

# **Data Navigation and Visualization: Navigating Coordinated Multiple Views of Data**

**Igoniderigha, Nseabasi Ekaette.**

A thesis submitted in partial fulfilment of the requirements of Edinburgh  
Napier University for the degree of Doctor of Philosophy.

March 2015

## **Abstract**

The field of coordinated and multiple views (CMVs) has been for over a decade, a promising technique for enhancing data visualization, yet that promise remains unfulfilled. Current CMVs lack a platform for flexible execution of certain kinds of open-ended tasks consequently users' are unable to achieve novel objectives. Navigation of data, though an important aspect of interactive visualization, has not generated the level of attention it should from the human computer interaction community. A number of frameworks for and categorization of navigation techniques exist, but further detailed studies are required to highlight the range of benefits improved navigation can achieve in the use of interactive tools such as CMVs.

This thesis investigates the extent of support offered by CMVs to people navigating information spaces, in order to discover data, visualize these data and retrieve adequate information to achieve their goals. It also seeks to understand the basic principle of CMVs and how to apply its procedure to achieve successful navigation.

Three empirical studies structured around the user's goal as they navigate CMVs are presented here. The objective of the studies is to propose a simple, but strong, design procedure to support future development of CMVs. The approach involved a comparative analysis of qualitative and quantitative experiments comprising of categorised navigation tasks carried out, initially on existing CMVs and subsequently on CMVs which had been redesigned applying the proposed design procedure. The findings show that adequate information can be retrieved, with successful navigation and effective visualization achieved more easily and in less time, where metadata is provided alongside the relevant data within the CMVs to facilitate navigation. This dissertation thus proposes and evaluates a novel design procedure to aid development of more navigable CMVs.

## **Presentations and Publications Arising from the Thesis**

Benyon, D., McEwan, T., & Igoniderigha, N. (2014) Designing Coordinated Multiple Views of Information Space. Proceedings of the British HCI conference, September 2014.

# Table of Content

<b>ABSTRACT .....</b>	<b>I</b>
<b>TABLE OF CONTENT .....</b>	<b>III</b>
<b>LIST OF BOX.....</b>	<b>VIII</b>
<b>LIST OF FIGURES.....</b>	<b>IX</b>
<b>LIST OF TABLES .....</b>	<b>XII</b>
<b>ACKNOWLEDGEMENT .....</b>	<b>XIII</b>
<b>CHAPTER 1 INTRODUCTION .....</b>	<b>1</b>
1.1 Research background.....	1
1.2 Motivation.....	3
1.3 Problem Statement .....	5
1.4 Research Questions .....	6
1.5 Hypothesis .....	7
1.6 Thesis Overview .....	7
<b>CHAPTER 2 BACKGROUND.....</b>	<b>9</b>
2.1 Introduction .....	9
2.2 Introduction to Navigation .....	9
2.3 Information Space.....	10
2.3.1 <i>Data</i> .....	12
2.3.2 <i>Metadata</i> .....	14
2.3.3 <i>Data models</i> .....	18
2.3.4 <i>Database Models</i> .....	19
2.3.5 <i>Entity Relationships</i> .....	24
2.3.6 <i>Metadata and E-R Model relationship</i> .....	27
2.4. Models of Navigation .....	29
2.4.1 <i>Navigation</i> .....	29
2.4.2 <i>The Navigation process</i> .....	36
2.4.3 <i>Navigation Activities</i> .....	39

2.4.4 Navigating interactive visual Information Spaces.....	42
2.4.5 Navigation Challenges on Information Spaces.....	45
2.4.6 Information foraging .....	49
2.5 Visualization.....	53
2.5.1 The visualization process.....	54
2.5.2 Types of Visualization .....	54
2.5.3. Data Analytics .....	62
2.6 Coordinate multiple views (CMVs).....	64
2.6.1 Visual Analytics .....	65
2.6.2 Visualization Techniques.....	68
2.6.3 Visualization, Navigation and Metadata.....	70
2.7 Conclusion .....	72
<b>CHAPTER 3 COORDINATE MULTIPLE VIEWS OF VISUALIZATION .....</b>	<b>73</b>
3.1 Introduction .....	73
3.2 Coordinate Multiple Views .....	73
3.2.1 Designing CMVs.....	77
3.2.2 CMV Visualization Techniques .....	80
3.3 The IIDI Data: A CMV of a Research Centre’s Database .....	84
3.3.1 The IIDI Data Visualization.....	85
3.3.2 Interface Events on the IIDI Data CMV.....	85
3.3.3 The IIDI User Interface A.....	87
3.4 Using the IIDI Data CMV .....	96
3.4.1 Task Design .....	97
3.4.2 Task Categorization .....	99
3.4.3 IIDI Data as an Information space.....	103
3.4.4 Metadata as Many-to-one (M:1) Relationship.....	105
3.4.5. Navigating the IIDI Information Space .....	107
3.5 A Redesigned CMV for the IIDI Database.....	109
3.5.1 Interface B.....	111
3.5.2 Interface C.....	114

<b>CHAPTER 4 EMPIRICAL STUDY ONE .....</b>	<b>121</b>
4.1 Introduction .....	121
4.2 Research Methodology .....	122
4.2.1 <i>Experimental Investigations</i> .....	122
4.2.2 <i>Descriptive Investigation</i> .....	123
4.3 The IIDI Data Visualization Interface .....	124
4.3.1 <i>The Experiment Design</i> .....	126
4.3.2 <i>Summary Data</i> .....	127
4.3.3. <i>Participants</i> .....	128
4.3.4 <i>Experiment Setup</i> .....	128
4.3.5. <i>Between Subjects Experiment Design</i> .....	131
4.3.6 <i>The Log File</i> .....	133
4.4 Quantitative Results.....	134
4.4.1 <i>Type of tasks completed</i> .....	135
4.4.2 <i>Task completion times</i> .....	141
4.4.3 <i>Navigation Pattern Analysis</i> .....	142
4.4.4 <i>Clicks Analysis</i> .....	143
4.4.5 <i>Discussion</i> .....	145
4.5 Qualitative results .....	147
4.5.1 <i>Informal Interview</i> .....	147
4.5.2 <i>Discussion</i> .....	148
4.6 Conclusion .....	151
<b>CHAPTER 5 EMPIRICAL STUDY TWO.....</b>	<b>152</b>
5.1 Introduction .....	152
5.2 Research Methodology.....	152
5.2.1 <i>Experimental Investigations</i> .....	153
5.2.2 <i>The Variables</i> .....	153
5.2.3 <i>Descriptive Investigations</i> .....	153
5.3 The IIDI Data CMV .....	154
5.3.1 <i>The IIDI Data Interfaces</i> .....	154

5.3.2	<i>The Experiment Design</i> .....	157
5.3.3	<i>Participants</i> .....	158
5.3.4	<i>Experiment Setup</i> .....	158
5.3.5	<i>The Interface key</i> .....	160
5.3.6	<i>Within Subjects Design</i> .....	161
5.3.7	<i>Task development</i> .....	162
5.3.8	<i>Tasks Categorization</i> .....	163
5.4	<b>Quantitative Results</b> .....	166
5.4.1	<i>Task Analysis of individual Participants</i> .....	166
5.4.2	<i>Task Category Analysis of Individual Participants</i> .....	167
5.4.3	<i>Analysis of Task Completion by Interface</i> .....	169
5.4.4	<i>Analysis of Task completion Times</i> .....	170
5.5	<b>Qualitative Results</b> .....	176
5.5.2	<i>Interface Analysis</i> .....	177
5.6	<b>Statistical significance</b> .....	179
5.7	<b>Conclusion:</b> .....	184
<b>CHAPTER 6 EMPIRICAL STUDY THREE</b> .....		<b>185</b>
6.1	<b>Introduction</b> .....	185
6.2	<b>User Interface</b> .....	186
6.2.1	<i>Design Guidelines</i> .....	187
6.2.2	<i>Proposed Guidelines for Designing CMVs</i> .....	189
6.3	<b>Evaluating the IIDI Data Interface C</b> .....	193
6.3.1	<i>The value of summary data and Interface key</i> .....	193
6.3.2	<i>Method</i> .....	194
6.3.4	<i>Participants</i> .....	195
6.3.5	<i>Experiment Setup</i> .....	196
6.3.6	<i>Procedure</i> .....	197
6.4	<b>Results</b> .....	199
6.4.1	<i>Success rate</i> .....	199
6.4.2	<i>Task completion times</i> .....	200

6.4.3 User Satisfaction .....	201
6.4.4 Utility.....	202
6.5 Discussion.....	203
6.6 Conclusion .....	204
<b>CHAPTER 7 VISUALIZING THE SIMD DATA: A CASE STUDY .....</b>	<b>205</b>
7.1 Introduction .....	205
7.2 The Scottish Index of Multiple Deprivation (SIMD) Data .....	205
7.3 SIMD Data Visualization as a CMV .....	207
7.3.1 Focus Group.....	208
7.3.2 User Stories.....	208
7.3.3 Personas .....	209
7.4 Chapter Discussion.....	217
7.5 Conclusion: .....	219
<b>CHAPTER 8 DISCUSSION.....</b>	<b>220</b>
8.1 Introduction .....	220
8.2 Information Spaces and Navigation.....	220
8.3 Metadata .....	223
8.4 Evaluating Interface C .....	227
8.5 Conclusion .....	232
<b>CHAPTER 9 CONCLUSION.....</b>	<b>234</b>
9.1 Introduction .....	234
9.2 Main Contributions.....	235
9.3 Future Work .....	238
9.4 Conclusion .....	239
<b>REFERENCES.....</b>	<b>240</b>
<b>APPENDICES .....</b>	<b>270</b>



## List of Box

Box 2. 1 Extract from Pirolli and Card .....	51
--	----

## List of Figures

Figure 2. 1: Diagram of a relation .....	19
Figure 2. 2: Diagram of a database created in terms of a relational model .....	20
Figure 2. 3: Diagram illustrating two related entities.....	22
Figure 2. 4: Diagram illustrating the entity employee with an attribute SSN .....	22
Figure 2. 5: Diagram illustrating a relation 'proved' with an attribute 'date'.....	22
Figure 2. 6: Diagram illustrating the use of metadata in an E-R diagram .....	24
Figure 2. 7: Diagram illustrating one-to-one relationship (1:1).....	25
Figure 2. 8: Diagram illustrating One- to- Many relationships.....	25
Figure 2. 9: Diagram illustrating Many-to-Many relationship .....	26
Figure 2. 10: Darken's model of navigation .....	32
Figure 2. 11: Spence's initial proposal of a model for a framework for navigation....	32
Figure 2. 12: Modification to Spence's original proposal .....	33
Figure 2. 13: Proposed framework for navigation.....	34
Figure 2. 14: A schematic model of Navigation .....	35
Figure 2. 15: Transformation of inherent data structure into imposed structure .....	38
Figure 2. 16: Visualization process.....	54
Figure 2. 17: Florence Nightingale's roses showing an outbreak of cholera .....	59
Figure 2. 18: Biography chart illustrating main world civilizations.....	60
Figure 2. 19: Diagram showing a scatterplot .....	61
Figure 2. 20: Diagram of visual data exploration .....	66
Figure 3. 1: IIDI Data CMV Interface A.....	88
Figure 3. 2: Upper left view of IIDI CMV Interface .....	89
Figure 3. 3: Upper left View of IIDI Interface showing a mouseover nodelink event.	90
Figure 3. 4: Upper left view of IIDI interface showing selected nodelink event.....	90

Figure 3. 5: Lower left view of IIDI CMV Interface .....	91
Figure 3. 6: Upper right view of IIDI CMV Interface A .....	92
Figure 3. 7: Diagram of Bar chart on URV of IIDI CMV Interface A.....	93
Figure 3. 8: Upper Right View of IIDI CMV Interface A, showing coding scheme ....	94
Figure 3. 9: Lower right view of the IIDI CMV interface .....	94
Figure 3. 10: Entity relationship diagram illustrating many-to-one relationships.....	105
Figure 3. 11: IIDI Data CMV Interface B.....	111
Figure 3. 12: Upper centre view of interface B .....	114
Figure 3. 13: IIDI Data CMV Interface C.....	114
Figure 3. 14: Upper right view of interface C .....	115
Figure 3. 15: URV of IIDI Data Interface C illustrating selected nodelink event.....	116
Figure 3. 16: URV of IIDI Data Interface C illustrating mouseover nodelink event .	116
Figure 3. 17: UCV of IIDI CMV Interface C illustrating the labels .....	117
Figure 3. 18: Mouseover matrix event on UCV cross centre publications window.	117
Figure 3. 19: Mouseover matrix on UCV cross –centre grants window. ....	118
Figure 3. 20: Mouseover matrix event on UCV centre by publication window.....	119
Figure 3. 21: Diagram of UCV showing centre by grants window on Interface C ...	119
Figure 4. 1: IIDI Data CMV Interface A.....	125
Figure 4. 2: IIDI Data CMV Interface B:.....	126
Figure 4. 3: Log file extract for Interface A, tutorial 1.....	134
Figure 4. 4: Navigation Pattern illustration for Interface A, Tutorial 2 .....	143
Figure 5. 1: IIDI Data CMV Interface A.....	155
Figure 5. 2: IIDI Data CMV Interface B.....	156
Figure 5. 3: IIDI Data CMV Interface C.....	156
Figure 5. 4: Interview results for 12 respondents for Interface A .....	178

Figure 5. 5: Interview results for 12 respondents, Interface C .....	179
Figure 6. 1: IIDI Data CMV Interface C:.....	195
Figure 7. 1: Diagram illustrating levels of deprivation in Edinburgh .....	207
Figure 7. 2: Map showing data zone S010001929 Craiglockhart, Edinburgh.....	211
Figure 7. 3: A CMV for SIMD data.....	213
Figure 7. 4: Holy Rood RC secondary school catchment area showing feeder primaries .....	215
Figure A: Map of most deprived areas in Scotland in 2012 .....	282
Figure B: Map of most deprived areas in Scotland 2012.....	283
Figure C: Map of most deprived areas in Scotland showing metadata in 2012.....	283

## List of Tables

Table 3. 1: Tasks categorisation .....	103
Table 4. 1: Task type completed by tutorial 1 on Interface A, (n=18) .....	135
Table 4. 2: Task type completed by tutorial 2 on Interface A, (n=37) .....	136
Table 4. 3: Task type completed by tutorial 3 on Interface B, (n=17) .....	136
Table 4. 4: Task type completed by tutorial 4 on Interface B, (n=28) .....	137
Table 4. 5: Comparative analysis of type of tasks completed by participants.....	138
Table 4. 6: Comparative analysis of task completion times by tutorials .....	141
Table 4. 7: Cross section of number of clicks by event made by tutorial groups ....	145
Table 4. 8: Broad categorization of responses from open-ended questions.....	150
Table 5. 1: Tasks categorization for empirical work 2.....	165
Table 5. 2: Cross section of uncompleted tasks on interfaces A and C.....	167
Table 5. 3: Cross section of task category completed by individual participants ....	168
Table 5. 4: Comparative analysis of task scores and task completion times of participants on interfaces A and C .....	170
Table 5. 5: Summary of results of participants (A – C order) on Interface A .....	171
Table 5. 6: Summary of results of participants on Interface C (A-C order) .....	172
Table 5. 7: Summary of results of participants (C – A order) on Interface A .....	173
Table 5. 8: Summary results of participants (C – A order) on Interface C .....	174
Table 5. 9: Comparative analysis of navigation steps on interfaces A and C .....	175
Table 5. 10: Table illustrating computation for Wilcoxon signed rank test.....	182
Table 5. 11: Critical table values for Wilcoxon signed rank test .....	183
Table 6. 1: Success rates of users on interface C .....	200
Table 6. 2: Task completion times (seconds) of users on Interface C .....	201

## **Acknowledgement**

I am grateful for the valuable support and continual encouragement I enjoyed from my director of studies: Prof David Benyon and my second supervisor: Tom McEwan throughout the duration of this research. The assistance I received from Martin Graham, as well as the useful advice from my panel chair, William Buchanan is treasured.

I am indebted to all the people who took part in the empirical studies carried out for the period of this research. I appreciate the computing department of Edinburgh Napier University for the support I received through the course of this research.

A big thank you to my friends who have supported me through the duration of this research: Pastor Edmund, Abekhe; Pastor Ethel, Abekhe; Dr. Tamuno, PBL and Dr. Tuaweri.

Special thanks to my parents, late Chief Abel N. Ettete and Mrs. E.E Ettete for their immense support, to my aunt Princess, and Nana, for their support and to my siblings, nieces and nephews, for being a part of my world.

The support, encouragement, sacrifice and understanding of my husband, Dengi, my children, Douye, Preye, Woye and Ifiye, through the course of this research has been extraordinary. You are special blessings to me from God. Thank you for adding color to my life.

Above all, I give all thanks and glory to The Lord God Almighty for his special grace, his mercies and for being my anchor.

# Chapter 1 Introduction

## ***1.1 Research background***

The design of coordinated multiple views (CMVs) of data is one area of human computer interaction that receives a lot of attention although from different perspectives. They cover a vast range of interactions from technical visualizations to 'dash boards' showing analysis of data such as Google Analytics. CMVs have a vast range of possibilities such as enabling visualizations that are easily managed to be created (Baldonado & Kutchinsky 2000; North & Shneiderman, 2000). However, its usage brings on various challenges, such as developing easy interaction methods for movement (Meiguins, Meiguins, Leandro, Augusto and Sergio, 2010). Shneiderman and his co-authors (2009) claims that powerful visualizations facilitate understanding and interpretation of data supports data discovery and information retrieval, however this research is of the opinion that these benefits may not be maximised if people are unable to navigate the data set within the information space.

In designing CMVs, developers are simply advised to “define the environment around their tasks”, which might involve writing SQL queries and selecting interface widgets to present the results in a given view. But with the coming on of “big data” on the one hand, and open ended exploration of that data on the other, and different people holding separate views on data, clearly the task can no longer be well defined. Furthermore, when a record in a relational database is equated with an object in the data set, it makes it difficult to distinguish data from metadata.

Good database design, using well established techniques, ensures that the tables of data are organised in such a way that the known relationships are easily and efficiently



accessed and modelled. What we are interested in, however, is how to design CMVs to support visual analytics particularly in the area of management information systems (MIS) whose intent could be to manage a research institute or a hospital support service. Rather than utilising conventional database methods which retrieve information through the use of query languages, analytical reasoning moves forward to discover patterns, find new relationships between objects and navigate the database in novel ways. The research seeks to establish guidelines for the novel design of CMVs to cover new patterns of use that include instances such as; when unknown relationships emerge, or when data about the data (metadata) reveal previously undetermined linkages, and answer questions that reflect unanticipated scenarios. Furthermore, the design could help users model the rate of change and as a result predict future relationships.

The concept of navigation of information space has been developed over several years (Benyon, 1998; Benyon, 2001; Benyon, 2006; Benyon, 2007). The concept compares the way people navigate through information spaces to the way they navigate physical spaces; they move through the physical or virtual world exploring data, finding their way in order to arrive at a required destination by means of some support. The role of navigation can be demonstrated in the need for people who according to Benyon, Hook & Nigay, (2010); Benyon & Hook, (1997) exist in their information spaces, and need to move through successfully to carry out various information seeking activities.

