express Olap*., USA: Charles River Media.

Ayres, L. (2007). Qualitative Research Proposal – Part I: Posing the problem. *Journal of Wound Ostomy Continence Nursing*, *34*, 30–32.

Babbie, E. (2010). *The Practice of Social Research*, 12th Edition. USA: Wadsworth Cengage Learning.

Baker, D., Bridges, D., Hunter, R., Johnson, G., Krupa, J., Murphy, J. & Sorenson, (2002). *Guidebook to Decision-Making Methods*, WSRC-IM-2002-00002, USA: Department of Energy.

Balakrishnam, N., Render, B. & Stair, M.R. (2012). *Managerial Decision Making with Spread sheet*, 2nd Edition. USA: Pearson International.

Barab, S., & Squire, K. (2004). Design-based research: putting a stake in the ground. *Journal of the Learning Sciences*, 13, 1-14. DOI: 10.1.1.128.5080

Battista, V., & Cheng, E. (2011). Motion Charts: Telling Stories with Statistics. *In Section on Statistical Graphics*, JSM 2011. Retrieved from http://www.bls.gov/osmr/pdf/st110110.pdf on 3rd October, 2016.

Benoit, G. (2012). Measuring Relevance Theory's Effects in an Image-Driven Explorative Information System, *Proceedings of iConference, 2012*, Toronto, Canada. ACM, pp. Pages 438 – 440.

Bier, E. A., Stone, M. C., & Pier, K., Buxton, W., & DeRose, T. (1993). Toolglass and magic lenses: The see-through interface, *in Proc. SIGGRAPH '93*, Anaheim, CA, 1993, pp. 73–80.

Bogard, M., & Green, B. (2013). *A Data Driven Analytic Strategy for Increasing Yield and Retention at Western Kentucky University Using SAS Enterprise BI* *and SAS Enterprise Miner*, Paper 044, presented at SAS Global Forum, 2013, pp. 1 – 12.

Bonner, N. (2008). *Acceptance of Systems Development Methodologies: Testing A Theoretically Integrated Model.* PhD Thesis. The University of Texas Arlington.

Bowen, J.P. (2013). Online Communities: Visualization and Formalization, *Proceedings of the Second International Workshop on Cyberpatterns: Unifying Design Patterns with Security, Attack and Forensic Patterns*. Abingdon, UK, 8-9 July 2013

Boy, J., Detienne, F., & Fekete, J-D. (2015). Can Initial Narrative VisualizationTechniques and Storytelling help Engage Online-Users with Exploratory Information Visualizations? *2015 Association for Computing Machinery, CHI 2015*, April 18 - 23 2015, Seoul, Republic of Korea http://dx.doi.org/10.1145/2702123.2702452

Bresfelean, V.P., Ghisoiu, N., Lacaurezeanu, R., & Sitar-Taut, D.A (2009). Towards the Development of Decision Support in Academic Environments, *Proceedings of International Conference on Information Technology Interfaces*, ITI, '09, pp. 343- 348

Brooke, J. (2009). *SUS: A quick and dirty usability Scale*, In Jordan, P.W., Thomas, B., Weerdmeester, B.A., & McClelland, I.L. (Eds) Usability Evaluation In Industry. London: Taylor & Francis.

Buja, A., Cook, D., & Swayne, D. F. (1996). *Interactive High-Dimensional*. UK: Wiley

Card, S.T., Mackinlay, J. & Shneiderman, B. (1999). *Readings in Information Visualization: Using Vision to Think*, United States of America: Morgan Kaufmann Publishers.

Carrasco, J-A., Hogan, B., Wellman, B., & Miller, E. J. (2006). Collecting Social Network Data yo Study Social Activity-Travel Behaviour: An Egocentric Approach, presented at the 85th *Transportation Research Board Meeting*, Washington DC, January 22-26, 2006.

Carroll, J. M. (2000). *Making use scenario-based design of human-computer interactions.* Cambridge, MA: MIT Press.

Chan, Y.H., Correa, C.D., & Ma, K-L. (2013). The Generalized Sensitivity Scatterplots. *IEEE Transaction on Visualization and Computer Graphics*, IEEE computer Society 19(10), pp. 1768 - 81

Chenail, R. J. (2012). Conducting Qualitative Data Analysis: Reading Line-by-Line, but Analyzing by Meaningful Qualitative Units. *The Qualitative Report*, 17 (1), 266 – 269. Retrieved from http://www.nova.edu/ssss/QR/QR17-1/chenail-line.pdf on 4th December, 2015.

Chi, h H. (2000). A taxonomy of visualization techniques using the data state reference model. *In Proceedings of the IEEE Symposium on Information Visualizations 2000*, INFOVIS ’00, pages 69 – 76 Washington, DC, USA, 2000. IEEE Computer Society.

Chin, J.P., Diehl., V.A., & Norman, K.L. (1988). Development of an Instrument Measuring User Satisfaction of the Human Computer Interface. *ACM CHI* ’88 proceedings, pp. 213-218

Christina, N., Julian, A., & Micheal, S. (2016). Visual Exploration of Media Transparency for Data Journalists: Problem Characterization and Abstraction, Retrieved from: http://www.validproject.at/wp--content/uploads/2015/11/151031_AusserhoferGutounigOppermann--Dubrovnik_Dubrovnik.pdf, 15 August, 2016.

Chuah, M. C., & Roth, S. F. (1996). On the Semantics of Interactive Visualizations, *Proceedings of IEEE Symposium on Information Visualization*, San Francisco, CA, pp. 29- 36.

Codd, E.F., Codd S.B., & Salley G.T. (1993). *Providing OLAP (Online Analytical Processing) to User-Analysis: AN I.T Mandate.* Hyperion Solution Corporation. Retrieved from: http://www.hyperion.com/solution/whitepapers.cfm on 15/08/13

Cohen, J. W. (1988). *Statistical power analysis for behavioural sciences* (2nd edn). Hillsdale, NJ: Lawrence Erlbaum Associates

Collier, K. (2015). You’ve Got the Data, Now What? Data Driven Decisions for School Improvement. Retrieved from http://engage.ok.gov/wp-content/uploads/2015/07/98-Data-Driven-Decision-Making-PPT-June-8-2015.pdf on 15th August, 2016.