In this research after the reviews of data analytics, and of the notions of navigation of information space, we introduce a case study of a management information system

(MIS) for a research institute to show how people are able to navigate the information space more effectively if they have access to metadata, through a CMV. From this study an initial set of design guidelines is derived, developed and evaluated and put forward as an addition to existing guidelines for the design of CMVs.

## **1.2 Motivation**

Coordinate and Multiple Views is a technology used in various domains like visual analytics. Coordination between views makes it easier to execute tasks that involve information retrieval and visualizing complex data sets. Nevertheless, users of interactive information spaces like CMVs are confronted with navigational problems that hinder their ability to explore data from multiple perspectives or achieve effective information visualization and retrieval. A number of these CMVs are designed to support user tasks, nonetheless, the problem is made more complex by the absence of sufficient techniques to aid successful navigation in order to visualise and retrieve adequate information. Dieberger (1997) points out that a significant setback for users of present information systems is the retrieval of new information from information systems. He notes that closely interlinked systems including the World Wide Web (WWW) do not communicate structures that support users to navigate to the information it holds.

A number of studies concerning visualization and information retrieval have been undertaken but not much work has been carried out on navigation. For instance, Roberts (2007) describes a state of the Art CMV but does not discuss the issue of navigation. Coordination helps users understand the relationship between views (Arora, 2013) and aids execution of tasks such as visualizing data sets. Coordination

is a significant aspect of interaction on CMVs and it involves navigation. Navigation described by Benyon (2006) as moving through an environment, plays a central role in coordinating multiple views, which involves various movements between views. Furthermore, it is vital to effective interaction between data sets for purposes of understanding and interpreting data in order to visualise and retrieve information.

Currently, users find it challenging to find their way to locations of interest and then find their way back to locations they visited earlier in an appropriate manner. Little support is provided in virtual worlds to enable people achieve successful wayfinding (Darken and Sibert, 1996b). Many users are unable to tell when they reach their desired destination (Elvins, 1997; Benyon, 2001). Choosing the right route causes a setback for many users or discovering patterns and relationships. A lot of users end up aggregating data themselves, a laborious strategy that yields insufficient information. Others become disoriented which means they lose their sense of direction, position or relationship with their surroundings and get lost in hyperspace. Benyon and Hook (1997) propose that rather than designing systems which support existing human tasks, the upcoming age will require that designers develop networks of interacting systems that support domain-oriented activities like navigation. He calls for a close examination of the 'big picture' in HCI.

Developing specific designs guidelines to create navigable CMVs would advance its primary objective which is to bring about effective data visualization and adequate information retrieval.

### **1.3 Problem Statement**

This thesis addresses significant challenges faced by people navigating existing CMVs. CMVs focus specifically on providing a platform for exploration but do not sufficiently integrate navigation which is a significant aspect of exploration. Existing CMVs focus more on user tasks however the overview provided to facilitate user's understanding of data is inadequate. Consequently, navigation is hindered and users are unable to perform the entire task they require.

In Chen's (2005) article on the top 10 unsolved Information visualization problems, he stipulates that there is an obvious need for the development of new evaluative measures that address challenges and requirements specific to data visualization.

McCormac, Parsons and Butavicius (2007) note that in spite of the development of research and data visualizations, not enough usability studies and empirical evaluations have been undertaken.

In addition, the members of the CHI '98 workshop believe that the design procedure for multiple view systems can be developed by usability heuristics (Baldonado, Michelle Q., Woodruff and Kutchinsky (2000).

Shneiderman (1987, 1996) and Nielsen (1994) have put forward guidelines for designing general user interfaces which are valuable. Nevertheless, designers of coordinate multiple views have insufficient guidance available to them. The designers find it challenging to choose the principles mostly relevant to CMVs from the various general guidelines available (Baldonado, Michelle Q., Woodruff and Kuchinsky, 2000).

Furthermore, CMVs are typically designed with two or more coordinate views to 'support the exploration of a single conceptual entity' (Baldonado, Woodruff and Kutchinsky, 2000). It presents data from different perspectives with the aim of enabling users discover novel relationships, interact with and understand the data presented better (North and Shneiderman, 2000). However, the theory behind this coordination is unclear and the issue of navigation is poorly addressed.

In spite of the extent of work carried out in the area of CMVs, there is little discussion about the type of data to be presented in the visualization. The visualization designer is expected to discern the tasks users will want to execute and tailor the CMV to support those tasks (Baldonado, Woodruff and Kutchinsky, 2000).

Navigation of information spaces rather than being task oriented encompass a general range of activities such as object identification, wayfinding and exploration in addition to tasks. Incorporating these activities along with tasks will make it possible to move away from task based CMVs to more general CMVs bridging the gap between user's navigation needs and the limitations of the CMVs.

The research questions outlined below have been derived from the reflections enumerated above.

#### ***1.4 Research Questions***

- Will integrating the concept of navigation of information space in developing novel CMVs bring about design adjustments that will enable the CMVs support users achieve successful navigation?

- Will providing metadata for content (summary data) and metadata for visual aspects (interface key) of an information space improve user's effectiveness to navigate, visualise and retrieve information?
- Will the redesigned CMV reduce amount of reasoning people have to make in order to navigate and bring about faster task conclusion times?
- To what extent will the novel set of design guidelines proposed by this research produce visible improvements on user's ability to navigate CMVs when applied?

### ***1.5 Hypothesis***

The research hypothesis has been developed to investigate appropriate methods to evaluate the research hypothesis:

- The research hypothesis states that when data is provided along with its summary data on CMVs, people can achieve successful navigation, visualize data effectively and retrieve adequate information.
- The null hypothesis states that the providing data along with its summary data would have no effect on successful navigation of CMVs, data visualization or information retrieval.

### ***1.6 Thesis Overview***

The thesis is structured as follows:

- Chapter 1: Chapter one provides a brief introduction of the problems addressed by the research. It explains the research motivation, research hypothesis and research questions.
- Chapter 2: This chapter presents the intellectual perspective for the research through a literature review. It defines related terms used and discusses related work.

- Chapter 3: Chapter three describes the model of the Institute of Informatics and Digital Innovation (IIDI) data's CMV. It touches on the IIDI data's entity relational model and the outcome of modifying the pre-existing one-to-one relationship to a many-to-one relationship to permit the provision of a summary data of metadata.
- Chapter 4: This chapter describes an initial laboratory based experiment which investigates a preliminary solution to the challenges of navigating a CMV.
- Chapter 5: Chapter five assesses in more detail through a user study the outcome proposed changes made on the IIDI CMV have on navigation.
- Chapter 6: Chapter six evaluates the design principles proposed by the research to determine its effectiveness in supporting the development of navigable CMV's.
- Chapter 7: Chapter seven gives a brief account of a case study carried out in a different domain to illustrate the practical benefit metadata provides to support people navigate CMVs.
- Chapter 8: Chapter Eight discusses the results obtained from the empirical studies in detail in relation to navigation, visualization and Information retrieval.
- Chapter 9: This chapter sums up the contributions this thesis makes to the future development of navigable CMVs in order to support successful navigation and outlines areas for further work.

## **Chapter 2 Background**

### ***2.1 Introduction***

This chapter provides a broad view of the thesis and forms the foundation for the Empirical studies carried out in this research. It discusses the concept of navigation of information spaces to give an overview of navigation and briefly introduces other relevant concepts to give a suitable background for the research. Next, the information space and how it is developed through data models are discussed. Navigation challenges encountered on information spaces are identified and discussed in detail. The visualization technique is discussed and the focus of this research is determined.

### ***2.2 Introduction to Navigation***

The challenge of how users can navigate through large information spaces successfully is central to achieving effective visualization and adequate information retrieval. Studies related to navigation have been undertaken in areas such as how spatial ability influences navigation (Ahmed and Blunstein, 2005), wayfinding (Arthur and Passini, 1992), designing navigable information spaces (Mark Foltz, 1998) and supporting the individual navigate information spaces with the aid of metaphors, 3D interfaces and adaptive interfaces (Benyon and Hook, 1997).

Research demonstrates that different approaches have been taken to tackle the difficulties encountered when people need to navigate information spaces. Some information spaces have provided keys, signs, maps and landmarks to help ease these challenges. Foltz (1998) points out that the regular way of organizing information spaces are not fully effective for users with the intention to navigate.



An alternative approach is the 'navigation of information space' view of humans and computers in the twenty-first century, which provides a new perspective to design for people who live in their information spaces (Benyon, 2001). One aspect of this design involves designing to allow people explore knowledge on virtual spaces in a way similar to how they navigate their physical spaces (Foltz, 1998).

### ***2.3 Information Space***

Information spaces are described by Benyon (2001) as spaces built to make available certain data and functions and to allow people supervise their actions. These spaces could be virtual or physical and they make it possible for people to perform activities such as search and retrieval. Benyon (2001) claims that people live and exist in information spaces. Ahmed and Blunstein (2005) agree with Benyon (2001) and explain that an information space is an organized gathering of information made up of different information artefacts and symbols from which people seek information for their activity space.

An overview of the theories of activity space and information artefacts will be given to explain information spaces further. Benyon (2001) defines activity space as a place where people take physical action and undergo physical experiences. He explains that people need access to information to enable them carry out activities in the activity space and that the activity space influences the information space by supporting its skilful development. In Benyon's words, 'the information space is intrinsically linked with the activity space'.

Information spaces offer these support through artefacts that could be interactive or non-interactive objects used to store, retrieve and in some cases, alter information and support the general interpretation of information spaces, (Green and Benyon, 1996). They can be described in two stages; a conceptual stage which presents an idea of the experienced world and a perceptual stage which gives a view of the structure in question. For instance, a library paper timetable which outlines its open hours is an information artefact which gives information about library open hours.

To design information spaces, information system designers imagine an aspect of an activity space and then make-up some signs, keys, symbols or artefacts which are used to convey the specific information the space represents to people. Information spaces store data in digital forms and people access these information spaces for purposes of browsing, searching and selecting assets of interest. Information artefacts support people through the use of signs to retrieve required information (Benyon and Hook, 1997).

In general these information signs, symbols, keys and artefacts define an information space. They define the objects, the configuration and the functions which permit the storage, retrieval and alteration of information (Benyon 2001).

One essential activity carried out in the information space is the movement from one point to another, also recognized as navigation. Information spaces hold a vast amount of data and are sometimes seen as a set of nodes, each representing a piece of information (Van Dyke Parunak, 1989). Virtual movement between nodes can be

likened to physical movement from one point to another. Sometimes the information space takes the form of a city offering numerous routes to reach a single destination.

Information spaces are diverse and include three dimensional cyberspace, which offers unrestricted movement, (Walker, 1990), the World Wide Web (WWW), a large and complex hypermedia system and paper documents such as books. In general, information spaces can be described as spaces created to allow people carry out various activities like searching, moving, living, working and resting. All miscellaneous interactions carried out by people in the real world can be described as an outcome of the discovery, substitution, organisation and manipulation of information. The zone within a computer is insufficient to define the breadth of an information space; it is the core of the daily occurrence of humans. Discovering one's way through a city centre or a holiday resort is a typical example of moving through the activity space, with the support of an information space (Benyon 2003).

This research investigates management information systems (MIS), a particular class of CMVs from the navigation perspective. It explores the navigation challenges encountered by people seeking to retrieve information from these CMVs in order to come up with additional recommendations for designing navigable CMVs.

### **2.3.1 Data**

Data are individual parts of information placed on information spaces in an organized way to communicate information. Different types of data are utilized to represent information, and the type of data used is considered carefully, they include graphs, nodes, matrix, tiles and words.

Data is represented through models known as data models and various information spaces including CMVs are created from data models. Data makes information available and may be collected in qualitative and quantitative forms. Both types of data are used to obtain information during empirical studies. For purposes of this research, both types of data will be utilized.

Qualitative data is descriptive in nature and is usually used to make comparisons given that it explains events, activities and people. It is gathered by means of nominal data, when ordering of the data is unimportant, such as telephone numbers, and in ordinal forms when the ordering of data is important, such as ages of individuals. On the other hand, quantitative data defines quantity and can be measured in terms of performance such as success rates and task completion times. It is collected numerically, and is referred to as interval data when the data does not need to be divided to make it meaningful such as a bank balance, however when division is required to make quantitative data meaningful, like length of time people take to retrieve information, it is referred to as ratio (Mitra, 2011). When both qualitative data and quantitative data are collected for analysis in a particular empirical study in the HCI field, in order to comprehend a research problem, it is termed a mixed method (Creswell, 2011). Qualitative data is gathered by means of observations, interview notes, focus groups and survey responses, while quantitative data is obtained by means of instruments, which include factual measures such as log file records and performance measures such as tasks. These data are then studied and analysed to ensure validity and reliability and then interpreted (Lazar, Feng & Hochheiser, 2010). The intent of the qualitative data collected in this research is to provide information about the behaviour

of people (Yin, 2003) as they navigate specific CMVs and the level of support the information space offers to enable them to navigate through successfully.

### **2.3.2 Metadata**

The term “meta” is used as a prefix in Information technology to explain or describe digital data using metadata standards. Cathro (1997) suggests that “an element of metadata describes an information resource, or helps provide access to an information resource”. He claims that a group of metadata parts could explain a single or multiple information resource. Tweedie, (1997) says metadata is descriptive information about data. Metadata is organised information that supports the uncovering, illustration and clarification of information resources. It enables people find and manage these information resources. It explains digital data by giving details about how, when, where and who collected a specific set of data, and goes on to give details about how the data was formatted.

Metadata is connected to the information resource it describes, for example a library catalogue record is formed by a group of metadata elements connected to the book through the card number Solodovnik, (2011). Metadata is used in data warehouses to describe the elements of the data warehouse and how they work together (Kimball 2008). It describes textual and non-textual objects, digital assets, academic books and materials. One of the very early references made of metadata was by Philip Bagley (1968) where he talked about structural metadata. After this many disciplines like Information Science, Information technology and Information Management accepted metadata to mean data about data.

For the purposes of this research, the term Metadata is described as the key to storing information resources in a manner which makes certain they can be accessed quickly and easily in future. It has the potential to shape the way people navigate through data in order to visualise it and retrieve adequate information easily and quickly. It is the key that will make sure that stored information will continue to exist and be available in future.

Metadata can describe the objects used in an information space to pass on information like in which direction user should navigate to retrieve what. According to Rada and Anthonio (1997) metadata can improve the interaction between people and digital information.

Metadata is a tool that can bring about accuracy during information navigation, visualisation and retrieval by identifying the major idea of the information resource like the author, centre, topic, title, publisher and date (Warwick, 1997).

A study on *metadata-enhanced visual interfaces* was carried out in School of Library and Information Studies, University of Alberta. In course of the study a particular group of digital library visual interfaces that support information seeking, exploration and retrieval based on metadata representations was examined and analyzed and showed that digital libraries improved with metadata are turning out to be more common (Shiri, 2007).

Similarly, metadata was enriched with extra textual and multimedia data to create Medio Vis, a user centred coordinate multiple views library metadata browser. It is a

visual information seeking system intended specially for users with the aim of encouraging the user to browse by making the user's information seeking procedure easier and enhancing its effectiveness. Evaluation results show a marked improvement of the efficiency of the information seeking process with the aid of Medio Vis (Grun, Gerken, Jetter, Konig & Reiterer, 2005).

Metadata provides information about parts of data that include: purpose of the data, the date and time it was created, the author, how it was created and where it can be found. For instance, metadata describes the size, colour depth, image resolution, creation date and time of a digital asset such as a picture. It could describe the origin, author, summary, type, date and time of a textual document such as a journal.

A number of distinctions have been made between three basic categories of metadata: structural, descriptive and administrative. Some other forms of metadata have also been identified and will be mentioned below.

### Structural Metadata

Structural metadata provides details on how the parts of the digital asset are organised such as the manner in which pages are arranged to form a book. Bretherton and Singley (1994) define structural metadata as one concerned with the design and measurement of data structures. They claim that it describes the structure of database objects like tables, columns, keys and indexes.

### Descriptive Metadata

Descriptive metadata are information utilized during the search process. It holds information that describes the information resource referred to by the digital asset and

supports people to identify and locate these assets. It includes details like title, author, subjects, keywords and publisher (Vellucci, 1998).

#### Administrative Metadata

This type of metadata includes technical information such as file type. It is divided into two sub-types: rights management metadata and preservation metadata. Preservation metadata is information that documents and assists the preservation of digital assets (Guenther and Radebaugh, 2004). Rights management metadata is information that directs the use of digital assets after sale Anderson (2001).

#### Other Forms of Metadata

In addition to the other types of metadata discussed earlier, Wodtke, (2003) recognizes *intrinsic metadata* as an additional type of metadata. Intrinsic metadata draws attention to the file size and resolution of graphical images. It can be split into structural or semantic metadata. *Geospatial metadata* describes geographic objects like maps, ecological and environmental metadata give details of who, what, when, where, why and how data was collected for a particular study.

*Guide metadata*, is a collection of keywords written in normal language to help people find definite items (Bretherton and Singley, 1994). *Business metadata* gives the user information such as the type of data user has retrieved the origin of the retrieved data and its relationship to other data in the data warehouse. *Process Metadata* explains the outcome of processes in the data warehouse and gives details like start time, end time, disk reads, disk writes and rows processed. *Technical Metadata* defines digital assets and procedures from a technical perspective (Kimball 2008). It defines data



models and its mode of display to viewers, in addition, it defines data structures like tables, data types, databases and fields. Other metadata include hash tags as used in social media. A picture on Instagram, for example, will have a hash tag such as 'sunset' allowing people to search and find all the pictures of sunsets. Combining with other tags such as the location 'Edinburgh' will find pictures of sunsets in Edinburgh. So, although this is classified as metadata, there is no intrinsic difference between the data and the metadata. The data about the picture (sunset in Edinburgh) is the metadata that will help people find the picture they want. This portrays the idea of one person's data being the other person's metadata. Most recently this distinction has been made in the context of the National Security Agency (NSA) in the USA accessing the metadata about e-mails. This includes the e-mail addressee where it came from, the subject, and information concerning when it was sent and received. Thus metadata about an e-mail becomes very useful data for the analytics in which the NSA is engaged.

### **2.3.3 Data models**

A data model is an object developed to provide a detailed definition and design to structure the data used within information systems. Data models ensure that compatibility of data is achieved when these models are used consistently across systems. They are used in two ways in software design; they give details about the objects held by the computer system such as products, suppliers, clients and order, the properties of the objects and how they are related. On the other hand, data models are a set of rules that define these models. During data modeling database requirements are developed, described and analyzed to generate a database (Simison, 2009). Data models consist of entity types, attributes, relationships and their

description. It is the beginning point for an interface or database design. It is founded on data, data relationships, the meaning of these data and the factors that limit the data (Smith & Sarsfty, 1993). A data model is usually expressed in a graphical form and provides details of the data structure in a clear manner (McCaleb, 1999).

### 2.3.4 Database Models

A database model is the type of data model that decides the method data is stored, ordered and maneuvered. It establishes the logical structure of a database. General types of database models include, the relational model (Codd, 1970), the hierarchical model, the entity set model (Senko, Altman and Astrahan, 1973), the network model (Bachman, 1975) and the entity relationship model (Chen, 1976).

#### Relational Model

The relational model was invented in 1969 by Edgar F. Codd to manage databases. It looks at objects in terms of the data the entities hold. In a relational model of a database, data is represented in form of an ordered list of elements referred to as tuples, which are grouped into relations. A relation (fig 2.1) can be expressed visually as a table which has tuples (rows) and attributes (columns).

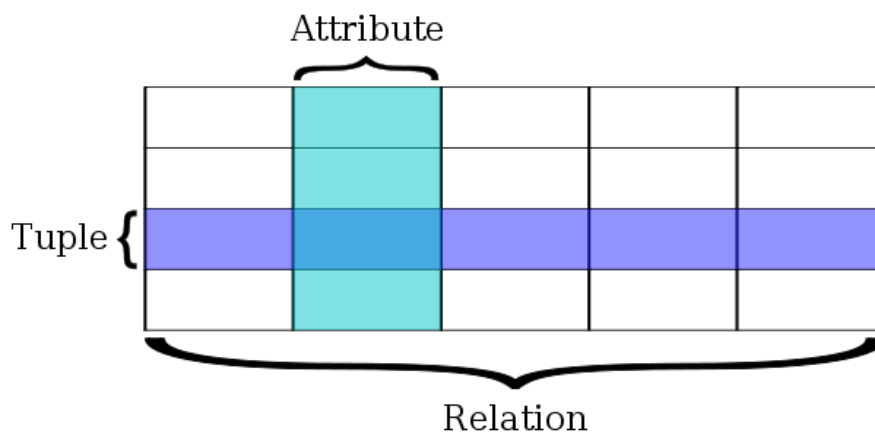


Figure 2. 1: Diagram of a relation

Source: Adapted from Wikipedia, the free encyclopedia.

When a database is organized according to the terms of a relational model, it is referred to as a relational database. The central purpose of relational model is to describe the characteristic function of the database and provide a means for data and queries to be written in detail, for instance, the information the database contains and the information that can be retrieved from it. Users can request data from these relational databases through queries. The relational model (fig 2.2) makes it possible for database designers to create a logical representation of information that is reliable. It is flexible in nature as it allows programmers write queries that were not expected by the database designers and it can attain a certain level of data independence; however it lacks the use of relevant language to portray information about real life. The information principle which sees information as being represented by data values is the fundamental principle behind the relational model.

### Relational Model

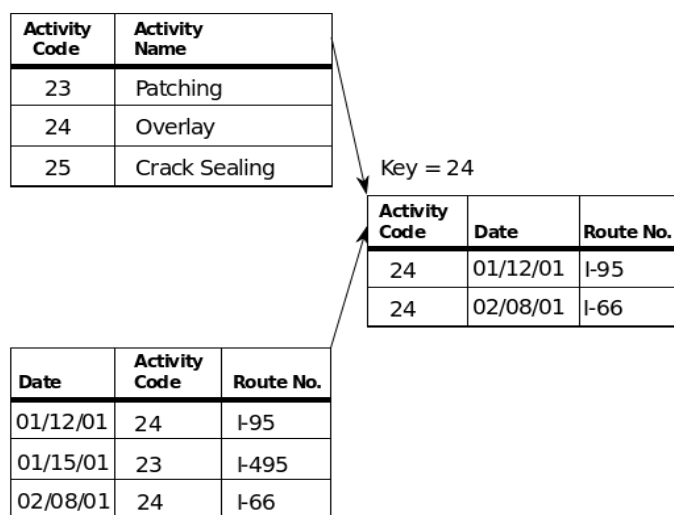


Figure 2. 2: Diagram of a database created in terms of a relational model

Source: Adapted from Wikipedia, the free encyclopedia.

## Entity Relationship Model

The entity relationship (ER) model was developed as far back as 1976 by Peter Chen and remains a significant method of presenting data logically. It is a database model that explains a database in a theoretical manner and displays the information about the physical and functional needs the design must be able to execute. The theoretical aspect explained by the ER model is prepared in a manner that allows it to be implemented as a database, such as a relational database.

From the point of view of the E-R model, the actual world is seen as a collection of fundamental objects called entities. These entities along with attributes and relationships are the main components of an ER model. Entities are connected to one another by relationships that explain the dependencies and requirements linking them. For instance: one centre may have many members, but one member can only belong to only one centre. An entity is a relation. It can be a tangible object such as a specific person or a house; it could be an event like a car sale or a concept such as an order placed by a customer. An entity outlines similar information and plays a role (Chen 2002). In addition, it exists independently and can be exclusively identified.

The characteristics that describe these entities are called *attributes* like for a specific person, the attributes would be persons – name, national insurance number, address and phone number. The association that exist between these objects are called relationships, and it shows how data is shared between these entities. An entity type describes a set of entities that share similar characteristics or properties. Each entity type is described using metadata; this is done by assigning a name usually singular in nature to it as it represents a collection of items. For instance an artist performs a

song, the artist and song are entities while performs is the relationship between them. Entities and relationship have attributes; the entity employee may have an attribute such as social security number, SSN, while date is the attribute of the relationship, proved.

An ER model (Figure 2.3) is developed through three levels, the conceptual, logical and physical phases. During the initial stage of system design, these models are used to describe the type of information to be stored in the database.



Figure 2. 3: Diagram illustrating two related entities

Source: Adapted from Wikipedia the free encyclopedia

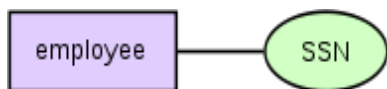


Figure 2. 4: Diagram illustrating the entity employee with an attribute SSN

Source: Adapted from Wikipedia the free encyclopedia.

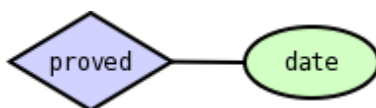


Figure 2. 5: Diagram illustrating a relation 'proved' with an attribute 'date'

Source: Adapted from Wikipedia the free encyclopaedia.

Conceptual data model

The conceptual model is the initial model created during the ER process, and is sometimes used as its foundation. It gives the general picture of the relationships required for the database and explains the entity groups, the characteristic of the attributes and relationships. It establishes the structural metadata and is also used to create relationships between ER models that have common attributes when integrating data models.

The logical data model

The logical data model is a diagram that illustrates the theoretical structure of an information space. It is more detailed than the conceptual model; it captures important relationships, and records data structures required to create the database. The details of data entities and the relationships between these data entities are established.

The Physical data model

A physical data model represents a database design. It allows data to be stored in tables in an orderly manner to facilitate access. It defines the data elements, their organization and the association between these data (West and Fowler, 1999). It is created from a logical model and explains a database.

When designing an information system that is founded on a database, the conceptual data model process brings about the logical data model such as a relational model which is mapped to a physical data model. E-R diagrams in computing symbolize the organisation of data in databases or information systems. The E-R model is a useful graphical approach of looking at objects. The major relationships that exist in E-R diagrams are described by Dennis, Wixom and Roth (2012) as one-to-one, one-to-

many, many-to-one and many-to-many. Every relationships is seen to have a parent entity and a child entity, it also has cardinality which shows the ratio of parent occurrences to child occurrences.

ERD Element	Kinds of Metadata	Example
Entity	Name	Item
	Definition	Represents any item carried in inventory in the supermarket
	Special notes	Includes produce, bakery, and deli items
	User contact	Nancy Keller (x6755) heads up the item coding department
	Analyst contact	John Michaels is the analyst assigned to this entity
Attribute	Name	Item_UPC
	Definition	The standard Universal Product Code for the item based on Global Trade Item Numbers developed by GS1
	Alias	Item Bar Code
	Sample values	036000291452; 034000126453
	Acceptable values	Any 12-digit set of numerals
	Format	12 digit, numerals only
	Type	Stored as alphanumeric values
	Special notes	Values with the first digit of 2 are assigned locally, representing items packed in the store, such as meat, bakery, produce, or deli items. See Nancy Keller for more information.
Relationship	Verb phrase	Included in
	Parent entity	Item
	Child entity	Sold item
	Definition	An item is included in zero or more sold items. A sold item includes one and only one item.
	Cardinality	1:N
	Modality	Null
	Special notes	

Figure 2. 6: Diagram illustrating the use of metadata in an E-R diagram

Source: Adapted from John Wiley & Sons, Inc. @ copyright (2011)

### 2.3.5 Entity Relationships

#### One-to-one

In a one-to-one relationship, a single occurrence of an entity (J) is linked to a single occurrence of another entity (K). A case in point is that of a database of members of

a research institute, each member name (J) is linked with only one MemberID (K). In summary one occurrence of parent entity is linked with one occurrence of child entity.

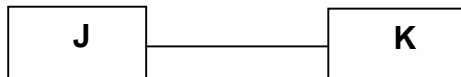


Figure 2. 7: Diagram illustrating one-to-one relationship (1:1)

### One-to-Many

In this instance, a single occurrence of an entity (J) is linked with zero, one or many occurrences of another entity (K). However for one occurrence of entity (K) there is only one occurrence of entity (J). For example all members of a centre work in a research institute; the name of the research institute (J) is linked with many different members (K), but all members share the same particular association with entity (J). In summary, a single occurrence of parent entity is linked to many occurrences of child entity. The one-to-many relationship is the most commonly used among databases.

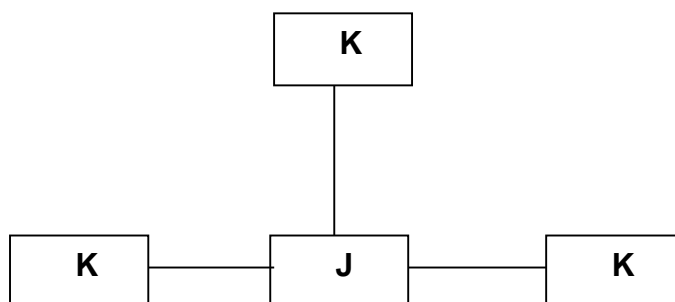


Figure 2. 8: Diagram illustrating One- to- Many relationships



## Many- to-Many Relationship

In this case a single occurrence of an entity (J) is linked with a zero or many occurrences of another entity (K), while one occurrence of entity (K) is linked with one, zero or many occurrences of entity (J). A case in point is that of a research institute in which all its members work on numerous publications. Therefore, every occurrence of a member (J) is linked with many occurrences of a publication (K), and simultaneously, every occurrence of a publication (K) has numerous members (J) linked to it. In summary many occurrences of a parent entity can relate to many occurrences of the child entity.

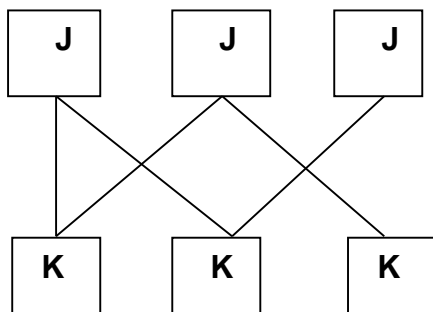


Figure 2. 9: Diagram illustrating Many-to-Many relationship

## Many to One Relationship

A many-to-one relationship is where one entity, usually a column or a set of columns holds values that refer to a different entity usually a column or a set of columns that has unique values. These many-to-one relationships are usually imposed in relational databases by using a foreign/primary key relationship and these relationships exist

between a piece of information and dimension tables or between levels in a hierarchy. The relationship usually describes a set of people sharing something in common (group) or categories. For example, in a MIS that holds tables such as Faculty, Institute and Centre, there will be many institutes in a given faculty, but no institutes are in two faculties. Similarly, a centre belongs to only one institute. The point worthy of note is that each centre exists in just one institute, but an institute may have many centres, consequently the phrase “many-to-one.”

### **2.3.6 Metadata and E-R Model relationship**

Metadata plays a vital role in the structuring of the entity E-R model. The metadata is stored in the data dictionary of the E-R model and includes information about entities, attributes and relationships. It is kept in order to allow users and developers share and make use of the information it provides throughout the systems development (Dennis, Wixom and Roth 2012). An example of this type of information as illustrated in fig 2.6 includes details about the data called metadata and administrative details on how the end user data is grouped.

The E-R database contains structures such as two dimensional tables that help to order the data it holds and these tables contain rows and columns. Categories of data described by metadata for each record are entered in columns and the table as a whole represents an entity in a database entity relationship diagram. A primary key usually unique in nature is used to identify each record in a table. For example, a record about an author is identified by a unique primary key, AuthorID.

Metadata describes data such as names given to attributes usually table headings; it defines the table and shapes the structure of the database. In addition, metadata adds meaning to values; it explains that a string of numbers "01829384754" recorded as a value in a field means telephone number.

In formatting the E-R diagram limits are placed on data held in databases; these limits are defined by metadata. It states clearly the size of fields and the type of data they should hold. It also defines the format data such as dates should take in some cases. For example in a relational database, the metadata identifies the table that holds certain data by listing its columns, column names, data types and the width. It also tells when a field in a column is optional or compulsory in a record.

The type of relationship that exists in the entity relational model of CMVs creates an impact on the navigation support the CMV can provide. In general the E-R diagram displays data, the individual pieces of information that have been converted to a form easier to process and metadata, which describes the data and facilitates interpretation, access and understanding of the information of an organization displayed. Metadata means different things for different people. On the one hand it is data about data where it describes these data; on the other hand it is access to data where it supports people to move in a specific direction to locate data in physical spaces such as libraries or conceptual environments such as information spaces. This research is interested in the aspect of metadata that facilitates navigation and location of objects on information spaces by people in order to visualize and retrieve adequate information.

## **2.4. Models of Navigation**

### **2.4.1 Navigation**

Navigation is described as the ability to move successfully through an information space in order to retrieve information and discover existing and new relationships. It acts as a support for interactive visual interfaces and is also described as an interactive technique which improves people's ability to communicate with data. Navigation of information space is influenced by key features that could enhance learning; they include shape, colour, size and location of objects. These features instil quality and time and offer the user space to move around on the information space, it encourages people to explore or on the other hand may overwhelm the user and create unwillingness in the user to explore further.

The concept of an information space has been developed over a number of years (Benyon, 1998, Benyon, 2001, Benyon, 2006, Benyon, 2007), Spence (1999), Darken and Paterson (2001), Dahlback and Lonqvist, 2000; Ahmed and Blunstein (2005) and some others all agree that navigation involves movement from one point to another and requires the navigator to acquire knowledge of the information space and interpret it correctly. However they present their views from different perspectives. Spence (1999) explains navigation as the process where people access an information space, gain knowledge of its contents, create an internal model, interpret this internal model and from their interpretation develop a browsing approach that enables them browse the information space.

Darken and Paterson (2001) describes navigation as the collective task of wayfinding and motion. He claims navigation must involve both; wayfinding which he explains as the reasoning aspect and motion, the motoric aspect. Benyon (2001) defines navigation as the act of moving from one position to another in a spatial location and being aware of the relative point in order to arrive at an expected destination. He claims that it involves understanding the location, separating it, discovering the correct route to the end from the start point.

Benyon (2006) further explains that navigation is moving through an information space to collect all the information needed. He explains it as movement targeted at information discovery and involves moving through an information space, discovering information and relating to its meaning. Similarly, navigation in information space is considered by Ahmed and Blunstein (2005) as the process where people move all the way through virtual information networks following connections from node to node and knowing their position in the network relative to the target node.

A group of psychologists, Kuipers (1982), Garling, Book and Ergesen (1982) hold the opinion that a phase of navigation entails “learning to find ones way in a new space”. One way to achieve this “learning” Lynch (1961) claims is by understanding key features of the new space. These features which include nodes, edges, districts, paths and landmarks communicate information to the users as they move through the information space. Nodes are described as points on information spaces linked by paths, they convey specific meanings and make up landmarks while districts are areas on the information space that can be identified and are defined by edges. It is noteworthy to point out that landmarks are mostly personal (Benyon, 2006).

Navigating the information space can be described in numerous ways. It is generally termed movement. It could be browsing, searching or scanning. It has also been referred to as a task (Dahlback and Lonqvist, 2000). They argue that navigation is not the same as information retrieval.

Spence (2001) concludes that people navigate in order to develop an internal model which is a significant factor in information visualization.

It can be inferred from Spence (2001) claim above that navigation is an action that precedes visualization. In other words if people are unable to navigate in order to create an internal model, their ability to visualise may be hindered. This statement contradicts results from a study by Dieberger (1997) which claims that visualizing an information space is necessary for effective navigation. However, this research is of the opinion that in order to attain effective visualisation, people need to be able to navigate through the information space successfully.

Navigation models have been created (Neisser, 1976; Downs and Stea, 1977; Passini, 1984; Darken, 1996; Jul and Furnas, 1997; Spence, 1999; Chen & Stanney, 2000) to illustrate the navigation process and break down navigation tasks however Darken (2001) argues that these models do not capture all aspects of the task. An awareness of how information spaces are designed is crucial to identifying a solution to navigation challenges such as disorientation, cognitive overhead and the inability to achieve successful navigation. It could also point to ways the navigation experience can be made enjoyable. Darken and his colleague Peterson, (2001) claims that if it is

possible to break down general navigation tasks to simpler forms, then identifying the point where people need support or training may be feasible.

Five navigation models of interest to this research are presented in this thesis to illustrate different navigation processes.

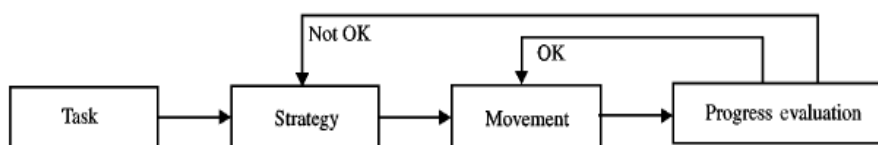


Figure 2. 10: Darken’s model of navigation

Darken’s model (fig 2.10) demonstrates a process where a user looks first at the task, forms a strategy, makes a move and performs a progress evaluation whether the strategy is ‘not ok’ or the movement is ‘ok’. Spence (1999) proposed an initial framework for navigation (fig 2.11), showing four navigation strategies; browsing, modelling, gradient perception and movement.

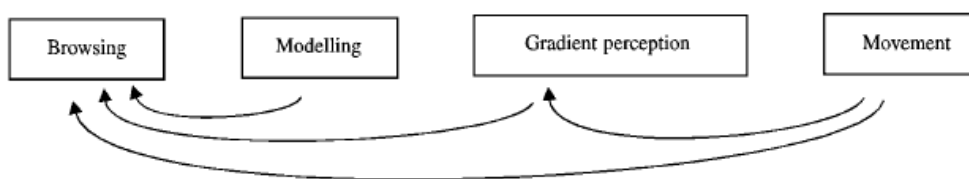


Figure 2. 11: Spence’s initial proposal of a model for a framework for navigation

Darken’s (2001) model in fig 2.10 and Spence’s (1999) model in fig 2.11 was later modified to reflect the goal of navigation as illustrated in Fig 2.12. The model (fig 2.13) proposed by Jul and Furnas (1997) and the modified navigation model (fig 2.12) by

Spence (1999) illustrate a more detailed process and shows the actions people may undertake as they navigate through an information space. Earlier studies (Darken and Peterson, 2001) have looked at the models of navigation from the point of view of understanding how navigation tasks are created with the aim of improving performance. This study will look at models of navigation from a different perspective. A close look at each model shows a different approach to navigation which demonstrates that navigation can be approached from several angles, and while one navigation model may suit a person's need, it may not be practicable for some others.