Connolly, T. & Begg, C. (2010). *Database Systems: A Practical To Design, Implementation, and Management.* Fifth Edition, Boston: Pearson Education Inc.

Conrad, C., Gasman, M., Lundberg, T., Nguyen, T-H., Commodore, F., & Samayoa, A.C. (2013). *Using Educational Data to Increase Learning, Retention, and Degree Attainment at Minority Serving Institutions (MSIs).* A Research Report of Penn Graduate School of Education, GSE.

Cooper, A., Reimann, R., Cronin, D., & Noessel, C. (2014). *About Face: Essentials of Interaction Design*, Fourth Edition. USA: Wiley.

Cordeiro, R.L., Traina, C., Traina, A.J.M., Lopez, J., Kang, U. & Faloutsos, C. (2011). Clustering Very Large Multi-dimensional Datasets with MapReduce, *in Proceedings of KDD, ACM,* San-Diego, California, USA.

Crapo, A.W., Waisel, L.B., Wallance, W.A. & Willemain, T.R. (2000). Visualization and The Process of Modelling: A Cognitive-theoretic View, *Proceedings of KDD*, USA, ACM.

Creswell, J.W. (2009). *Research Design: Qualitative, Quantitative and Mixed Methods Approaches.* Third Edition, United States of America, Sage.

Croll A. (2013). *Big Data Is Our Generation's Civil Rights Issue and We Don't Know It (*In Eds) Big Data Now: A Collection of O' Reilly Perspectives on Big Data.

Delavari, N., Phon-Amnuaisuk, S., & Beikzadeh. M.R. (2008). Data Mining Application in Higher Learning Institutions, *Vilnus*, *7*(1), 31-54.

Denscomber, M. (2014). *The Good Research Guide: For Small-Scale Social Research Projects.* UK: Open University Press

Dix, A. & Ellis, G. (1998). Starting simple: adding value to static visualisation through simple interaction, *In the Working conference on Advanced visual interfaces* (AVI '98), L'Aquila, Italy, pp. 124-134, 1998.

Druin, A. (2002). The Role of Children in the Design of New Technology. *Behaviour and Information Technology*, 21(1), 1-25. DOI: 10.1080/01449290110108659

Druzdzel, M., & Flynn, R. (2002) *Decision Support System*. Encyclopedia of Library and Information Science, Second Edition, Allen Kent (ed.) New York: Marcel Dekker, Inc.

Dumbill, E. (2012). *Big Data Now: A Collection of O' Reilly Perspectives on Big Data In Eds).*

Eduventure Inc. (2013). *Predictive Analytics in Higher Education: Data-Driven Decision Making for the Student Life Cycle,* White Paper Publication by Eduventure Inc., Boston.

Ehlers, K., Joubert, M., Kinghorn, J., & van Zyl, A. (2009). Decision Support System for Institutional Research Management in Higher Education: Data Mining to determine Research Focus, Intensity and Synergy, *Proceedings of International Conference on Computational Science and Engineering*, pp. 885-890

El-Fattah, M.A. (2012). Visual Model for Managing Educational Capacity Utilization in Egyptian Universities, *in Proceedings of International Conference on Information Technology Based Higher Education and Training*, pp. 1-4.

Elmqvist, N., Stasko, J., & Tsigas, P. (2007). DataMeadow: A Visual Canvas for Analysis of Large-Scale Multivariate Data, *Proc. IEEE Symposium in Visual Analytics Science and Technology* (VAST '07), pp. 187-194

Elmqvist, N., Dragicevic, P., & Fekete, J.-D. (2008). Rolling the dice: Multidimensional visual exploration using scatterplot matrix navigation. *IEEE Transaction on Visualization &. Computer Graphics*, 14 (6), 56 – 60.

Elo S, Kyngäs H. (2008). The Qualitative Content Analysis Process. *Journal of Advanced Nursing*, *62*, 107–115.

Ewell, P.T. (2012). Using Information in Higher Education Decision-Making: Modes, Obstacles, and Remedies. Retrieved from: www.spencer.org/resources/content/3/3/8/documents/Ewell-paper.pdf on Aug. 18, 2013

Faisal, S., Craft, B., Cairns, P. & Blandford, A. (2008) Internalization, Qualitative methods, and Evaluation, *In Proceedings of the 2008 conference on Beyond time and errors: novel evaluation methods for Information Visualization*, pp. 1-8

Fekete, J-D. & Plaisant, C. (2002). Interactive Information Visualization of a Million Items, *Proceedings of the IEEE Symposium on Information Visualization* (InfoVis '02), pp. 117-124.

Fekete, J-N., van Wijk, J.J., Stasko, J.T. & North, C. (2008). The Value of Information Visualization. *Information Visualization: Human-Centred Issues and Perspectives* (E.d.), pp. 1-18.

Fishkin, K., & Stone, M. C. (1995). Enhanced dynamic queries via movable filters, *Proceeding Human Factors in Computing Systems*, CHI '95, pp. 415–420

Forsell, C. (2012). Evaluation in Information Visualization: Heuristic Evaluation, In *Proceedings of 16th International Conference on Information Visualization*, pp. 136-142

Frankel, L. & Racine, M. (2010). The Complex Field of Research: for Design, through Design, and about Design. In E.S. David Durling, Rabah Bousbaci, LinLin Chen, Philippe Gauthier, Tiu Poldma, Seymour Roworth-Stokes (Eds.), Design & Complexity: *International Conference of the Design Research Society.*

Friendly, M. & Denis, D. (2005). The Early Origins and Development of the Scatter plot. *Journal of the History of the Behavioural Sciences*, *14* (2), 103-130.

Gelman, A. & Unwin, A. (2012). *Infovis and Statistical Graphics: Different Goals, Different Looks.* Readings from a departmental seminar held at University of Kentucky.

Gill, B., Borden, B.C., & Hallgren, K. (2014). *A Conceptual Framework for Data-Driven Decision Making*. Final Report of Research conducted by Mathematica Policy Research, Princeton, submitted to Bill & Melinda Gates Foundation, Seattle, WA.

Goss, P., & Hunter, J. (2015). *Targeted Teaching: How better use of data can improve student learning.* A White Paper Publication of Gratian Institute. Retrieved from http://grattan.edu.au/wp-content/uploads/2015/07/827-Targeted-Teaching.pdf on 15 August, 2016

Graham, D. (2005). Information Visualization Theory and Practice, *Proceedings of the Ninth International Conference on Information Visualization* pp. 3-8

Greenberg, S. (2002) Working through Task-Centered System Design, *in Diaper, D. and Stanton, N. (Eds) The Handbook of Task Analysis for Human-Computer Interaction*. Lawrence Erlbaum Associates.

Hackos, J. T., & Reish, J. C. (1998). *User and Task Analysis for Interface Design*. Canada: Wiley.

Halverson, C. A. (2002). Activity Theory and Distributed Cognition: or What does CSCW need to do with Theories? *Computer Supported Cooperative Work*, *11*, 243–267

Hao, M. C., Dayal, U., Sharma., Keim, D., & Janetzko, H. (2010). Visual Analytics of Large Multi-Dimensional Data Using Variable Binned Scatter Plots, *in Visualization and Data Analysis Proceedings of IS & T/SPIE Electronic Imaging, Science and Technology*, Belligham.

Hao, M., Rohrdantz, C., Janetzko, H., Keim, D., Dayal, U., Haug, L. & Hsu, M. (2012). Integrating Sentiment Analysis and Term Association with Geo-Temporal Visualization on Customer Feedback Streams, *Proceedings SPIE 8294, Visualization and Data Analysis,* 2012

Harrison, K., Bowen, J.P. & Bowen, A.M. (2013). Electronic Visualisation in Chemistry: From Alchemy to Art, *Proceedings of EVA London 2013 Conference, Electronic Workshops in Computing (eWiC),* British Computer Society, 29-31 July 2013.

Healey, C.G. (1996). *Effective Visualization of Large Multidimensional Datasets*. A PhD Thesis in Partial Fulfilment for the Degree of Doctor of Philosophy (Ph.D). The University of British Columbia.

Hecksel, D. (2004). *Methodology evaluation and selection*. Sun Microsystems. Retrieved Jan 15, 2014 from http://www.davidhecksel.com/projectcontext/whitepaper.html

Hevner, A. R., March, T. S., Park, J., & Ram, S. (2004). Design science research in information systems. *Management Information Systems Quarterly*, *28* (1), 75-105. DOI: 10.2307/249422

Hollan, J. D., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: Toward a new foundation for human-computer interaction. *ACM Transactions on Computer-Human Interaction*, *7* (2), 98 - 104

Hollingworth, L., Sullivan, A.M., Condon, C., Bhatt, M. & Brandt, W. C. (2012). Data-Driven Decision Making in Higher Education: One University's Process of Revamping the Superintendent Licensure Program, *Journal of Research on Leadership Education, 7* (1), 78–97.

IBM Data Profile (2012). *IBM Big Data Success Stories*. A White Paper Publication of IBM Inc.

Inselberg, A., & Dimsdale, B. (1990). Parallel coordinates: A tool for visualizing multi-dimensional geometry, *in Proc. Visualization 90, San Francisco*, CA, pp. 361- 370.

Institute for Higher Education Policy (IHEP). (2009). *Impact of College Rankings on Institutional Decision Making: Four Country Case Studies*. Retrieved at: http://www.ihep.org/publications/publications-detail.cfm?id=126 on August, 18, 2013.

Isenberg, P., Tang, A., & Carpendale, S. (2008). An Exploratory Study of Visual Information Analysis, *CHI Proceedings: Visualization to Support Information Work*, pp. 1217-1226

Jackson, S. S. (2013). *A continuous Improvement Framework: Data Driven Decision-Making in Mathematics Education*, A White Paper Publication of Data-Driven Innovation Consulting, Inc.

Janetzko, H., Hao, M.C., Mittelstadt, S., Dayal, U. & Keim, D. (2013). Enhancing Scatter Plots Using Ellipsoid Pixel Placement and Shading, *Proceedings of the 46th Annual Hawaii International Conference on System Science*, IEEE Computer Society, pp. 1522- 1531.

Johansson, J., Nest, T.S. & Linner, B. (2010). Evaluating Climate Visualization: An Information Visualization Approach, *Proceedings of IEEE 2010 14th International Conference on Information Visualization*, pp. 156- 161.

Julie, A., Marsh, J. F., & Laura S. H. (2006). *Making Sense of Data-Driven Decision Making in Education.* Rand Occasional Paper, RAND Corporation.

Kaufman, T., Graham, C. R., Picciano, A. G., Wiley, D., & Popham, J. A. (2014). Data-driven decision making in the k12 classroom. *In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), Handbook of research on educational communications and technology* (4th ed., pp. 337–346). New York, NY: Springer

Karine, T., Diane, L., & Deborah, R. (2012). *Assessment of Higher Education Learning Outcomes*. Feasibility Study Report. Volume 1. OECD Better Policies for Better Lives. Retrieved from http://www.oecd.org/education/skills-beyond-school/AHELOFSReportVolume1.pdf, on 3rd October, 2016

Keim, D. & Ward, M. (2002). *Visual Data Mining Techniques. Published in Intelligent Data Analysis: An Introduction*/ Micheal Berthod (eds), Berlin. Springer.