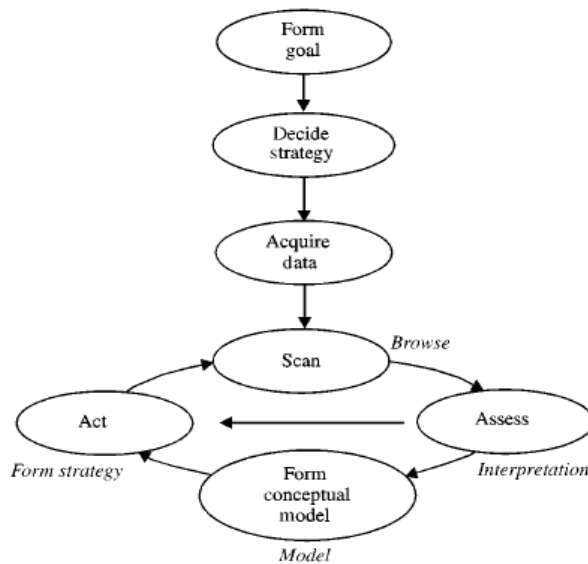


Figure 2. 12: Modification to Spence's original proposal



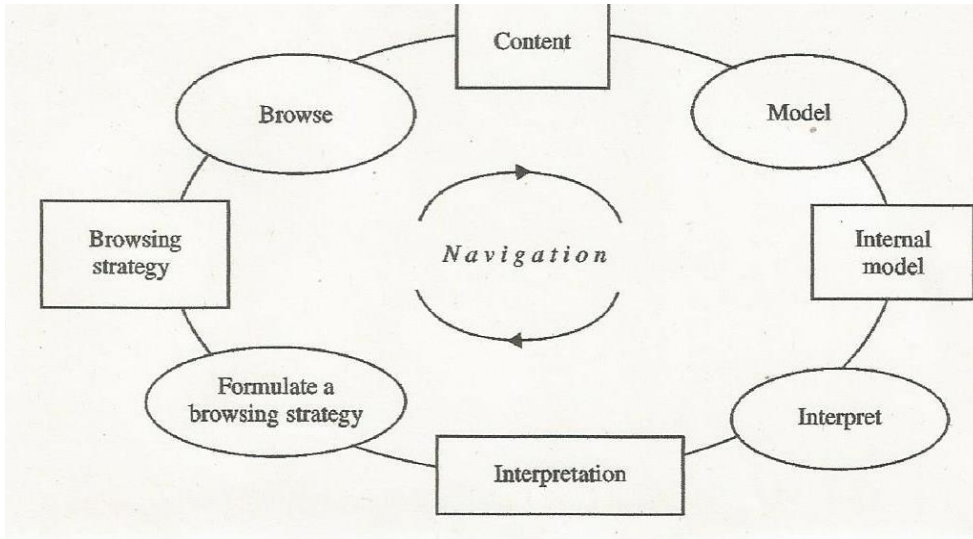


Figure 2. 13: Proposed framework for navigation

Source: Adapted from Spence 1999

The diagram of the framework of navigation in fig 2.13 is utilized by Spence (2001) to explain the process involved in navigating an information space. He claims that in general, navigation can be defined as the formation and understanding of an internal mental model. The framework for navigation shows that the navigation process involves developing and exploring the internal model, which supports his claim that the internal model facilitates visualization.

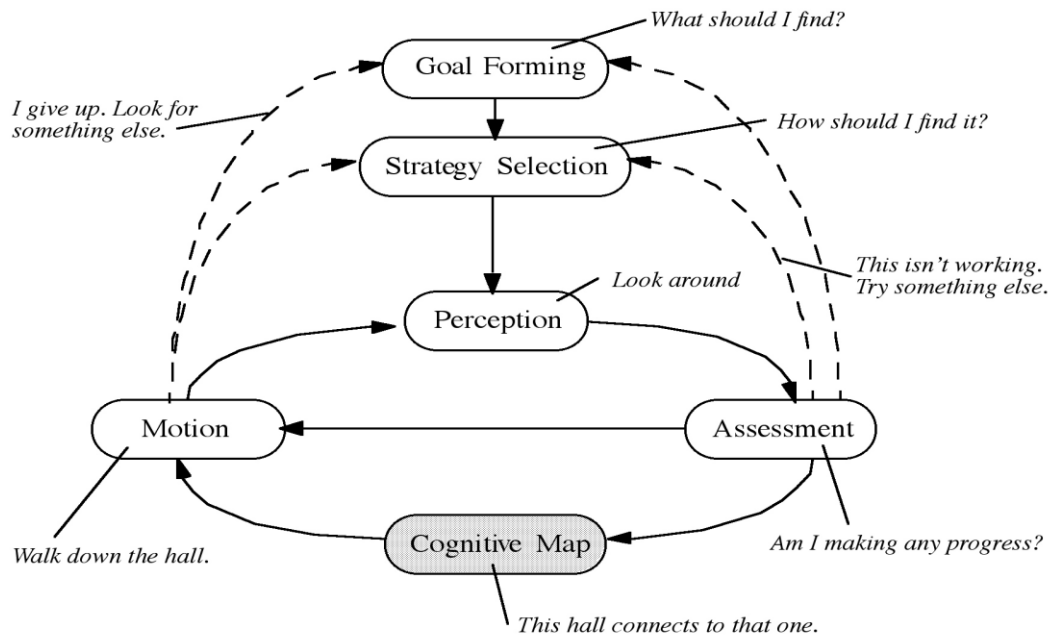


Figure 2. 14: A schematic model of Navigation

Source: Adapted from Jul and Furnas (1997)

The need for navigation arises when a user needs to visualise an information space to retrieve adequate information Spence (2001). He concludes by defining navigation as “cognitively directed movement in information space based on the interpretation of a mental model and/or externalized data.

Each of the models presented above show a method people can utilize to navigate. One notable difference between the models is that some are more detailed than others; however, they all demonstrate that navigation is a flexible and iterative process and requires a lot of movement. This research explores the idea that people may choose to approach navigation using any of the methods illustrated by the models, and that people will most likely make more progress navigating an information space designed to provide users with adequate support, irrespective of the navigation model

a user chooses to utilize. For instance, a person navigating the physical world may choose to ask questions or use a map, in the virtual world people's navigation behaviour may reflect Darkens (2001) model or the model proposed by Jul and Furnas (1997). Whichever model is chosen, this study hopes to demonstrate that people can achieve successful navigation on information spaces where adequate support is provided on information spaces.

#### **2.4.2 The Navigation process**

The navigation process incorporates several stages such as browsing, strategy selection, modelling, movement, assessment and goal forming. Some of these processes will be explained below.

##### **Browsing**

Browsing means scanning or looking around, it addresses and answers the question "what's there?" and is expressed by Spence (1999) as the "assessment of content".

The user browsing may not need to "search" or have a definite goal like a person looking through the menu of a computer program like AutoCAD to determine the possible operations but may scan through the morning newspaper and then choose a story to read. As users' browse, they assess information which they put together to create an internal model or cognitive map. The possibility of different people creating different internal models exists. The user's look at the internal model along with the data provided on the information space, form an opinion and reach a conclusion. The user decides the task is complete which means no further browsing is necessary or the user decides the model is not sufficient and goes on to develop a browsing strategy. Going back and forth through this process repeatedly is described by Spence

(1999) as the navigation process. Solso (1998) claims that browsing and the act of being perceptive are alike, this agrees with Spence's (1999) point of view that navigation is a cognitive process.

On the other hand, Carmel, Crawford and Chen (1992) hold the opinion that browsing are of various types, they discuss three types; "scan-browse" which they describe as people scanning for exciting information, "review-browse" as when people compile previously browsed information while "search-browse" involves exploration, seeking out information.

### Modelling

Modelling is when people create an internal model of the data they have browsed for example a tourist by browsing a map will have an idea of the area in a city where a castle, a cinema or mall can be found. An internal model is thus formed when people browse an information space to check what is available. They may be looking at a particular city to see the patterns that come into view. The way the pattern is represented ends in the internal model or cognitive map. The two events, browsing and modelling usually take place simultaneously. Modelling is similar to review-browse (Spence, 1999) and is influenced by externalization. The manner in which the data is externalized or displayed directly impacts on modelling. It either advances or impedes the formation of an internal model (Spence 1999). Externalization is the process of converting data to a form that makes it suitable for people to process. It involves three structural stages identified as inherent.

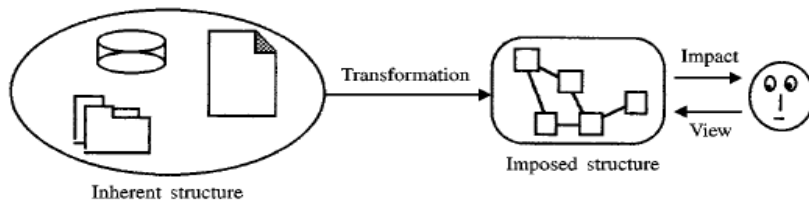


Figure 2. 15: Transformation of inherent data structure into imposed structure

Source: Adapted from Spence, 1999.

In the inherent structure stage, the nature of the data is transformed through the process of selection, encoding and interpretation to form an imposed structure which influences the nature of the cognitive map developed by the user as illustrated in fig 2.15. This imposed structure is where browsing is performed and then an internal model is created in order to support navigation (Jul and Furnas 1997). The externalization is created to improve the user's internal model of certain data in order to improve the user's interpretation of the data.

The interpretation stage is where the user makes a choice based on the externalized data and internal model on how navigating the information space should proceed. Here strategies which are methods people utilize to move around the information space are required.

Strategy selection is utilized to help the user decide if further browsing is required? Or the task solving approach should be changed? A progress evaluation can also be carried out at this point.

Navigation is a significant component of countless human-computer interactions comprising search and problem solving (Spence, 1999). It facilitates visualization and enhances information retrieval. Navigation involves certain activities termed *navigation activities* that determine its outcome. Three of these activities are revealed by Benyon and Hook (1997) as object identification, wayfinding and exploration.

### **2.4.3 Navigation Activities**

Navigation activities can be expressed as the actions people undertake as they move from one location to another on the information space. Several opinions as to what exactly make up these actions have been explained by several Information experts. The navigation models discussed earlier give an insight into some of the processes people undertake in order to achieve navigation. Navigation activities are the actions people carry out to accomplish these processes. An overview of the processes illustrated by the navigation models (Figures 2.11, 2.12, 2.13, 2.14 and 2.15) are discussed below.

#### Object identification

Object identification as the name implies involves identifying objects on the information space, getting to know what's available and what they mean. Objects include nodes, keys, symbols, landmarks, matrixes, tiles, information artefacts or anything used to represent information. In order to identify objects users need to gain sufficient knowledge of the information space to be able to discover collections of objects scattered all over the environment, group these objects, spot significant trends about the way they are arranged and retrieve information about them. It is similar to the internal model formation described by Spence (2001). The objects available on the

information space are utilized by user's to create these internal models which help them interpret the information space.

### Exploration

This is when people seek out information about locations on the information space and how these existing locations relate to other locations. Object identification and exploration are similar however their purposes are clearly different. Object identification discovers groups of objects while exploration aims to understand these objects that exist in the environment and the conceptual or physical relationship between them (Benyon, 2006). Take the case of a tourist standing on a mountain at a point which overlooks the city. The tourist is positioned to discover locations of interest like the shopping centre, a river and the train station and therefore determine their relationship in terms of distance, similarities such as colour, size and structure. This is exploration. The same tourist can identify objects like the castle and the tourist bus. Exploration involves searching to find relationships that exist within the surrounding and other surroundings, locating common patterns and trends that exist. The actions that take place during browsing such as finding out "what's there" in an environment show it's akin to exploration.

### Wayfinding

The wayfinding activity is one that has to do with how user's find their way, user's need to work out how to reach their destination, recognise the routes they take and be able to tell when they achieve their goal which is to arrive at a specific destination. This illustrates that people understand their location on an information space, have knowledge of the target destination and know the route required to get there.

Dahlback and Lonqvist (2000) claim navigation and wayfinding have been used in ways that suggest they are the same. However, Benyon (2006) argues that wayfinding is distinct and describes it as when people have identified a destination and move in the direction required to lead them to the path. Darken (2001) holds the view that wayfinding is the cognitive aspect of navigation. He claims that it is the creation and use of a mental map of the environment.

It involves user's ability to interpret the information space, observe events that take place on the different views and find their way around as they search for specific information. It is defined by Passini (1984) as "spatial problem solving" and is related to interpretation, one of the navigation processes described by Spence (1999). Downs and Stea (1973) and Passini (1984) claim that wayfinding involves four steps; becoming familiar with the information space, selecting the proper route, keeping an eye on the route and being aware as soon as the destination is reached. The ability to recognize paths that one has taken before to reach required destinations is wayfinding. The manner information spaces are structured have the potential to support individuals to achieve successful navigation or on the other hand make navigation complicated.

Benyon and Imaz (1999) discuss how different underlying metaphors inform the theories and origins of various subjects. For example, the changes of perception from a metaphor of 'processes are machines' was replaced with the metaphor of 'objects are people' when the discipline of software engineering moved to an object-orientated paradigm. The challenge to bring about a similar change is essential for HCI and Navigation of Information Space which sees 'people as navigators'.



When the term 'Human-Computer Interaction' (HCI), is referred to, one immediately imagines a person sitting at a visual display unit staring at the world of 'information'. It is believed that the person is very much outside the space of information. But when other activities such as going shopping, having a meeting or driving across town are thought about, the person is seen to be inside a space of activities, surrounded by, and interacting with, assorted artefacts and people. Benyon (2001) argues that "Navigation of Information Space is an alternative conceptualisation of HCI that sees people as existing inside information spaces" (p.425). This concept could impact on the way information spaces are created. An information space designed for a person existing in a space could have some variations from one designed for people existing outside the same space. Benyon (2001) goes on to suggest that looking at HCI in this way will bring about a change in HCI design. The design will be seen as the creation of information spaces. He claims that "Navigation of Information Space is not only a metaphor for human-computer interaction; It is a 'paradigm shift' that changes the way HCI is perceived" (p.425). This way of reasoning has brought about some changes to the way information systems are designed (Munro, Hook and Benyon, 1999), to usability (McCall and Benyon, 2000) and information gathering (Macaulay, Benyon and Crerar, 2000).

#### **2.4.4 Navigating interactive visual Information Spaces**

The concept of the information space described earlier leads to the helpful perspective that rather than think of people as retrieving items from a database, the spatial metaphor makes us think of people as navigators. The way people navigate through information spaces is similar to the way they navigate through physical spaces.

Spence also uses this idea of navigation as a central part of his views on visualization (Spence, 2001). He illustrates visualization as when a person forms a model, determines where to go next, progresses and restructures the process in an iterative fashion. In the real, physical world people do not simply retrieve objects; people are navigators. For instance, people navigate physical information spaces by imagining various routes and planning how to move from one location to another. They utilize various strategies such as measures of distance and fixed direction. They use landmarks and signs. The strategies and mechanisms people utilize to enable them move around the information space are referred to as navigation systems. The choice of navigation mechanisms are influenced by the size, orientation and the way the information space is arranged (Dahlback and Lonqvist 2000). These navigation mechanisms are grouped broadly into Egocentric navigation, Allocentric navigation and Agent driven navigation.

In Egocentric Navigation people move through the information space searching, identifying, exploring, wayfinding, panning and zooming. In order to navigate a route, local landmarks, left and right directions which make it easier to use in well-known rather than strange information spaces (Andreano and Cahill, 2009).

Allocentric navigation uses fixed directional terms such as north, south, east and west to measure the distance from one point to another in order to determine the best route. It is beneficial in large and strange information spaces because it uses mental maps instead of physical signals to navigate (Geary, 1998).

In Agent Driven navigation people are directed by agents such as search engines in the World Wide Web as they navigate the information space Robertson (1997) navigation systems They move through the world exploring, finding their way, arriving at destinations, engaging with objects and moving on.

In a comparable style, navigation of information space sees people as moving through information spaces as they interact with the data through different visualizations of that data (Benyon, 2006; 2007). In some situations they have a clear need for some information about an object; this is the activity of object identification. For example some facts about an individual who works in a research centre may be required. At other times they undertake an exploration of the space. Here the activity concerns browsing around to see what information the database holds and how it is structured. At other times they know where they want to go and how to get there; this is termed wayfinding. For instance if information about the name of the paper that was co-authored between Dr Y and Dr X is required, Dr X can be looked up and the list of publications co-authored found and the specific one that is co-authored with Dr Y is selected and read off the title, thus wayfinding from person to co-publications to names of publications.

Navigation is not a novel interaction technique but has the potential of being applied in novel ways to improve data visualization and information retrieval on interactive visual information spaces such as CMVs. One way to optimize these Interactive spaces, is to support people navigate through successfully, more easily and quicker. Successful navigation will lead to data being visualised more effectively and adequate information retrieval which is ultimately the goal of information seekers. One of the

research aims is to identify novel navigation methods to support people on CMVs and enable them achieve faster task conclusion times and move successfully through data. A case in point from one of the studies carried out in this research shows a user following a route supported by the centre data provided in order to retrieve specific information.

#### **2.4.5 Navigation Challenges on Information Spaces**

The factors that affect people's ability to achieve successful navigation on an interactive information space have been identified by earlier research as cognitive skill, spatial ability, mental models and others. These factors are well known in navigation in the physical world (Passini, 2000a) and have been applied to information design (Passini, 2000b). The designers of physical spaces provide maps so that people can get an overview of the whole space, signposts to guide people to different parts of the space, and place information signs at strategic points so that people know when they have reached their destination. In the same manner, virtual information spaces should provide support to aid users navigate through.

In face of the growing size and dimension of information spaces like the World Wide Web (WWW), the challenges encountered by user's as they navigate is being taken seriously by the HCI community. Huge information spaces such as the WWW can concurrently confound and divert people searching for information as they navigate. Some challenges encountered by users which the research seeks to address include cognitive overhead and lost in hyperspace (Edward and Hardman, 1999) amongst others are discussed below.

## Cognitive Overhead

Cognitive overhead is defined by Conklin (Thuring, Hannemann and Haake, 1995), as “the additional effort and concentration necessary to execute several tasks or trails at one time” (p.40). Well-designed information spaces usually reduce the amount of reasoning people need to make. Cognitive overhead relies on the ability of people to process information and so it differs between individuals for a specific information space.

People have different levels of cognitive ability which reflects in the way they think, perceive, resolve issues and memorize information. Hook & Dahlback (1997) found that people that possess a high level of spatial cognition are able to navigate better than people with lower spatial abilities. The spatial ability of individuals cannot be determined at face value, moreover these spaces are designed for anyone who needs them and therefore should incorporate people with different spatial abilities.