Keim, D. A. (2002). Information Visualization and Visual Data Mining, *IEEE Transactions on Visualization and Computer Graphics, 8*, 1-8

Keim, D. A., Kriegel, H.-P., & Ankerst, M. (1995). Recursive pattern: A technique for visualizing very large amounts of data, in Proc. Visualization 95, Atlanta, GA, pp. 279–286.

Keim, D., Andrienko, G., Fekete, J-D., Gorg, C., Kohlhammer, J. & Melancon, G. (2008). Visual Analytics: Definition, Process, and Challenges. A Kerren et al. (Eds): *Information Visualization*, *LNCS* 4950, pp. 154-175. Springer-Verlag Berlin Heidelberg.

Keim, D., Hao, M.C., Daya., U., Janetzko, H., & Bak, P. (2009). Generalized Scatter Plots. *Information Visualization*, pp. 1-11, Macmillan Publishers Ltd.

Keim, D., Kohlhamer, J., Ellis, G. & Mansmann, F. (2010a). *Mastering the Information Age: Solving Problems with Visual Analytics.* Germany, Eurographics Association.

Keim, D., Pras, A., Schonwalder, J., Wong, P.C. & Mansmann, F. (2010b). Report on the Datgstuhl Seminar on Visualization and Monitoring of Network Traffic, *Journal of Network and Systems Management*, *18* (2), 232-236.

Keim, D.A., Mansmann, F., Oelke, D. & Ziegler, H. (2008). Visual Analytics: Combining Automated Discovery with Interactive Visualization, *Proceedings of the 11th International Conference on Discovery Science* (DS,2008), pp. 2-14

Kerzner, H. (2006). *Project Management: Best Practices: Achieving Global Excellence*. USA: Wiley.

Koh, L.C., Slingsby, A., Dykes, J., & Kam, T.S. (2011). Developing and Applying a User-centred Model for the Design and Implementation of Information Visualization Tools, *Proceedings of 15th IEEE International Conference on Information Visualization*, IEEE Computer Society, pp. 90-95.

Kohlhammer, J., Ruppert, T., Darvey, J., Mansmann, F. & Keim, D. (2010). Information Visualization and Visual Analytics for Governance and Policy Modelling, *CROSSROAD Call for Contribution on Future Internet on ICT for Governance and Policy Modelling*, 2010, pp. 251-252

Kopáčková, H. & Škrobáčková M. (2006). *Decision Support Systems or Business Intelligence: What Can Help In Decision Making*?

Krstajic, M., Najm-Araghi, M. Mansmann, F. & Keim, D. (2012). Incremental Visual Text Analytics of News Story Development, *Proceedings of Conference on Visualization and Data Analysis* (VDA '12), pp. 308-323

Kucuktunc, O., & Ferhatosmanoglu, H. (2013). λ-Diverse Nearest Neighbours Browsing for Multidimensional Data, *IEEE Transaction on Knowledge and Data Engineering, 25* (3), 481-493

Lam, H., Bertini, E., Isenberg, P., Plaisant, C. & Carpendale, S. (2012). Empirical Studies in Information Visualization: Seven Scenario, *IEEE Transaction on Visualization and Computer Graphics, 18* (9), 1520-1536

Lange, C., Range, B., & Welsh, K. (2012). *Conditions for Effective Data Use to Improve Schools: Recommendations for School Leaders*. Retrieved at: http://cnx.org/content/m45021/latest/ on August, 18, 2013

Lau, A. & Moere, A. V. (2007). Towards a Model of Information Aesthetics in Information Visualization, *in Proceedings of 11th International Conference on Information Visualization,* IEEE Computer Society

LeBlanc, J., Ward, M. O., & Wittels, N. (1990). Exploring n-dimensional databases, *in Proc. Visualization '90, San Francisco*, pp. 230–239.

Lee, B., Czerwinski, M., Roberston, G. & Bederston, B.B. (2005). Understanding Research Trend using PaperLens, *Proceedings CHI 2005*, Portland, Oregon, USA.

Leventhal, L., & Barnes, J. (2007). *Usability Engineering: Process, Products and Examples*. USA: Pearson

Lewis, J.R. (1995). IBM computer usability satisfaction questionnaire: Psychometric evaluation and instruction for use. *International Journal of Human-Computer Interaction, 7* (1), 57-58

Li, I., Hook, K., Dey, A., Forlizzi, J., & Medynskiy, Y. (2011). Personal Informatics and HCI: Design, Theory and Social Implications, *Proceedings of CHI, 2011,* Vancouver, Canada. ACM.

Lin, X. (1997). Map displays for information retrieval. *Journal of the American Society for Information Science, 48* (1), 40–54

Lirong, X., Mengium, W. & Jing, F. (2011). A Visualization System for Web Retrieved Credit Information, *Proceedings of seventh International Conference on Natural Computation,* pp. 728-732.

Liu, Z. & Stasko, J. (2010a). Theories in Information Visualization: What, Why and How, VisWeek 2010, The Role of Theory in Information Visualization. ACM Workshop.

Liu, Z., & Stasko, J. (2010b). Mental Models, Visual Reasoning and Interaction in Information Visualization, *IEEE Transaction on Visualization and Computer Graphics, 16* (6), 999- 1008

Liu, Z., Nersessian, N. J. & Stasko, J. T. (2008) Distributed cognition as a Theoretical Framework for Information Visualization, *IEEE Transactions on Visualization and Computer Graphics, 14* (6), 1173 – 1180

Luna-Reyes, L. F., & Andersen, D. L. (2003). Collecting and Analysing qualitative data for system dynamics: methods and models. *System Dynamics Review, 19* (4), 271-296. DOI: 10.1002/sdr.280

Maciejewski, R., Rudolph, S., Hafen, R., Abusalah, A., Yakout, M., Ouzzani, M., Cleveland, W.S., Grannis, S.J., & Ebert, D.S. (2010). A Visual Analytics Approach to Understanding Spatiotemporal Hotspots, *IEEE Transactions on Visualization and Computer Graphics, 16* (2), 205-220.

Mandinach. E.B., Gummer, E.S & Muller, R.D. (2012). *The Complexities of Integrating Data-Driven Decision Making into Professional Preparation in Schools of Education: It's Harder than You think.* A Report Summary

Mansmann, F., Fisher, F., Keim, D.A. & North, S.C. (2009). Visual Support for Analyzing Network Traffic and Intrusion Detection Events using TreeMap and Graph Representations, *Proceedings of the 3rd ACM Symposium on Computer Human Interaction for the Management of Information Technology* (CHiMiT '09), page 19-28.

Matsui, K., Yamanouchi, M., & Sunahara, H. (2011). A Proposal of Framework for Information Visualization in Developing of Web Application, *Proceedings of 2011 IEEE/IPSJ International Symposium on Application and the Internet*, pp. 457- 462, IEEE Computer Society

Mazza, R., & Berre, A. (2007). Focus Group Methodology Evaluating Information Visualization Techniques and Tools (2007). *Proceedings of IEEE 11th International Conference on Information Visualization*, 2007. IV '07, pg. 74-80., Zurich.

Merriam-Webster, Inc., Merriam-Webster Online Dictionary, First published on 2007, retrieved from http://www.m-w.com/, on 22nd, September, 2013

Meyer, M. (2012). *Designing Visualization for Biological Data*, A Keynote talk at 3rd Leonardo Satellite Symposium of Arts, Humanities and Complex Networks (NetSci, 2012)

Microsoft Inc. (2004). *Data Driven Decision Making in Higher Education*. Discover, A White paper publication of Microsoft Incorporation, released in February, 2004.

Modified Assured Career Progression Scheme (MACPS) (2009). *Government of lndia Ministry of Personnel, Public Grievances and Pensions (Department of Personnel and Training)* North Block, New Delhi, the 19th May, 2009.

Neuman, L.W. (2007). *Basics of Social Research: Qualitative and Quantitative Approaches*, 2nd Edition. Pearson International, USA.

Nielsen, J. (1993). *Usability Engineering*, Boston, Massachusetts: Academic Press, Inc.

Ning, Z., Jiaxin, W., Bing, W., & Shaolong, Z. (2008). A Visualization Model for Information Resources Management, *in Proceedings of 12th International Conference on Information Visualization*, pp. 57- 62, IEEE Computer Society

Ning, Z., Wenxing, H., Siting, Z. & Li, T. (2012). A Solution for an Application of Information Visualization in Telemedicine. *Proceedings of the 7th International Conference on Computer Science & Education (*ICCSE, 2012), Melbourne, Australia, pp. 407-411.

Norman, D. (2002). *The Design of Everyday Things*, New York: Basic Books.

O'Reilly Media (2012). B*ig Data Now: A Collection of O' Reilly Perspectives on Big Data*. O'Reilly Strata Making Data Work

Oelke, D., Bak, P., Keim, D.A., Last, M. & Danon, G. (2008). Visual Evaluation of Text Features for Documents Summarization and Analysis, *Proceedings of IEEE Symposium on Visual Analytics Science and Technology*, Columbus, Ohio, USA

Oelke, D., Hao, M., Rohrdantz, C., Keim, D.A., Dayal, U., Haug, L.E. & Janetzko, H. (2009). Visual Opinion Analysis of Customer Feedback Data, *Proceedings of the 2009 IEEE Symposium on Visual Analytics Science and Technology* (VAST '09), pp.187-194

Oelke, D., Janetzko, H., Simon, S., Neuhaus, K., & Keim, D. (2011). Visual Boosting in Pixel-based Visualizations, *Eurographics/IEEE Symposium on Visualization*, 30 (3), 6- 15.

OLAP Council (1998). APB-1 OLAP Benchmark Release II. Retrieved from http://www.olapcouncil.org/research/bmarkco.html on August 15, 2013 .

Oracle Data Sheet (2011). Oracle Higher Education Constituent Hub (HECH). Retrieved on 6th June, 2013 from http://www.oracle.com/us/products/applications/master-data-management/education-research-hech-ds-164687.pdf

Parsons, P., & Sedig, K. (2011). Human-information interaction: An emerging focus for educational cognitive tools, *Proceedings of FORMATEX, in Education in a technological world: communicating current and emerging research and technological efforts, A.Medez-Vilas (Ed)*, pp. 245- 251

Pinto, M., & Carvalho, D. (2014). U-TRACER: *The Use of Communication Technology in Higher Education*, Universidade de Aveiro, Ano, 2014. Retrieved from https://ria.ua.pt/bitstream/10773/13140/1/u%20tracer_%20o%20uso%20das%20tecnologias%20da%20comunicacao%20no%20ensino%20superior.pdf on 3rd October, 2016.

Pinto, M., Raposo, R. & Ramos, F. (2012). Comparison of Emerging Information Visualization Tools for Higher Education, *Proceedings of 16th International Conference on Information Visualization*, pp. 100-105

Playfair, W. (1801). *The Statistical Breviary; Shewing, on a Principle Entirely New, the Resources of Every State and Kingdom in Europe*, Wallis.

Pocius, K. & Reklaitis, V. (2000). Web-based Services for Technology Education Management and Planning, *Proceedings of International Workshop on Advanced Learning Technologies, IWALT, 2000*, pp. 284- 285.

Protheroe, N. (2009). *Improving Teaching and Learning with Data-Based Decisions: Asking the Right Questions and Acting on the Answers*. Education Research Centre, Research Report.

Protheroe, N. (2010). *Schools as Effective Data Users*. Research Report of www.naesp.org, October/November, 2010.

Purao, S. (2002). *Design research in the technology of Information system: Truth or Dare*. GSU Department of CIS working paper. Atlanta, GA.