According to Dahlback & Lonqvist, (2000) a number of studies have been undertaken to ascertain the association existing between cognitive ability and navigation in the information space. Benyon and Murray (1993) and Dahlbäck, Hook and Sjolinder (1996) all discovered connections between how users navigate an information space and their spatial cognitive ability. This discovery points to the possibility that metaphor may not be the chief solution to navigation problems on the information space.

## Spatial Ability

The spatial ability of a person can be expressed as the ability to think about and understand spatial properties such as location, size, distance, direction, shape,

movement, and so on. The awareness of the information space gained by the user as a result of earlier visits and the ability to assimilate the available information also influences navigation abilities. In order to understand an information space, a mental model influenced by the user's spatial ability is created. Studies carried out earlier propose that users create three types of mental models; landmark, route and survey during navigation (Dillion & Vaughan, 1997). Landmark awareness is acquired at the initial stage of interaction. The user obtains information on the exclusive properties of the information space. Route knowledge is defined by Dillion, McKnight & Richardson (1993) as "the ability to navigate from point A to point B utilizing the landmark knowledge acquired to make decisions about when to turn left or right" (p.173). Survey knowledge is developed in the final stage of navigation and helps the user find landmarks and routes. Thuring, Hannemann and Haake (1995) claim that coherence and Constancy of objects such as colour, size and visual origin of object (Raubal, n.d) are elements considered to affect the formation of mental models and spatial ability.

### Lost in Hyperspace

The vast amount of information user's encounter in interactive information spaces such as hyperspaces can make them feel lost (Smith, 1996). Lost in hyperspace occurs when the user becomes disorientated as a result of jumping from one part of the space to the other as they navigate. Bernstein as seen in Schulmeister, (n.d) believes that lost in hyperspace stems partially from the interface designer and also from the initial layout of hypertext. Some people think it is not a vital issue in the information world while others look forward to a way out in future (Theng, 1999). An

experiment carried out by Theng (1999) confirms that lost in hyperspace is dominant in digital library navigation which has led to ongoing research in this area.

### Serendipity

Serendipity can be expressed as the other side of lost in hyperspace. In many cases users arrive at a point other than their intended destination only to find it more exciting than their target on the information space. Users do not see themselves as lost as a result of their exciting discoveries. In view of this users may not take this as a problem but it is a problem in the sense that the user has deviated from their major focus of looking for specific information. This problem could be as a result of an insubstantial navigation support however according to Cooper-Kuhlen (1992) it is seen by some as a “delightful accidental discovery”, even though some others hold a different opinion.

### Segmentation and Contextualization

Segmentation and Contextualization are factors that may contribute to challenges encountered by people as they navigate information spaces (Cooper-Kuhlen, 1992). When the Information space holds too much information usually represented in form of atoms and nodes along with lots of sections it makes the system difficult to navigate. On the other hand in contextualization, each information node or entity is grouped and could result in a substandard navigation practice.

A few other constraints to individual navigation are gender, physical challenges and learning tools. Males have been found to navigate virtual worlds more effectively but females have been found to possess the ability to cope with a larger view of the information space (Czerwinski, Tan and Robertson, 2002) hence they navigate the

real world by taking separate approaches (Halpern, 2000) and (Kimura, 1999). Another relevant factor in navigation is the learning technique. Dillon (1991) conveys the idea “that readers of scholarly works” assume an Introduction, Method, Results and Discussion (IMRD) approach. Many of the established documents are organized to compel users to learn the IMRD layout and they believe they will find a similar layout when they navigate new documents.

This research proposes that where users receive substantial navigation support from the information space, the navigation challenges they encounter will lessen. Two theories; the information scent and the information foraging are presented to illustrate this point. Information scent was first invented by researchers working at the Xerox’s Palo Alto Research Centre (PARC). Their studies showed that the way people search for information is akin to the way animals hunt for food.

Trepess (2013) points out that one challenge the information retrieval community face is to design information spaces that effectively support information foraging (comparable to navigation) concepts. Information spaces that offer rich sources of metadata, helpful navigation cues, effective content categorisation that support suitable strategies for different information requirements.

#### **2.4.6 Information foraging**

The Information foraging theory illustrates the significance of navigation as regards information retrieval. It was derived from the optimal foraging theory, a food foraging theory that aids biologists understand the factors that determine the food choice and feeding methods of animals (Pirulli and Card, 1999). The theory is concerned with



information; nevertheless it can be applied to objects like text, video, audio and image. The foundation of this theory concerns the cost and benefit evaluation of achieving a goal. Cost refers to the amount of resources consumed when carrying out a specific activity while benefit is the resource gained by carrying out that activity. For instance, the amount of effort a predator uses to track and chase a particular species of prey and the maximum benefit obtained when the lowest amount of energy is expended. The amount of energy expended depends on how the environment is structured. Optimal foraging can be achieved by understanding the strategy to be utilized. The analogy here is this, an information space structured to support navigation will require less cognitive ability, and adequate information will be retrieved when the navigator understands which strategy to use.

Information foraging is developing into an acceptable concept that describes and gives an understanding of web browsing behaviour (Chalmers, 2004; Dix, Howes and Xiao, 2003). It could influence researchers in the information visualization field in a significant way by supporting them to discover effective ways to represent vast amounts of data and provide effective methods for navigating through it.

## Box 2. 1 Extract from Pirolli and Card

Imagine a predator, such as a bird of prey, that faces the recurrent problem of deciding what to eat, and we assume that its fitness, in terms of reproductive success, is dependent on energy intake. Energy flows into the environment and comes to be stored in different forms. For the bird of prey, different types of habitat and prey will yield different amounts of net energy (energetic profitability) if included in the diet. Furthermore, the different food-source types will have different distributions over the environment. For the bird of prey, this means that the different habitats or prey will have different access or navigation costs. Different species of birds of prey might be compared on their ability to extract energy from the environment. Birds are better adapted if they have evolved strategies that better solve the problem of maximizing the amount of energy returned per amount of effort. Conceptually, the optimal forager finds the best solution to the problem of maximizing the rate of net energy returned per effort expended, given the constraints of the environment in which it lives, (Pirolli and Card 1999).

The information foraging theory is based on the comparison drawn from an animal deciding what to eat (goal formation), where it can be found (information space), and the best way to obtain it (navigation and visualization) and how much energy (information) the meal (data set) will provide. The analogy is explained in greater detail in the extract (Box 2.1) from Pirolli and Card (1999).

The aspect of information foraging of interest to this research is one of the key concepts (information scent) that have come into view from the above analogy. The

food source (data set), the location where it can be found (information space) the strategies (the concept of) employed to find it, the tools (navigation and visualization features) available to find it and the benefit (Information retrieval) that can be gained from its utilization. These concepts have been interpreted into an information-seeking context using the following terms:

- Information: The item of information being sought after and the value it has in satisfying the information need;
- Information patches: The chronological and spatial manner in which information is grouped;
- Information scents: The process of determining the value of information based on navigation cues and metadata;
- Information diet: The process of choosing which information source to follow.

### Information Scent

Supporting people find the particular information they need on an information space within vast data is challenging. Research has shown that people navigate by following one path and when they do not retrieve the information they require, they retrace their steps and take other paths which usually turns out to be frustrating as they are unable to retrieve the information they require. Information scent is a concept which supports people to navigate information spaces easily.

It explains how people assess their options when searching for information on an information space. People will likely choose the option that gives them the clearest clue (or strongest scent) that will take them a step closer to the information they need. The stronger the information scent, the easier it is for people to navigate (Barker,

2005). Information scent can be illustrated in the use of an interface key which acts as a guide and gives clues to people navigating these spaces. The more details the interface key provides, and the richer the metadata source, the easier it will be for people to navigate.

## ***2.5 Visualization***

Visualization is the procedure that involves creating visual images such as graphs, diagrams and animations, which respond to actions and sometimes make the first move to represent information and make it easier to understand. Visualization makes it possible to acquire insight and understanding. Spence (1996) says it improves people's imagination, visual skills and abilities. Stuart, Mackinlay and Shneiderman (1999) describe visualization as the use of computer to support interactive, visual representation of data in order to enhance cognition. Visualization alters theoretical and physical data to visual forms (pictures) and makes it easier to gain knowledge and insight in order to make decisions. It is a phenomenon that has been in use for many years, it uses interactive, sensory representations, typically visual, of abstract data to build reasoning. The modern study of visualization was introduced using computer graphics, and it can also be described as the way an image data is represented and manipulated using a computer. Visualization is a form of imagination; it is creating the image of something in the mind. The Encarta dictionary defines it as "creating a positive mental picture of something like the desired outcome of a problem". The usefulness of visualization was limited initially as the power of graphics had not been mastered however; computer graphics and the development of animation have improved its usefulness.

Visualization can also be described as a cognitive process which involves creating mental images to facilitate understanding of items with no direct significance (Spence, 2001). Mackinlay and his co-authors, Stuart and Shneiderman (1999) suggests that more research should be carried out towards incorporating visualization into the daily management and use of the web since it makes information easier to understand and supports decision making.

### 2.5.1 The visualization process

The visualization process is described by Ware (2000) as involving four stages; collection and storage of data, pre-processing to change data into an understandable form, display of hardware and software and human perceptual and cognitive system. The feedback loops (fig 2.16) illustrate data gathering, data manipulation and data exploration. Ware (2000) and MacEachren (1995) agree that visualization is a cognitive action which engages the human brains and produces mental models.



Figure 2. 16: Visualization process.

Source: adapted from Colin Ware (2000)

### 2.5.2 Types of Visualization

### Information Visualization:

The information visualization field came alive as a result of research in human-computer interaction, which allows visualization to be applied in more general ways to education and business. It is the study of interactive visual representations of data and how they support people acquire knowledge and understand data. It has become an essential element in scientific research, digital libraries, data mining, financial data analysis, market studies, manufacturing production control, and drug discoveries (Bederson and Shneiderman, 2003). Information visualization focuses on creating different methods to pass on theoretical information in perceptive ways (Thomas and Cook, 2005). For example the digital moving pictures used to demonstrate weather reports.

On television broadcasts, technical visualizations are used to show animated productions of natural and unexpected disasters. Some examples are images generated by computers illustrating real spacecrafts in action. These changing forms of visualization that include animation and timelines can make learning easier.

### Software Visualization:

Software visualization aims to support users understand the structure, algorithms and the analysis of software systems as well as their anomalies. Information is represented in visual forms of static or animated (Diehl, 2002, Diehl, 2007), 2-D and 3-D (Marcus, Feng and Maletic, 2003). Based on their size, (Staples and Bieman, 1999), Structure, (Stasko, Brown, and Price, 1997) behaviour (Keim, 2002), or history (Soukup, 2002).

### Scientific visualization

Scientific visualization is a subdivision of computer science, its purpose is to illustrate scientific data in a graphical manner in order to help scientists understand and gather insight from scientific data. Friendly, (2008) says it is fundamentally concerned with the visualization of three dimensional events and is used in medical, architectural, meteorological, and other disciplines.

Information Graphics (Info graphics):

Here information, data or knowledge is represented using visual graphics. The graphics present complex information in a clear manner making it easier to understand for example the map of the city centre, Edinburgh. Though complex, it has been presented in a simple way that a first time visitor can understand it. Maps, traffic signs, technical writing, manuals, weather, site plans, books, statues, icons and so on can be presented in this manner. Information that would have gone unnoticed and therefore not read in form of text when presented as a sign becomes significant.

Interactive Visualization:

This is a branch of graphic visualization in Computer science which studies how humans interact with computers to create graphic illustrations of information and how this process can consistently be improved until it is made more efficient. Generally, it is considered to be a soft real-time task. Some types include virtual reality where information is represented in three dimensions and collaborative visualization where multiple people who are often apart interact with one computer, sharing ideas. An interactive visualization satisfies two criteria: Human Input and Response Time.

The human must have control of some aspect of the visual representation of information; it requires physical devices like keyboards, touchpad, graphics tablet, and trackballs. A range of inputs provided by humans include:

- They can pick some part of an existing visual representation
- They can locate their point of interest
- They can make a choice from a list of options.
- They can determine the value by inputting a number.
- They can write by inputting text.

Response Time involves incorporating changes made by people into the visualization in a timely manner. Interactive visualization works best with systems that provide feedback to users within seconds of input. This response time is a difficult target to meet and several approaches are still being explored to this end. Schneiderman (1996) gives a guide called the visual information seeking mantra illustrating how data should be presented on screen to speed up visual exploration as “Overview first, zoom/filter, details on demand”. The existing datasets are huge and are difficult to create an overview of the visualization to contain relevant patterns. Visualizations do not tell users what steps to take during their investigation which results in techniques like zooming and filtering being underutilized. From the perspective of visual analytics, visual information seeking mantra could be extended to “Analyse first, display the significant, zoom/filter, analyse further, details on demand” (Keim, Mannsman, Schneidewind and Zeigler 2006). This shows that using just a visual metaphor is inadequate to search for and display data. Instead the data needs to be analysed with respect to its level of significance, illustrating the most important parts of the data and supplying details on demand simultaneously.



When people visualise data, their goal is to find information, understand this information and gain knowledge. The process of obtaining information from a body of data is the challenge information user's face which led to the creation of visualization. Moving from one point to another on an information space featuring visualizations is a challenge, navigation aims to address Spence (2001). The earlier visualizations available did not offer the opportunity for the data to be changed to create a more valuable view of it which is what interactive visualizations offer. Data can be rearranged interactively using visualization tools (Spence, 2001). Spence holds the opinion that by rearranging the manner in which data is presented, further insight into the data in question can be gained. In addition, the characteristic of interactivity that improves the tools effectiveness refers to navigation. Looking at two separate interfaces where the more effective of both allows easy movement while movement in other is slight.

### Data visualization

Data visualization entails the design and study of the visual representation of data. Its main aim is to communicate information in a clear and effective manner with the aid of graphics (Friedman, 2008). Data visualization is utilized to display connections, articles and resources, data, news, websites, mindmaps, tools and services (Modern Approaches, 2007). In Data visualization, bar charts, steamgraphs, treemaps, gaantt charts and scatterplots are used to represent data and tell stories. Some examples are illustrated below. The diagram (fig 2.17) shows a data visualization of Florence Nightingale's map which tells the story about the causes of mortality in the army in the east during the Crimea war from April 1855 to March 1856 where large numbers of

deaths were recorded from cholera a preventable disease. The blue wedges represent deaths that occurred from preventable diseases; the red areas represent death from wounds while the black wedges show deaths from all other causes.

Illustration is a type of data visualization and usually involves work of Art like drawings, paintings, photographs and other art works created to pass across information by visual or graphical representation to explain specific parts and convey certain information through visual means, this enhances the user’s ability to visualise and understand the technique (Ivan Viola and Meister E. Groller 2005).

## War Mortality – Florence Nightingale

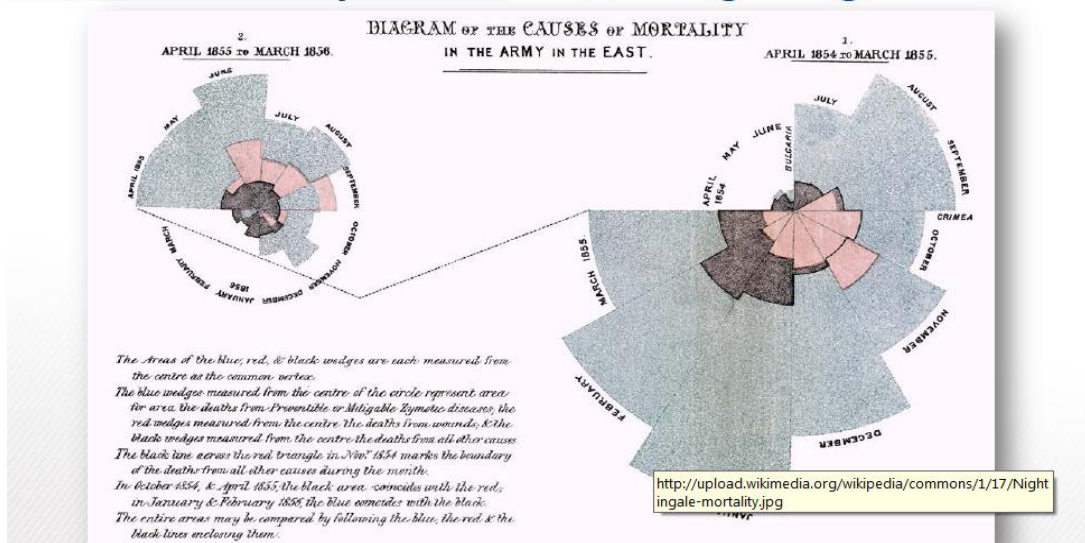


Figure 2. 17: Florence Nightingale’s roses showing an outbreak of cholera

Source: Adapted from Wikipedia, the free Encyclopaedia.

Tufte (1983) outlines the benefits of data visualisation as 1. It shows data. 2. It allows large set of numbers to be presented in a small space. 3. It encourages the viewer to think about the real meaning communicated by the visualization instead of the method

used. 4. It allows large data sets to be put together in a coherent manner. 5. It encourages the eye to make comparisons between different pieces of data and 6. It reveals data at various levels of detail, from a wide overview to the minute structure.

The biography chart illustrates the groups of biographies and their relationship to significant periods in human history of the Greeks, Romans and the Enlightenment and records the rise and fall of empires and the unique people, who defined them.

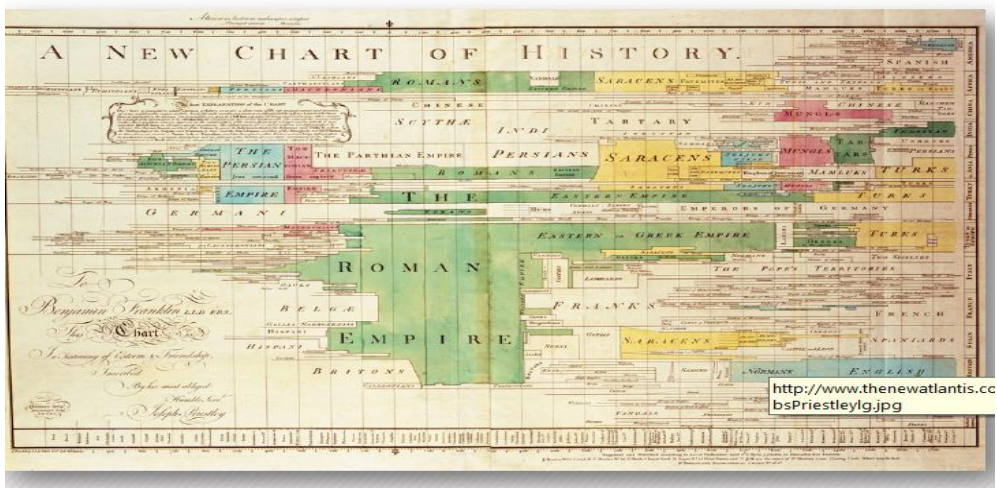


Figure 2. 18: Biography chart illustrating main world civilizations

Source: Adapted from Priestly J. (1975).