Purchase, H. C., Andrienko, N., Jankun-Kelly, T., & Ward, M. (2008). *Theoretical foundations of information visualization*. In Information Visualization - Human-Centered Issues and Perspectives

Quigley, A.J. (2002). Experience with FADE for the Visualization and Abstraction of Software Views, *Proceedings in 10th International Workshop on Programme Comprehension*, IWPC '02

Quinlan, P. T. & Humphreys, G. W. (1987). Visual search for targets defined by combinations of color, shape, and size: An examination of task constraints on feature and conjunction searches, *Perception & Psychophysics, 41* (5), 455-472.

Reigeluth, C. M. (2008). Instructional-design theories and models: Building a common knowledge base, pp. 83.

Ren, L., Cui, J., Du, Y., & Dai, G. (2013). Multilevel Interaction Model For Hierarchical Tasks in Information Visualization, *In VINCI 2013*, Tianjin, China, pp. 11- 15

Reshef, D.N., Reshef, Y.A., Finucane, H.K., Grossman, S.R., McVean, G., Turnbaugh, P.J., Lander, E.S., Mitzenmacher, M., & Sabeti, P.C. (2011). Detecting Novel Associations in Large Data Sets. *Journal of Science Mag*, *33* (4), 1518- 1524

Robertson, W., & DeFerrari, C. (1999). Systematic approaches to visualization: is a reference model needed? Retrieved from http://www.siggraph.org/education/materials/HyperVis/concepts/ref_model.htm on 30th August, 2016

Robinson, A. C., Chen, J., Lengerich, E. J., Meyer, H. G. & MacEachren, A. M. (2005). Combining Usability Techniques to Design Geovisualization Tools for Epidemiology, *Cartography and Geographic Information Science*, *32* (4), 243–255

Rogers, Y., Sharp, H. & Preece, J. (2011). *Interaction Design: Beyond Human Computer Interaction*, 3rd Edition, United Kingdom. John Wiley & Sons, Limited

Roth, R. E., Ross, K. S., Finch, B. G., Luo, W., & MacEachren, A. M. (2010). A user-centered approach for designing and developing spatiotemporal crime analysis tools, in *Sixth International conference on Geographic Information Science*

Sabri, A., Hela, L., & Mounir, B. A. (2015). Towards an intelligent evaluation method of medical data visualizations, *2015 15th International Conference on Intelligent Systems DeSign and Applications* (ISDA), pp. 673 - 678

Sarker, Farhana, Davis, Hugh, Tiropanis, & Thanasis (2010). The Role of Institutional Repositories in addressing Higher Education Challenges, *The Second International Workshop on Semantic Web Application in Higher Education,* University of Southampton, UK

SAS (2011). *Data Visualization Techniques: From Basics to Big Data with SAS® Visual Analytics*, SAS (The Power To Know), White Paper.

Schaefer, M., Wanner, F., Kahl, R., Zhang, L., Schreck, T. & Keim, D.A. (2011). A Novel Explorative Visualization Tool for Financial Time Series Data Analysis, *Proceedings of VAW,* University College London, UK

Schmidt, S., Sch¨ogl, O., Kirchberger, R., Doleisch, H., Muigg, P., Hauser, H., Grabner, M., Bornik, A. & Schmalstieg, D. (2005). Novel visualization and interaction techniques for gaining insight into fluid dynamics in internal combustion engines, *In Proc. of the NAFEMS World Congress Conference* 2005

Sedig, K., Parson, P. & Babanski, A. (2012). Towards a Charaterization of Interactivity in Visual Analytics. *Journal of Multimedia Processing and Technologies, Special Issue on Theory and Application of Visual Analytics, 3* 1), 12-28.

Sedlmair, M., Meyer, M., & Munzner, T. (2012). Design Study Methodology: Reflections from the Trenches and the Stacks, *IEEE TVCG, 18*, 2431-2440

Semiu A. & Zulikha J. (2015a). Measuring InfoVis' Decision Support Effectiveness: From Theory to Practice. *Proceeding of Science and Information Conference,* pp. 560-564. July, 28 - 30, 2015, London, United Kingdom

Semiu A., & Zulikha, J. (2015b). A Theoretical Approach towards Designing InfoVis for Decision Support Effectiveness. *Special Issue of Jurnal Teknologi (In print)*

Semiu, A. & Zulikha J. (2014). Towards Developing An Instrument In Measuring The Need for InfoVis, i*n Proceeding of Knowledge Management International Conference*, Langkawi, pp. 628-631, Malaysia- KMICe 2014, Aug., 12-15, 2014

Semiu, A. & Zulikha, J. (2013). Information Visualization Techniques Usage Model, *in Proceeding of 4th International Conference on Computing and Informatics*, Sarawak, Malaysia- ICOCI, 2013, pp.247 - 252, Aug., 28-30, 2013.

Semiu, A. & Zulikha, J. (2015c). A Conceptual Framework for Students' Data-Driven Decision Making Process in Higher Education Institutions. *Advanced Science Letters, 21* (7): 2256-2260. DOI:10.1166/asl.2015.6275, American Scientific Publishers.

Semiu, A., & Zulikha, J. (2015d). Structural Components of Students’ Data-focussed Information Visualization, in *Proceeding of 5th International Conference on Computing and Informatics*, Istanbul, Turkey- ICOCI, 2015, pp. 665 – 670. Aug. 22-24, 2015.

Semiu, A., & Zulikha, J. (2016). Students' Data-driven Decision Making in HEI: The Explicit Knowledge Involved. *International Journal of Information and Education Technology*, *6* (1), 71 - 75.

Sheneiderman, B. & Plaisant, C. (2010). *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, 5th Edition, U.S.A. Pearson Higher Education.

Sherwood, C., & Rout, T. (1998). A structured methodology for multimedia product and systems development, *in Proceeding of the ASCILITE98 Conference*

Shneiderman, B. (1996). The eyes have it: a task by data type taxonomy for information visualizations, *IEEE Symposium on Visual Languages, 1996,* Boulder, CO, USA, pp. 336-343.

Simon, S., Oelke, D., Landstorfer, R., Neuhaus, K. & Keim, D. (2011). Visual Analysis of Next-Generation Sequencing Data to Detect Overlapping Genes in Bacterial Genomes, *Proceedings of IEEE Symposium on Biological Data Visualization,* October 23-24, Providence Island, USA, page 47- 54

Soukup, T., & Davison, I. (2002). *Visual Data Mining: Techniques and Tools for Data Visualization and Mining*, United States of America, John Wiley & Sons, Inc.

Spence, R. (2007). *Information Visualization: Design for Interaction,* 2nd Edition, England. Pearson Education Limited.

Stacey, M., Salvatore, J. & Jorgensen, A. (2013). *Visual Intelligence: Microsoft Tools and Techniques for Visualizing Data.* Indianapolis, John Wiley & Sons, Inc

Stasko, J. (2008). Visualization for Information Exploration and Analysis, *Proceedings of IEEE Symposium on Visual Languages and Human-Centric Computing* (VL/HCC ’08)

Stoffel, A., Spretke, D., Kinnemann, H. & Keim, D.A. (2010). Enhancing Document Structure Analysis using Visual Analytics, *Proceedings of the 2010 ACM Symposium on Applied Computing* (SAC’10), pp. 8-12, 2010

Stoffel, A., Zhang, L., Weber, S.H. & Keim, D.A. (2012). AMPLIO VQA- A Web-Based Visual Query Analysis System for Micro Grid Energy Mix Planning, *Proceedings of International Workshop on Visual Analytics, the Eurographics Association, 2012*

Strobelt, H., Oelke, D., Rohrdantz, C. & Stoffel, A. (2009). Document Cards: A Top Trumps Visualization for Documents. *IEEE Transaction on Visualization and Computational Graph, 15* (6), 1145 - 52. doi: 10.1109/TVCG.2009.139.

Syamsul, B.Z. (2011). *Mobile Game-Based Learning (mGBL) Engineering Model*, an Unpublished PhD Thesis submitted to Awang Had Salleh Graduate of Arts and Sciences, Universiti Utara Malaysia.

Tan, P-N, Steinbach, M. & Kumar, V. (2006). *Introduction to Data Mining*, 1st Edition, U.S.A. Pearson Education, Inc

Tatu, A., Albuqerque, G., Eisemann, M., & Bak, P. (2011). Automated Analytical Methods to Support Visual Exploration of High-Dimensional Data, *IEEE Transaction on Visualization and Computer Graphics*, *17* (5), 584- 596

Teegavarapu, S., & Summers, J.D. (2008). Case Study Method for Design Research, *Proceedings of International Design Engineering Technical Conferences & Computers and Information in Engineering Conference,* August 3-6, 2008, New York city, New York, USA, pp. 1-9

Teets, J.M., Tegarden, D.P., & Russell, R.S. (2010). Using Cognitive Fit Theory to Evaluate the Effectiveness of Information Visualizations: An Example Using Quality Assurance Data, *IEEE Transactions on Visualization and Computer Graphics, 16* (5), 841-853

Tekir, S., Mansmann, F. & Keim, D. (2011). Geodesic Distances for Web Document Clustering, *Proceedings of IEEE Symposium on Computational Intelligence and Data Mining (CIDM, 2011)*

Thacker, B.H., Doebling, S.W., Hemez, F., M., Anderson, M.C., Pepin, J. E., & Rodrigues, W., A. (2004). *Concepts of Model Verification and Validation*, South Research Institute & Los Alamos National Laboratory

Thomas, J.J., & Cook, K.A. (2005). *Illuminating the Path: The Research and Development Agenda for Visual Analytics*, Richland, WA: IEEE Press.

Thomsen, E. (2002). *OLAP Solutions: Building Multi-dimensional Information Systems.* John Wiley & Sons

Toker, D., Conati, C., Steichen, B., & Carenini, G. (2013). Individual User Characteristics and Information Visualization: Connecting the Dots through Eye Tracking, *In CHI 2013*, pp. 295-303

Tongco, M.D.C (2007). Purposive Sampling as a Tool for Informant Selection, *Ethnobotany Research & Applications*, *5*, 147-158

Tory, M. & Staub-French, S. (2008). Qualitative Analysis of Visualization: A Building Design Field Study, *ACM Proceedings of BELIV*, 2008, Florence, Italy, pp. 1-8.

Tory, M., & Moller, T. (2005). Evaluating Visualizations: Do Expert Reviews Work? Visualization Viewpoints, *IEEE Computer Graphics and Application,* September/October, 2005 Edition pp. 8-11

Triesman, A. (1985). Preattentive Processing in Vision. *Computer Vision, Graphics and Image Processing*, 31, 156- 177

Truckenbrod, J. R. (1981). Effective Use of Color in Computer Graphics, *Computer Graphics*, *15* (2), 83-90

Tullis, T., & Albert, B. (2008). *Measuring The User Experience: Collecting, Analyzing and Presenting Usability Metrics*. USA: Morgan Kaufmann Publishers

Turban, E. (1998). *Decision Support Systems and Intelligent System*. USA: Prentice Hall.

Turban, E. Sharda, R. & Delen, D. (2010). *Decision Support and Business Intelligent Systems,* 9th Edition. USA: Prentice Hall.

Vaishnavi, V., & Kuechler, W. (2007). *Design Science Research Methods and Patterns: Innovating Information and Communication Technology*. New York: Auerbach Publications.

Wagner, E. & Ice, P. (2012). Data Changes Everything: Delivering on the Promise of Learning Analytics in Higher Education, *EduCause Review,* July/August pp. 33- 42.

Wang, H. (2013). *Asynchronous Visualization of Spatiotemporal Information for Multiple Moving Targets,* an unpublished PhD Dissertation in School of Information Sciences, University of Pittsburgh.