The information presented in graphical forms are easier to visualise and understand, it facilitates comprehension of complicated information and people can swiftly take in the meaning of data and make adequate decisions. It compels us to notice unexpected patterns. Information specialists utilize visualizations and other tools to find relationships, significant information and precisely what the data has to say. The data illustrated above brings life into the information represented; the framework represents data simply and tells the story.

The scatterplot (fig 2.19) allows the visualization of multivariate data which displays two or more variables for a set of data.

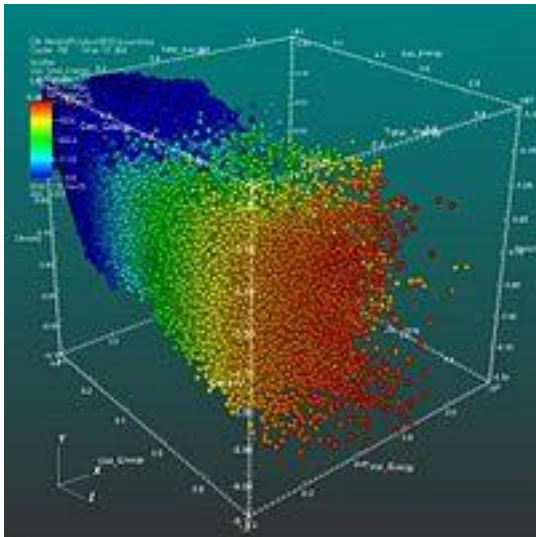


Figure 2. 19: Diagram showing a scatterplot

Source: Adapted from Wikipedia the free encyclopaedia

A number of visualizations are created to solve specific problems for a category of people and have a life expectancy. Some examples of these visualization problem types include; data discovery, data quality, storytelling, dashboards and tools, trends and predicative.

Trends and Predicative visualizations have a wider audience and can be used as storytelling or dashboard and tools. They are employed to illustrate combined results of decisions made or activities carried out including those out of the control of the audience. It requires visual tools to display sufficient relevant data to enable the audience gain thorough knowledge of the trends and predictions (Sacolick, 2009).

### **2.5.3. Data Analytics**

Data navigation is defined by Dourish and Chalmers (1994) as the means by which users can describe movement between pieces of information or the routes they follow on the information space. It is moving from one data to another on the information space.

Data analytics is a technique that involves studying unprocessed data with the aim of retrieving required information, it utilizes expressive models to understand data and make it possible to detect important patterns of data. It is a valuable way of communicating knowledge during data visualization and directing actions taken and decisions made. Data analytics in general is more interested in the methodology applied to the field of data visualization.

A practical example of the use of data analytics can be observed in most higher institutions of learning where data analytics application are applied to give a predictive perspective of future challenges for the students and institutions. A case in point is where research suggests that students who move directly from high school to university stand a higher chance of graduating than students who do not. As a result the first year is the object of interest of many schools and the application of data analytics can guide the use of economic resources and aid to identify where to invest resources in order to achieve the best results (Educause, 2010).

Data analytics brings together massive data sets, statistical tools and predictive modelling. It is used by some organizations to turn information into knowledge and direct the actions taken ultimately in order to deliver value. The focus of this research

is concerned with management information systems such as coordinate multiple views (CMVs). The research interest lies in developing CMVs to support visual analytics: 'the science of reasoning facilitated by interactive visual interfaces' (Thomas and Cook, 2005).

The area of interest of this research is concerned with how CMVs of a data set can be designed to support visual analytics in the specific field of information management. These are extremely common circumstances where the aim may be managing a research institute, or a service industry (such as care in the community services), or a garden centre. The centre of attention of conventional database design is retrieving information through the use of queries on the database using a query language such as the structured query language, (SQL), (Elmasri and Navathe, 2007).

Analytics focuses more on identifying patterns, spotting new relationships between data and navigating the database in novel interesting ways, an objective that can be accomplished through a CMV developed to support successful navigation.

Previous studies have discussed some important areas concerning CMV development such as view generation (Roberts (2007) and the issue of tasks (Shneiderman & Aris, 2006), nevertheless, the issue of the type of data to be presented and have not been addressed. The challenge of how to accumulate data may necessitate designers to exploit the benefits metadata can add to creating CMV's.

Metadata can create additional information that can help the user gain knowledge of the information space, interact with understanding, navigate, visualise, and complete tasks more easily. It can be programmed to give details about how the information

space was set up and instructions on how to use it. It can also provide supplementary information about the data such as what information the different views hold, what views are coordinated and how the information space is organized. This type of information is described as useful information that requires visualization such as visualizing the conceptual interactive structure of a system (Weaver, 2004). One example uses arrows with notes written at the top to point to the linked views of the visualization (Roberts, 2007).

## ***2.6 Coordinate multiple views (CMVs)***

In HCI CMV sometimes referred to as MCV is an exploratory visualization technique that is useful to information seekers in the area of data exploration. According to Roberts (2007) CMV is a developing area still requiring more research. It is distinct in structure as a result of its multiple views. It represents data in multiple windows and operations on the views are coordinated hence it can be referred to as multiple coordinated views. These multiple views allow users to compare data easily as views are placed side by side (Convertino, Chen, Ryu & North, 2003).

CMV is useful to users who need to interact with complex data with the aim of discovering facts. It affords the user the opportunity to evaluate several scenarios and compare visualizations generated from many data sets. Users can collect, retrieve data and put these data together to create new information. CMVs are useful in several domains for analytical investigations, it allows several professionals look at the same data, compare this data, discuss their findings and reach conclusions.

CMVs are designed to make several windows, parameter choices and information available to the user. It could help the user interact with and understand their tasks better when implemented with helpful techniques (Roberts, 2007). They have the prospect of promoting analysis of management information systems and supporting visual analytics which is of interest to this work. CMVs will be discussed in more detail in the subsequent chapter.

### **2.6.1 Visual Analytics**

Visual Analytics is an outcome of the field of information visualization and cuts across multiple fields like cognitive and decision science, knowledge management and statistical analysis. Visual analytics is defined by Thomas and Cook (2006) as ‘the science of analytical reasoning facilitated by interactive visual interfaces’. A more recent definition by Keim, Kohlhammer and Mannsmann (2010) claim that “Visual analytics combines automated analysis techniques with interactive visualizations for an effective understanding, reasoning and decision making on the basis of very large and complex datasets”. Visual Analytics allows the creation of conceptual visual metaphors that work together with human information dialogues on large dynamic information spaces. This form of interaction aided by the wide-band visual interface to the mind promotes the location of the likely and discovery of the unlikely (Wong and Thomas, 2004). In addition it allows people synthesize information and obtain insight from enormous, confusing and contradictory data. The visual analytics procedure merges mechanical and visual analysis process with a firm coupling during human interaction so as to derive understanding from data. A general idea of the process is shown by the steps (illustrated by oval structures) and their evolutions (illustrated by arrows) in the visual analytics procedure (Fig 2.12).



The goal of the visual analytics procedure is to create a firm coupling between mechanical analysis processes and interactive visual representations. An important attribute of the visual analytics process is that it creates interaction between data, visualizations, models about the data, and the users to gain knowledge. Visual analytics presents complicated information in a visual form making it easier for the human brain to comprehend thus making analytical reasoning possible

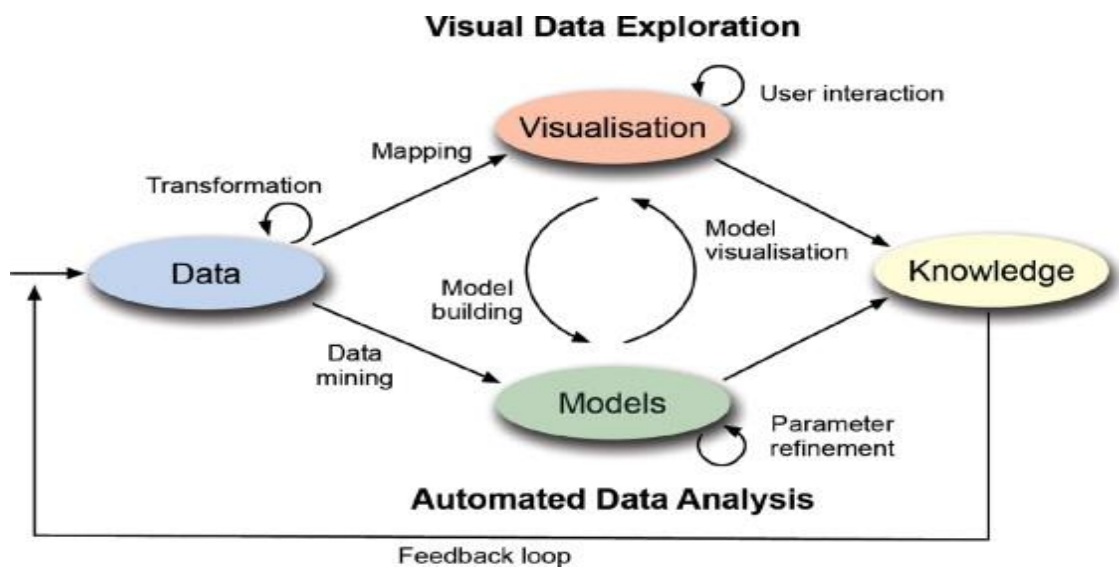


Figure 2. 20: Diagram of visual data exploration

Source: adapted from solving problems with visual analytics, science direct article.

### Interactive Visual Interfaces

Interactive visual interfaces are characterised by visual representations which give people access to data and support human acquisition of knowledge. Visual representations make it possible to solve complex assignments as it allows people observe significant parts of data by changing it to a discernible form that draws attention to relevant features. In addition it enables people make out more easily

unknown or fragile relationships between assets. Tufte (1983) illustrates quantitative data plotted in a time series showing how visual representations express relevant relationships between several sets of data. Nonetheless, these relationships will be difficult to identify on information spaces where navigation is not adequately supported.

Interactive visual interfaces make use of interactive techniques described by Foley, Van Dam, Feiner and Hughes (1995) as the use of an output device to carry out a general task by users on a computer. Interaction techniques such as navigation are elements that enable users move through data, communicate with the data and interpret representations directly or indirectly. Visual analytics is used by many subject areas and covers the following areas:

Analytical reasoning techniques, which enhance the user's ability to acquire deep understanding, support appraisal, development and decision making.

Visual representations and interaction techniques that make use of the human eyes ability to perform many tasks to help users see discover and obtain vast knowledge concurrently.

Data representation and alteration change various incompatible and dynamic data to ways that sustain visualization and analysis. These techniques make inventions, presentations and the spread of analytical results and exchange of information possible in an acceptable manner to various people (Cook and Thomas, 2005).

Analytical reasoning, visual representations and interactive techniques can be advanced by successful navigation. An information space that supports successful

navigation has the potential of producing a positive impact on people's ability to visualise effectively and retrieve adequate information. In order to optimise the visualization technique, vital aspects of navigation should be addressed.

### **2.6.2 Visualization Techniques**

Visualization techniques are interactive tools that support people on information spaces to interact with and visualise data. Some of these techniques include brushing and linking, zooming and panning, animation and distortion based techniques.

#### **Search Function**

The search function is one of the most important features of an information space. It helps find certain aspects of the data, determined by users input during information retrieval. Grun et al (2006) point out users can confine their search to specific aspects of the data such as an area of interest.

#### **Brushing and Linking**

In the technique termed brushing and linking, an object selected in one view of the interface causes related objects to be highlighted simultaneously in another view. Brushing changes the colour of the selected element (Ross, Morrison and Chalmers, 2004) and helps user's observe separate parts of the system between multiform (Lawrence, Cook, Hofmann and Wurtele, 2006). A multiform describes data that is displayed in two or more different forms (Robert, 2007). Brushing is widely used in interfaces that involve text analysis (Hearst, 2009). It allows users make comparisons and go on to obtain the fine details of the piece of information in question.

## Maps

Maps are figurative representations of elements that exist within an information space such as objects, regions and themes and how they are related. Some maps are two dimensional in nature while others are three dimensional in nature and can be interactive and dynamic. They may be used to represent virtual or physical information spaces.

## Tables

Tables allow data to be organized into rows and columns in order to make it easier to understand and interpret. A row in a table is referred to as a record while a column is usually assigned a name and referred to as a field. Tables are used as communicating tools across many disciplines but they differ in structure, flexibility, representation and function (Morgan, 2004). Tables may be simple or multidimensional in nature.

## Charts

Charts are techniques that have been in use for decades; however, new interactive charts have been developed to make information easier to manage. They include bar charts, pie charts, histogram and scatterplots. For instance, a scatterplot uses coordinates to display two variables for a set of data. Scatterplots display data as a collection of points; each point has the value of one specific variable which determines its position on either the horizontal or vertical axis.

## Graphs

Graphs are used to present structural information in diagrammatic forms. They include tree diagrams, network diagrams and flowchart diagrams. They are usually created as

node-links and used to communicate information across several disciplines. Graphs reveal data; they allow information to be represented in simple ways and facilitate comparison of information.

### **2.6.3 Visualization, Navigation and Metadata**

Visualization aims to change the way content is manipulated. Research using digital libraries can be made easier if a central metadata repository can be accessed. Metadata enhances the exploratory use of visual interfaces and plays an indispensable role in the acquisition of data for problem solving. It can provide users with enough information to facilitate navigation and data visualization. It could also help them decide how well the available datasets meet the intended use. The interactive exploration of metadata can be utilized in a way that different characteristics of datasets studied simultaneously can support users in their navigation and visualization of data. As metadata becomes more useful to information spaces, the ability to manage it to the advantage of users becomes increasingly important. Visualization of the metadata of an existing information system is a valuable tool for information management.

On semantic information spaces, the use of Metadata clearly indicates what data means and can provide computers with substantial information to handle such data. Large information spaces contain enormous amounts of data and metadata and are more complex in nature. The question is how will such information flow be managed? Two complementary solutions have been proffered and again at this point, navigation and visualization techniques come to play.

The first solution will see machines turned into nonhuman type web users such that they understand the meaning of data on the web and how they should be used without any human input. This is the main purpose of the web to be developed.

The second solution sees the web being made more functional for users by presenting data and metadata in a comprehensible visual form. Navigation is vital to accessing and moving through large complex information spaces, on the other hand Visualization is central to understanding and handling large complex information resources. Visualization and navigation are critical to the semantic web because of its complex nature.

This research will explore the use of a novel type of metadata, termed summary data to develop a navigable CMV. The CMV will afford users the opportunity to find their way around with adequate support, visualise data effectively and retrieve adequate information easily and quickly. The design of visual interfaces for information retrieval is a challenging area in practical information spaces. Navigation and Information retrieval in current information systems is difficult, the large volumes of data, their lack of structure and multidimensionality make these systems challenging (Spence 2001).The development of a variety of tools for visualization also presents an issue. The quality and type of metadata provided brings about a noticeable effect on navigation and visualization.

Visualization, transforms data to a visible form which clearly highlights its important features and makes it easier for users (Thomas and Cook 2005, Thomas and Cook, K. Ed. 2005). It also improves accessibility, supports user tasks such as analysis,

search and exploration. It may be a superior alternative for the future web than machines. Machines only give back what was programmed into them but humans can think, visualise and be creative. The computer visualization technique of interest to this research is the CMV.

## ***2.7 Conclusion***

From the literature review, it is essential to recognize that navigation plays a vital role in visualising and retrieving data. Navigation is an important aspect of interaction, visualization and information retrieval. Focusing on creating several visualizations to address these issues will not bring about the level of information management required for digital assets. The concept of navigation for information space moves from designing for people who live outside their information spaces to designing for people who live in their information spaces. The main problem of designing systems is coping with individuals; the design needs to provide for individuals with different needs and those with different goals. It also needs to accommodate users whose goals are not properly developed, or have different experiences and are familiar with different domains (Benyon, 2006). According to Spence (1999), the problem of developing interaction design to support navigation is becoming more crucial and challenging. Designers need to create information spaces that support the general navigation needs of people to enable them achieve their objectives on these spaces. It also looks at incorporating some design aspects of the physical world into the virtual world in order to improve navigation. In the subsequent chapter, the concept of navigation of a CMV of a management information system will be discussed in detail.

## **Chapter 3 Coordinate multiple views of Visualization**

### ***3.1 Introduction***

Chapter two discusses the background of this research. It gives an overview of navigation, navigation models and information spaces. This chapter discusses CMVs in general which dates back to the late 1990s, when a CHI workshop led on to papers in the Advanced Visual Interfaces (AVI) conference in 2000. Its significance is evident in the five specialist conferences held from 2003 to 2007.

Chapter 3 offers a detailed description of the IIDI visualisation, employed for the empirical studies carried out in this research. It describes the redesigned IIDI visualization interface, highlighting its features in a comparative style to the initial IIDI visualization interface. The redesigned IIDI interface draws from the concept of navigation of information spaces which focuses on creating navigable information spaces.

### ***3.2 Coordinate Multiple Views***

Coordinate multiple views of visualizations (CMVs) sometimes referred to as multiple coordinate views (MCVs). During an AVI conference in 2000, it was defined as a visualization where more than one view was provided to 'support the investigation of a single conceptual entity' (Baldonado, Woodruff and Kutchinsky, 2000). They pointed out that several forms exist in many different domains. A high level of taxonomy was made available by North and Shneiderman (1999), it was based on the whether the emphasis was on selecting items or navigating views and whether the different views represent similar or dissimilar information. They went on to highlight significant



features of CMVs and the advantage it offers to user' performance, the new relationships it points to, and the help it provides to enable manage complicated information through interaction.

During the same conference, (North and Shneiderman, 2000) proposed that the relational data model (Codd, 1970) provided a good foundation for coordination, likening the object of a visualization with a tuple in a relation. Different CMVs could be automatically 'snapped together' depending on the relationships between relations as either one-to-one, or one-to-many. (Boukhelifa, Roberts, & Rodgers, 2003) however argue that it is limiting to use only the relational model. They go on to develop their view of coordination focusing on how the different views are coordinated.

CMV is a visualization method that is tailored to help users explore their data. It is believed that users understand data better if they interact with the available information and look at it from different perspectives. On the one hand, users want to view complicated data, explore the information space and discover some hidden facts. This kind of studies may need the user to look at various circumstances and compare visualizations developed from various datasets, to collect and retrieve the data. They may also need to pull together data from several datasets to create new information. On the other hand different information experts may be studying the same data to evaluate and discuss patterns in order to make decisions. CMVs provide a design layout that allows user to consider data from various angles and can control their interactions and harmonize operations between views. This will lead to users being able to observe new relationships and information from the data. It will also draw the attention of users who are too conversant with visualization methods and could easily

overlook the information the data offers. The underlying theory is to provide an interface for the user to communicate with the data while its objective is to discover information and understand the vast separate datasets. Many people communicate with interactive visualizations to find information, discover irregularities, compare and contrast data (Thomas and Cook, 2005). Furthermore, users want to investigate different circumstances and develop theories and examine them. These can be achieved by developing an extremely interactive visualization space which is the aim of CMVs to enable the discovery of knowledge where user's interaction with data brings about inventing a problem and finding a solution to it simultaneously (Spence, 2001).