Wang, H., Zhao, G., Wang, W. & Chen, C. (2012). The Design and Implementation of the Platform of Interactive Information Visualization on Aircraft Product Development, *Proceedings of 3rd International Conference on System Science, Engineering Design and Manufacturing Informatization*, pp. 180-184

Wanner, F., Rohrdantz, C., Mansmann, F., Stoffel, A., Oelke, D., Kristajic, M., Keim, D.A, Lou, D., Yang, J. & Atkinson, M. (2009). Large-Scale Comparative Sentiment Analysis of News Articles, *Proceedings of InfoVis Conference* '09.

Ward, M., Gristein, G., & Keim, D. (2010). *Interactive Visualization: Foundations, Techniques, and Applications.* India: A.K. Peters, Ltd

Ware, C. (2000). *Information Visualization: Perception for Design*. United States of America: Morgan Kaufmann Publishers

Web Center for Social Research Publication. Available at: https://www.syr.edu/gradschool/pdf/assessmentpdfs/Evaluation%20Defined.pdf. Retrieved on 19th January, 2015.

Wen-ting, W., Heng-yu, Z., Yi, R., Mi-mi, T., & Li-rong, X. (2009). Study on Plant Landscape Information Visualization and Its Key Technologies, *in Proceedings of International Conference on Management and Service Science,* pp. 1- 4

White, S. (2008). Interaction and Presentation Techniques for Situated Visualization, *UIST'08*, pp. 3-7, October 19-22, 2008

Wiegers, K. E. (2002). Seven Truths about Peer Reviews. *Cutter IT Journal*, 15(7).

Wilkinson, L. (2005). *The Grammar of Graphics*, 2nd ed. New York, NY, USA: Springer.

Willems, N., H. van de Wetering, J.J., & van Wijk. (2009). Visualization of Vessel Movements, *Proceedings of EuroVis*, *28* (3), 959-966.

Williams, N. & Amin, G. (2006). Higher Education in Sudan and Knowledge Management Application. *Information and Communication Technologies*, 2(1), 60-65

Wolfe, J. M. & Cave, K. R. (1989). Deploying Visual Attention: The Guided Search model, *In AI and the Eye, Troscianko, T. and Blake, A., Eds*., 79-103. John Wiley & Sons, Inc., Chichester, United Kingdom

Wright, H. (2007). *Introduction to Scientific Visualization*. United Kingdom. Springer Science and Business Media, LLC.

Wu, X. & Ren, F. (2010). Mechanism and Methods of Outdoor AR Spatial Information Visualization Representation, *in Proceedings 2010 Second International Conference on Computer Modelling and Simulation*, pp. 272-276, IEEE Computer Society

Xie, Y., Hofmann, H., & Cheng, X. (2014). Reactive Programming for Interactive Graphics. *Statistical Science, 29* (2), 201 – 213. DOI: 10.1214/14-STS477

Yanosky, R. (2009). *Institutional Data Management in Higher Education*. A Report from EduCause Centre for Applied Research. Retrieved on 6th June, 2013 from http://net.educause.edu/ir/library/pdf/EKF/EKF0908.pdf

Yi, J. S. (2010). Towards Practical Information Visualization Theory: Lessons from the History of Decision Science, *VisWeek 2010, The Role of Theory in Information Visualization.* ACM Workshop.

Yi, J.S., Kang, Y., Stasko, J.T., & Jacko, J.A. (2007). Toward a deeper understanding of the Role of Interaction in Information Visualization, *IEEE Transactions on Visualization and Computer Graphics*, *13* (6), 1124–1231.

Yu, C., Yurovsky, D. & Xu, T. (2012). Visual Data Mining: An Exploratory Approach to Analyzing Temporal Patterns of Eye Movement, *Infancy*, *17* (1), 33-60.

Zadrozny, P. & Kodali, R. (2013). *Big Data Analytics Using Splunk: Deriving Operational Intelligence from Social Media, Machine Data, Existing Data Warehouses and Other Real Time Streaming Sources*. Apress

Zalta, E. N., Nodelman, U., Allen, C. & Perry, J. (2013). The Problem of Induction, *Standard Encyclopedia of Philosophy*, *36* (7), 1-5

Zhao, D. & Zhou, N. (2008). Analysis of the Present Applications of Information Visualization in E-commerce Websites

Zheng, X-Z., Wang, S-L., Luo, J-H., Cao, Y., & Qing, D-Z. (2011). Research on the Framework of Information Visualization System of Equipment Support Decision-making. *Proceedings of 2011 International Conference on Computational and Information Science*, pp. 780- 784. IEEE Computer Society

Zhong, C., Wang, T., Cairns, S., & Arisona, S.M. (2012). Using an Interactive Dymaxion Map to Convey Research Information through Visualization, *Proceedings of 16th International Conference on Information Visualization,* IEEE Computer Society

Ziegler, H., Nietzschmann, T. & Keim, D. (2008). Visual Analytics on the Financial Market: Pixel-based Analysis and Comparison of Long-Term Investments, *Proceedings of the 12th Conference on Information Visualization (IV 2008), IV-VDM Symposium on Visualization and Data Mining,* pp. 287- 295

Ziemkiewicz, C., Ottley, A., Crouser, R.J., Chauncey, K., Su, S.L., & Chang, R. (2012). Understanding Visualization by Understanding Individual Users, IEEE *Journal of Computer Graphics and Applications, 32* (6), 88-94

Zikmund, W.G., Babin, B.J., Carr, C.J., & Griffin, M. (2010). *Business Research Methods,* 8th Edition, Canada. South-Western Cengage Learning Centre

Zikopoulous, P., Eaton, C., Deutsch, T., Deroos, D., & Lapis, G. (2012). *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data*. USA, IBM.

Zins, C. (2007). Conceptual Approaches for Defining Data, Information, and Knowledge. *Journal of the American Society for Information Science and Technology*, 58(4), 479–493