The CMV interface is designed to provide the user with many views to walk through data, various parameters to select from and a variety of information to reason through. The relevance of CMVs is seen in its usefulness in bringing different perspectives of data for analysis, analysing huge datasets and developing bits of information. Users require CMV interfaces that are easy to navigate, allow swift visualization of information and information review. CMVs face the challenge of providing strategies which enable people collect, compare and manage data efficiently in order to produce meaningful information. Designing CMVs whose component parts work together successfully and can be extended will improve its purpose (Roberts, 2007).

The process of analysing multidimensional data requires proper investigation of relationships existing across each dimension. Coordinated multiple view strategies are becoming more popular in visual analysis tools because they use easy interactions to explain complicated multidimensional queries. However developing these tools are

challenging as certain requirements must be met like plotting specific data structures of an area to interdependent data and visual concepts required to show certain trends in data (Weaver, 2004).

CMVs have drawn the attention of the information visualization (infovis), Human computer interaction (HCI), advanced visual interfaces (AVI) and other communities. The infovis has nonetheless focused on CMVs from the perspective of drawing graphs and other mechanical methods for generating striking visualizations, which is insufficient. The benefits of the CMV technique have the potential to advance visual analytics which is enhanced by interactive visual interfaces (Wong & Thomas, 2004). However despite the extent work in this area, there is amazingly little discussion about *which* data should be presented in the visualisation. It is generally assumed that the visualisation designer will discover the tasks that users are trying to do and will design a CMV to support those tasks (Baldonado et al., 2000). There are many examples of this, but little generalisation.(Chen, 2005) summarises these problems, identifying specifically a lack of evaluative methodologies for usability in CMVs, a need to modify design thinking from structural to dynamic, and an in general lack of strong empirical work in the area. We seek to rectify this by proposing simple, but powerful design heuristics for a particular class of CMVs and evaluating these through adequate empirical studies.

Stated plainly, the initial design heuristic of this research is 'providing data *and metadata* for people using a CMV will improve their ability to navigate a data set'.

### 3.2.1 Designing CMVs

CMVs are difficult to design due to the discreet interactions existing among the several aspects of the design space. Developers are faced with the challenge of making various design decisions which include shaping the layout and building complicated coordination mechanisms (Baldonado, et al., 2000). A number of visualization researchers have presented valuable guidelines for general user interface design (Shneiderman, 1987, 1996, Nielsen, 1994), and for multimodal systems (Coutaz, Nigay, Salber, Blandford, May & Young, 1995). However, the many general guidelines in existence are not sufficient to help designers to choose the one most relevant to creating CMVs which prompted the decision by Baldonado and his colleagues to present specific guidelines for using multiple views (Baldonado et al, 2000), however, the subject of the type of data to be presented in CMVs and design strategies to support navigation were not addressed.

The interest of this research is to design CMVs to incorporate the different aspects of navigation which (Benyon & Mival, 2010) characterised as object identification, wayfinding and exploration and what (Spence, 2002) terms the process where people can gain knowledge of an information space, are able to interpret this knowledge effectively in order to retrieve adequate information.

The focus here is on the use of CMVs for MIS, rather than scientific data, geographical data or pictorial databases. The research is interested in the analysis of management data and in developing CMVs to support visual analytics: 'the science of analytical reasoning facilitated by interactive visual interfaces' (Thomas & Cook, 2005). It is a hugely ambitious and comprehensive attempt to create a new science of analytical

reasoning and involves theories of reasoning, representations, understanding and interaction at the many different levels of detail that establishing such a new science requires.

The area of this work significant to us is in how CMVs of a data set can be developed to support visual analytics, particularly the area of management information such as a research institute or a hospital management system. Analytics focuses on trends, seeing new relationships between objects and navigating the database in new and interesting ways rather than the approach of the traditional database which focuses on data retrieval.

In his state-of-the-art review of CMVs, (Roberts, 2007) discusses “view generation” as an important part of the developing CMVs: designers need to consider the form of the visualization, how to map information to the form, how to abstract and aggregate data, and how the user interacts with the data. However, he does not consider what data or relationship to present in the first place as well.

As pointed out earlier (p.65), the issue of how to aggregate the data requires designers to consider the development and use of metadata, which traditionally is seen as a different type of data from the underlying database.

For instance, descriptive metadata identified by Wodke (2003) is used as hash tags in social media. A picture on Instagram is described by a hash tag (a metadata tag) such as ‘sunset’ allowing people to search and find all the pictures of sunsets. Combining with other tags such as the location ‘Edinburgh’ will find pictures of sunsets in

Edinburgh. So, although this is classified as metadata, there is no intrinsic difference between the data and the metadata. The data about the picture (sunset in Edinburgh) is the metadata that will help people find the picture they want. This means that one person's data is another person's metadata. Most recently this distinction has been made in the context of the National Security Agency (NSA) in the USA accessing the metadata about e-mails (Landau, 2014). This includes the e-mail addressee, where it came from, the subject, and information concerning when it was sent and received. Thus metadata about an e-mail becomes very useful data for the analytics in which the NSA is engaged. Plainly patterns of metadata allow inference and conclusion and also identification of data for further scrutiny with an enhanced probability of relevance.

There are two key problems that arise for the design of visual analytics. Baldonado and his associates say that designers 'necessarily' begin by establishing a clear view of the user's task, when, of course, different people will have very different views, and hence different tasks, on any data set (Baldonado, et al., 2000).

Secondly, equating a tuple (a row) of a relational database with an object in the data set mixes up the data and metadata, as a relation provides metadata for the object that is the primary key of that relation.

The issue of user tasks is critical. Shneiderman & Aris, (2006) list a number of tasks that users of visualizations will want to do, such as 'count the number of nodes and links', 'count the degree (the number of links) for each node', 'find the distance from one node to another (count the number of steps from one node to another)', before finally concluding that 'there are an unlimited number of tasks that could be defined'. Collins and Carpendale (n.d) discuss the relationships between data, relations and

visualizations, giving CMVs as one example of a multiple view. Step 1 of their method for creating visualizations is 'choose a relationship', without saying which relationships might be useful to choose. Javed and Elmqvist (2012), in their review of choosing and using multiple, coordinated views of a data set, look for recurring design patterns. They discuss the importance of 1:1 and 1: M links between relations but, again, stop short of saying what data needs to be available for people to undertake which sorts of tasks. So, whilst many writers describe the different stages of developing CMVs such as data definition, layout strategy, rendering choices, and so on (Roberts, 2007; Shneiderman & Aris, 2006), no-one really talks about the general ideas of what data is needed by whom to do what. To rectify this we take a different view of visualizations and see it in terms of navigation, we propose a simple but well-built design procedure for a distinct class of CMVs. The design will provide data alongside its metadata in order to enhance navigation of a data set and information retrieval.

### **3.2.2 CMV Visualization Techniques**

Visualization forms can be described as the structure of the visualization that gives it a distinctive character. They could be in form of maps, networks, charts and graphs such as scatterplots, bar charts, line graphs and parallel coordinate plots. The IIDI visualization utilizes barcharts, nodelinks, a scatterplot, matrix, tag cloud and parallel coordinate plots. According to Roberts (2007) parallel coordinate plots form the standard element of most CMV systems.

Visualization techniques are interactive tools that support people on information spaces to interact with and visualise data. Some of these techniques include brushing and linking, zooming and panning, animation and distortion based techniques. CMV

techniques such as interaction and manipulation are utilized by most CMVs including the IIDI visualization in order to support people explore data sets.

### Interaction and Manipulation

Interaction is a vital strategy used by the CMVs and is a two way communication between the user and the system. Users can modify the system, navigate through information, reorder the windows, filter the data and select something of interest (Andrienko and Andrienko, 2003). It offers two types of manipulation; direct and indirect.

### Direct Manipulation

This technique supports the user filter through the data and select elements from the visualization directly. Brushing is the main type of direct manipulation and all modern CMVs offer some form of brushing (Roberts, 2007). Scatterplots were used to carry out the initial work on brushing (Carr, Littlefield and Nicholson, 1986., Becker and Cleveland 1987).

### Brushing and Linking

In this technique, an object selected in one view of the interface causes related objects to be highlighted simultaneously in another view. Brushing changes the colour of the selected element (Ross, Morrison & Chalmers, 2004) and helps user's observe separate parts of the system between multiform (Lawrence, Cook, Hofmann & Wurtele, 2006). A multiform describes data that is displayed in two or more different forms (Robert, 2007). Brushing is widely used in interfaces that involve text analysis (Hearst, 2007). This technique is used on the IIDI data visualization and according to



Hearst; the technique acts as a support to people to improve their understanding and enable them discover interesting relationships within the data.

Other direct manipulations techniques include widgets and manipulators fastened to objects to alter their properties like making available handles used to manage parameters directly (Andrienko & Andrienko, 2003, Chuah, Roth, Mattis & Kolojejchick, 1995, Ericson, Johansson & Cooper, 2005).

### Indirect Manipulation

Indirect manipulations are dynamic queries (Shneiderman, 1990) that offer the user a platform to interact with sliders, menus and buttons to enable them filter available data and alter the manner in which data is displayed. The sliders symbolize a way users can interact with and decide on what they want to visualise and how much limit they want to place on the data they want to view. An example of a slider is the range slider (Shimabukuro, Flores, Oliveira & Levkowitzet, 2004) used in a number of CMVs. The IIDI data visualization utilizes buttons as an indirect manipulation strategy.

### Manipulation Attributes

A number of attributes that can be manipulated exist on CMV interfaces like the IIDI visualization. They include selection, label, sort, filter and colour, transparency and size. These attributes can be explained as designs that define characteristics of any object found on the interface. Selection allows people choose a data item; data items selected in one view are highlighted in other views. This is an aspect of filtering as it allows people focus on specific parts of the data. The label makes the multiview aspect of the visualization possible. It allows different views of the data to be displayed and decides which perspective of data the view will present. The sort attribute decides the

order the visual displays of data will appear. Colour, transparency and size are visual properties. They make it possible for people to interpret and understand differences between the objects on the interface.

### Multiple Views

The phrase multiple views describe an occasion where data is illustrated in multiple windows. When different representations of data are placed in successive windows and operations carried out on the data are coordinated, they are called coordinated multiple views (Roberts, 2007). CMVs could be dual views using two side by side views or they could have four or more views like the IIDI data visualization. Many types of visual forms exist each having its shortcomings and benefits.

### Multiple view Functions

Multiple view functions are tools employed by interactive systems to create multiple views (Roberts, 2007). They include buttons, menus like CommonGIS (Andrienko & Andrienko, 1999), snap (North and Shneiderman, 2000) and Mondrien (Theus, 2002). Some systems make dynamic queries available to enable users explore interactively changing display parameters when required such as the Improvise (Weaver, 2004) and Ross and Chalmers (2003) visual workspace. Other systems use a modular method and are described as Modular Visualization Environments (MVE's). Their functions are usually defined beforehand and can be connected to form a visual plan which offers a suitable method to extend and create the visualization (Ross, Morrison & Chalmers, 2004; Upson, Faulhaber, kamins, Schlegel, Laidlaw, Vroom, Gurwitz and Vandam, 1989 and Walton, 1996).

### ***3.3 The IIDI Data: A CMV of a Research Centre's Database***

The Institute of Informatics and Digital Innovation (IIDI) is one of the nine institutes for research and innovation at Edinburgh Napier University, it is situated in the Merchiston Campus and its chief aim is to structure the digital future. It acts as a support to organisations to help them understand innovative digital challenges and openings that crop up. The institute works across various aspects of computer and information technology such as mobile networks, data intensive applications which involve intelligent procedures and filtering. It goes further to look at interactions between people and technology and the impact this creates on the society.

The Institute has five centres; the centre for interactive design (CID) whose major focus is to design new methods of interaction. The centre for emergent computing (CEC) investigates the use of biologically and socially inspired systems to solve industrial and commercial challenges. The centre for information and software systems (CISS) aims to find new and useful solutions for software and information systems. The centre for distributed computing, network and security (CDCS) focuses on security issues in areas such as e-health, improving mobility for devices and group communications. The centre for social informatics (CSI) studies the interdisciplinary creation and use of information and communication technologies (ICT's) that is concerned with cultural and institutional trends.

The members of the institute comprise staff, research students, associates, an advisory board and past members. Its areas of expertise include information visualization, networks, future interactions, business and technology strategy, optimisation, security and cybercrime, software engineering and data intensive

systems. The institute provides opportunities for personal, professional and business development tailored to meet the needs of people working in the public, business and industrial sectors ([www.iidi.napier.ac.uk/c/groups/site/iidi](http://www.iidi.napier.ac.uk/c/groups/site/iidi)).

### **3.3.1 The IIDI Data Visualization**

The IIDI data visualization is a management information system (MIS) created by a database designer to provide information about the individuals and groups within the institute and facilitate effective and efficient evaluation, analysis, design, implementation and management. It consists of names of past and present members, research students and a board of directors within the institute. It gives details about the centres they belong to, individual and joint publications, keywords developed and grants received within the centre.

The IIDI data is a large set of diverse data and was chosen for this study because it is a live data with coordinate multiple views and accessible for this study. In addition, where the need to effect changes crop up, it can be carried out and the success recorded will create a live impact on the users. The IIDI visualization is illustrated below.

### **3.3.2 Interface Events on the IIDI Data CMV**

Interface events are described by this research as the actions that people take to enable them carry out navigation activities such as object identification, wayfinding and exploration. These events are dynamic and they occur when people move from one view to another on CMVs in order to search, browse, explore, identify objects,

create internal models, and find their way to desired destinations. The specific events that occur on the IIDI Data CMV Interfaces are outlined below:

- Selected Matrix (SLM) - when a matrix is selected, it indicates a choice has been made and detailed information is being retrieved.
- Sorted Matrix (SM) - when the centre, publications, grants or label button is selected in order to access the required window.
- Selected Nodelink (SN) – when a nodelink is selected, like the SLM, it signifies a decision has been taken and specific information is being retrieved.
- Mouseover Matrix (MM) – when a mouse is placed over a matrix, it points to the possibility that information about publications, centres and grants is being sought.
- Mouseover Nodelink (MN) – when a mouse is placed over a nodelink, it suggest that a participant is searching for information about an individual, or a centre
- Mouseover Barchart (MB) – when a mouse is placed over a bar chart. It shows participants are searching for details about publications.
- Filtered Year Barchart (FYB) – when a mouse is placed over the year on the bar chart to retrieve publication date.
- Filtered Pu Barchart (FPB) – When a mouse is placed over the publications on the bar chart to find the total number of a specific publication.
- Reset Selections (RS) – when the clear selection icon is selected. This action returns the interface to its default state.
- Task Start All (TSA) – when the reset task timer icon is selected to show the participant wants to commence a new task.
- CMA – when the upper centre view is selected to retrieve information on centres.

### 3.3.3 The IIDI User Interface A

A CMV representation of the IIDI data is shown in Figure 3.1 (which is presented in low resolution so as not to identify individuals). This is a genuine example where the designer has taken a number of decisions about what to represent and how to represent it. The CMV shows the people in the institute in the top left window, colour-coded by centre, with lines showing the co-published relationship and the thickness of the line showing the number of co-publications. The bottom left window shows the keywords, with the size indicating the number of times they occur in the database. The top right pane shows publication by year as a bar chart, with colour indicating the publication type, and the bottom right pane shows a matrix representation of the co-publishing relationship. Coordination across the views is managed so that clicking on a node in the node-link representation in the top selects the appropriate keywords for the individual and highlights their co-publications in the matrix and individual details in the bar chart in the top right.

Interface A, (Fig 3.1) is the initial CMV designed to visualise the IIDI data. It comprises four distinct displays described as outlined below for purposes of this research;

Upper left view of the interface (ULV)

Lower left view of the interface (LLV)

Lower right view of the graph (LRV)

Upper right view of the interface (URV).

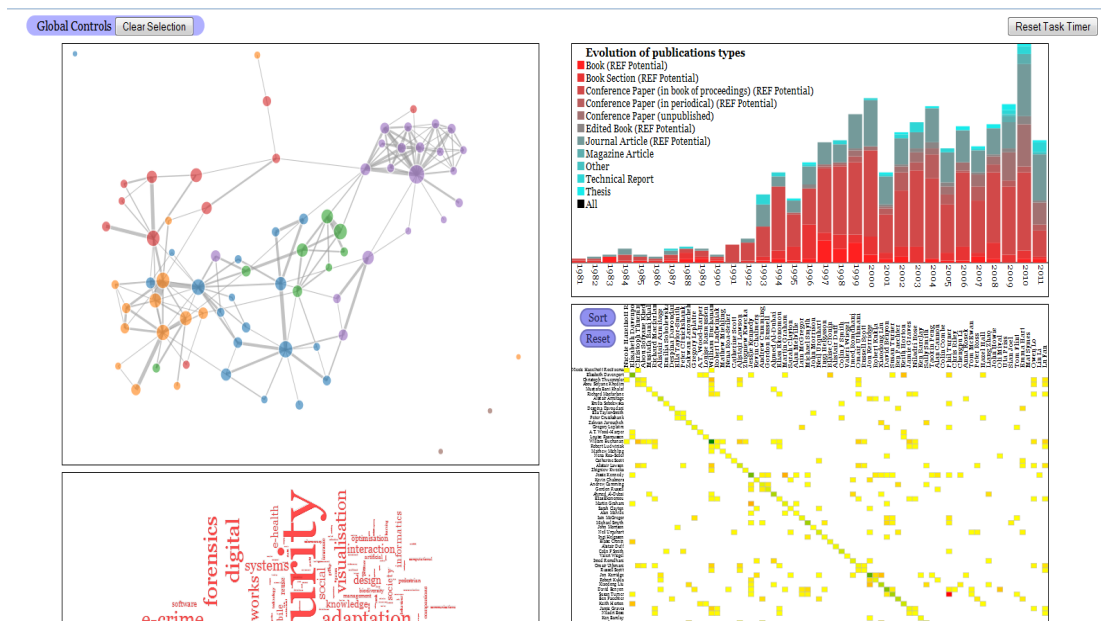


Figure 3. 1: IIDI Data CMV Interface A

Source: [www.soc.napier.ac.uk/~CS22/test\\_2/iidivislog.html](http://www.soc.napier.ac.uk/~CS22/test_2/iidivislog.html)

### The Upper Left View (ULV)

The upper left view (Fig 3.2) consists of nodes distinguished by different colours: blue, red, green, orange and purple to enhance visualization. Each node represents an individual member within the centre while each colour represents a centre within the institute. The *centre for interaction design* is denoted by *green* while *orange* stands for the *centre for information and software systems*. *Purple* signifies the *centre for social informatics* whilst *red* and *blue* represent the *centre for emergent computing* and *centre for distributed computing, networking and security* respectively. The colour acts as a filter and can be used to narrow the scope of data presented to the user. The user can visualise a data set of a specific centre by selecting a specific colour which allows the user concentrate on the elements that satisfy a required condition.

The node links show the existence of specific relationships such as joint publications between members whilst the nodes with no links indicate an absence of joint publications. The thickness of the node link demonstrates a variation of value of that relationship. For instance, the thickness of the node link is proportional to the number of publications co-authored. The size of the node points to certain trends in the publication history of the member represented by that specific node.

The ULV has a button, *Global control, clear view*, which allows users to return the view to its default state.

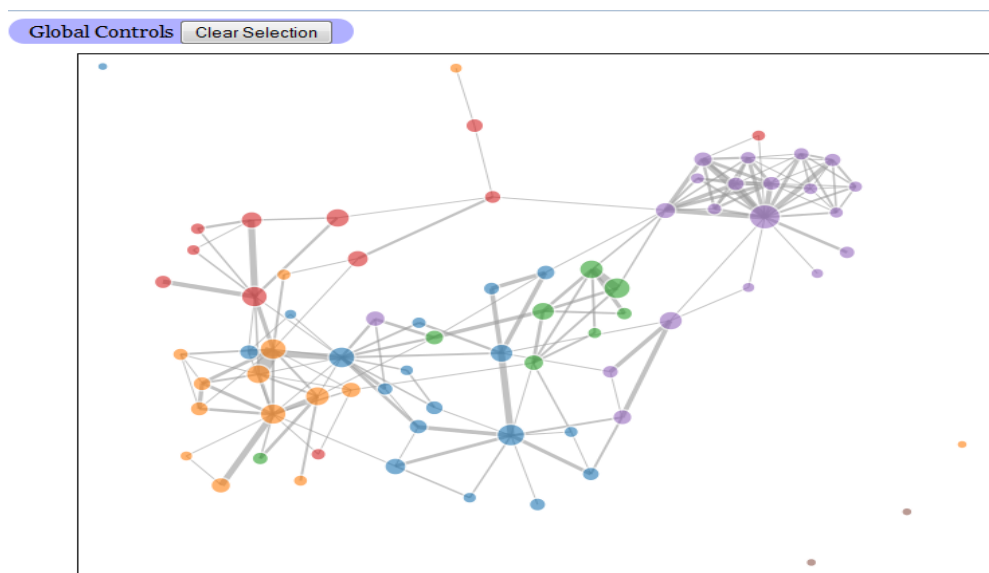


Figure 3. 2: Upper left view of IIDI CMV Interface

The main events that take place in the ULV are mouseover nodelink (MN), and selected nodelink (SN). The MN events occurs when a user hovers over a node to see the data provided before performing the SN event which is involves selecting the node. A mouseover a node yields data, in this instance the name of the individual member represented by that node (Fig 3.3).



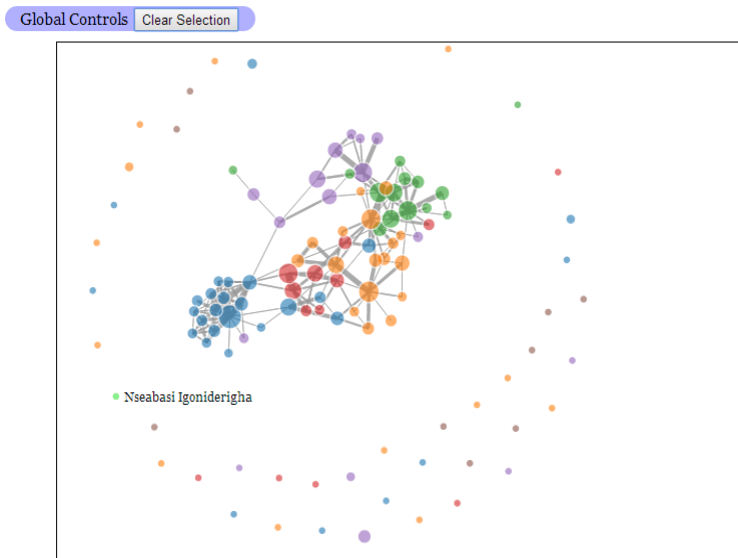


Figure 3. 3: Upper left View of IIDl Interface showing a mouseover nodelink event

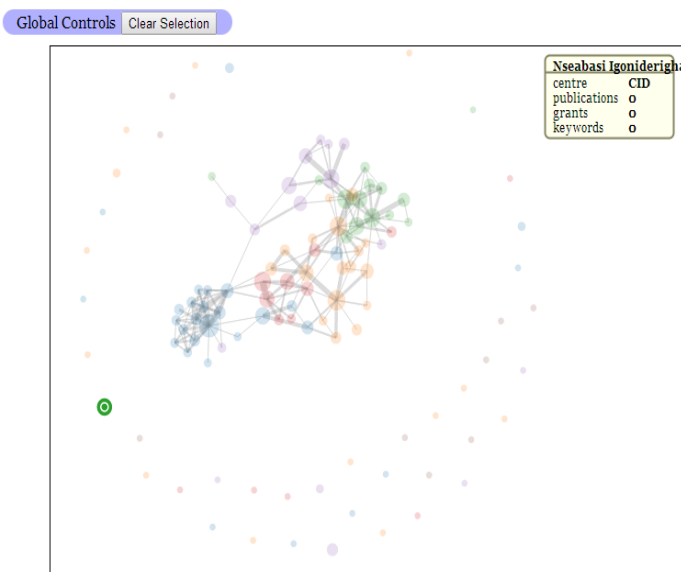


Figure 3. 4: Upper left view of IIDl interface showing selected nodelink event

Selecting a node provides metadata such as details of the member's name, centre they belong to, number of publications made, number of grants received and number of keywords (Fig 3.4). This view answers typical queries such as, how many

publications has person X made? Or how many grants has person y received? However, it is challenging to find a specific person, the user has to navigate to each node, perform several mouseovers until the person in question is identified. It is also difficult to interpret the data set because there are no clues to help the user understand the visual aspects of the interface.

### The Lower Left View (LLV)

The lower left view of interface A can be described as a tag cloud; it illustrates data in form of keywords and when a node is selected in the ULV, the keyword it is associated with is highlighted in LLV and appears bolder to catch the users attention. This view provides keywords associated with centres, groups and individuals.

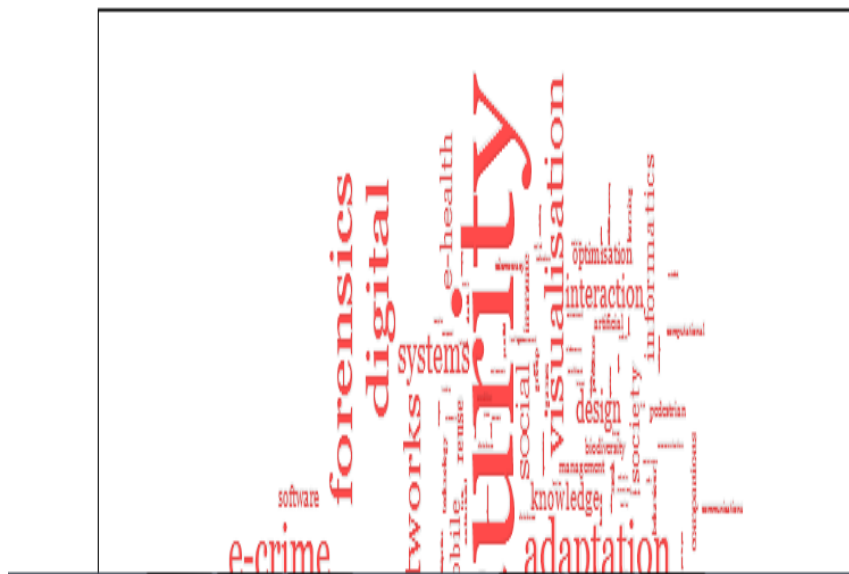


Figure 3. 5: Lower left view of IIDI CMV Interface

## The Upper Right View (URV)

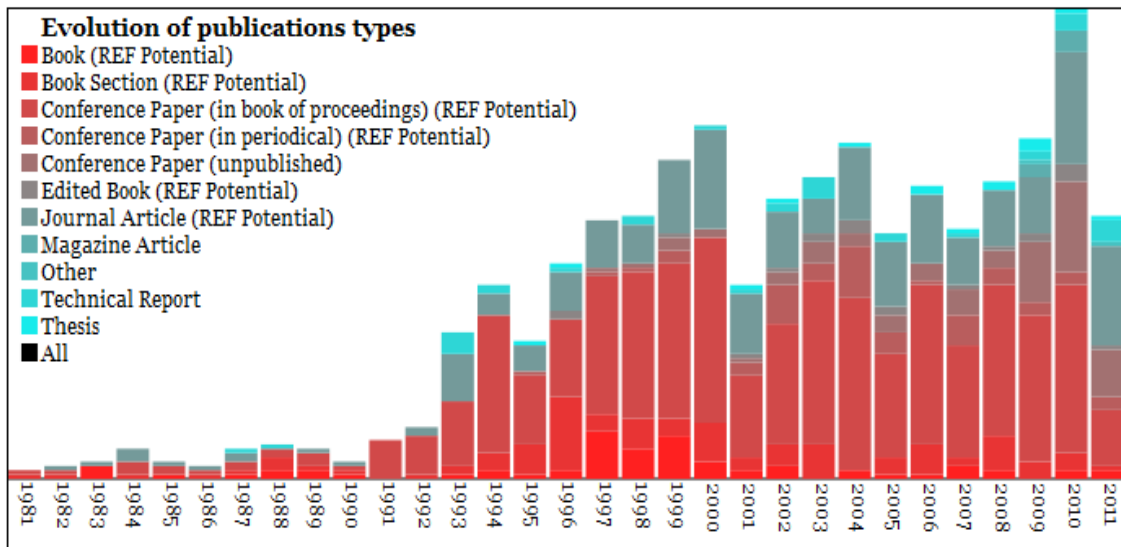


Figure 3. 6: Upper right view of IIDI CMV Interface A

The URV portrays data detailing evolution of publication types from 1981 to 2011 represented by a bar chart; each bar represents a specific type of publication. The main events performed on this view are mouseover Barchart (MB), filtered year Barchart (FYB) and Filtered publication Barchart (FPB). The MB is when a mouse is placed on the bar chart (fig 3.7) which yields data on the number of each publication type made while the FYB provides the date the publication was made. The URV has a manipulation attribute a reset task timer button used to reset the time, however currently it is not functional.

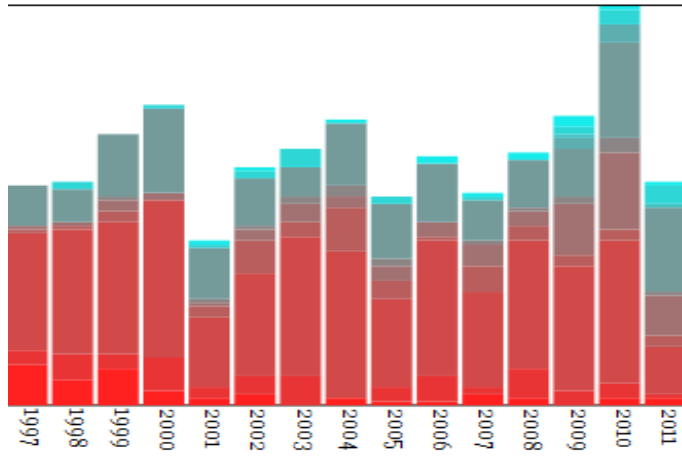


Figure 3. 7: Diagram of Bar chart on URV of IIDI CMV Interface A

The events that take place on the bar chart include MB and FPB. The FPB is when a mouse is placed over a bar representing the evolution of publication types (fig 3.8) to retrieve the total number of publications made for each publication type. It is colour coded to differentiate the twelve publication types and an explanation of the colour code referred to as metadata about the visual aspect of the data is made available to users, which makes it easier to understand and interpret the data set. This view is interesting as it illustrates the added advantage supplying a code can make to users navigating CMV's. The code acts as a support to help users navigate the view. Several interpretations can easily be drawn from this view; the bar chart (fig 3.6) clearly shows that the highest number of publications was recorded in 2010 while the least was obtained in 1981.

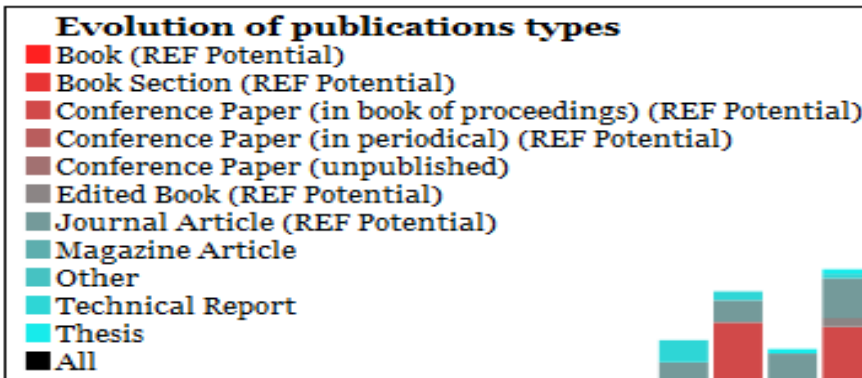


Figure 3. 8: Upper Right View of IIDI CMV Interface A, showing coding scheme

The Lower Right View (LRV)

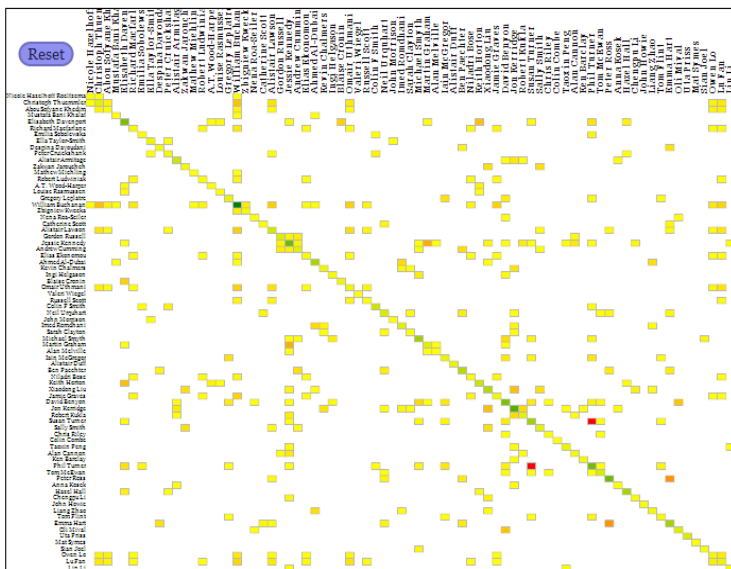


Figure 3. 9: Lower right view of the IIDI CMV interface

The lower right view (fig 3.9) is a scatter plot matrix illustrating the publishing relationship between members. Each matrix represents an individual member of a centre in the institute. The matrixes are red, yellow, light or dark green in colour. The chief events that take place within this view are mouseover matrix (MM) and selected

matrix (SLM). A mouseover matrix provides data such as a members name and number of publications made or names of two members that have made mutual publications and the number of publications made. When a selected matrix is performed, a group of nodes are highlighted in the ULV; these highlighted nodes indicate members who have mutual publications. The data entry selected in the LRV and related data in all other three displays are highlighted.

The “Reset” button on LRV does not cause any change however the “Sort” button gives the order in which the data will be represented by giving access to other buttons such as “centre”, “publications”, “grants” and “label”. However, these buttons when selected show different visual representations of the same data. The centre button presents the data (individual names) on the horizontal and vertical axis of the graph according to centres. Publication button presents the data (individual names) in order of publications made and the Grants button presents similar data (individual names) in order of grants received but does not provide any data on number of grants received. It is not clear what role the “label” button plays.

The type of navigation required to answer queries people request on a CMV such as the IIDI Data visualization is one that allows people interact with an information space the way they interact with physical spaces. People should be able to move through the data as they move through the real world, exploring, finding their way, and reaching their destinations and moving on when the need arises.

The Interface should offer support to people to navigate the information space whether they have a clear need for information about an object, or they need to explore the

space or on the other hand they need to arrive at a known destination. In order to help them understand the information space, descriptions of the space should be provided in form of maps, interface keys, navigation charts (), signposts should be provided to guide them to different parts of the space and information signs should be provided to tell them when they have reached their destination.

### ***3.4 Using the IIDI Data CMV***

The IIDI CMV was designed to make it easy to manage the IIDI data set. It should typically answer a range of queries for different categories of people. For instance, the university manager who requires information at the centre level such as data about grants; awarding bodies, recipients and so on, data about publications such as number of individual and mutual publications and pieces of publication. The institute director who requires data at the institute level, and someone else who may want to know the subject that the keywords link to.

The CMV shown in figure 3.1 was evaluated and found to support a range of user tasks such as 'how many conference papers were published in 2010', 'how many publications has person x made', 'who has person X co-published with', 'which centre does person X belong to', and so on. It supports tasks at the individual level in general which may not answer queries for the university manager or institute director. Interface A, does not support tasks at the centre or institute level. It does not answer queries such as 'how many publications have been made by centre A', 'how many publications are there', or 'which keywords are associated with centre B'.

Users have to make a lot of movements from one node to another and from one view to the other to query the individual people in order to aggregate the required data

themselves. This method is time consuming and many users get disoriented, and are unable to navigate successfully in order to visualise or retrieve adequate information.

### **3.4.1 Task Design**

In order to evaluate the IIDI CMV the following steps were taken; first, it was checked to see the extent to which it improves user performance, enables user manage complex information through interaction and discover new relationships. Secondly, a range of representative tasks to be performed on the IIDI interface were developed through a focus group and individual discussion organised by the researcher with the intent of finding out the type of questions the target user group would like to find answers to on the IIDI CMV.

#### The Focus group

The focus group comprised of ten postgraduate students from a common social background and was held in the Merchiston library of the Edinburgh Napier University. The researchers selected were peers and comfortable with themselves. The interface was introduced to them and they were given thirty minutes to familiarize themselves with it, communicate among themselves and come up with tasks and activities they would like to perform on the IIDI interface. A list of 20 tasks and activities were drawn up and after further evaluation; the list of tasks was scaled down to 15. The members of the focus group attempted to perform one task each during the session on the interface A of the IIDI data CMV and a success rate of 2 out of 7 was obtained. The session was rounded off with suggestions from the members on how to make the interface more navigable. For instance members agreed that the key provided for the URV made it easier to navigate. The tasks and activities were



interpreted by the researcher who acted as the facilitator. The focus group lasted for two hours.

The focus group method was selected to allow a wide range of queries to be collected, and give an opportunity for individuals to offer their opinion about the interface. As people tend to become influenced by others during a focus group and results obtained could be biased, a one to one discussion was carried out at separate times with five different participants. The participants include; one university staff, three research students, and one centre director and 10 similar questions were obtained which were reduced to 5 at a later date.

Following the focus group, the researcher performed a walkthrough of all 20 tasks obtained from the focus and individual group in order to ascertain their feasibility and select representative tasks for each navigation activity identified. After the walkthrough, 15 representative tasks were selected and discussed with a centre director and two members and the questions were narrowed down to 10 for the empirical study.

The questions selected are typical questions and are presented in form of representative tasks that are comparable to tasks a member, a centre director, university manager or general user may perform on the IIDl CMV in order to obtain information. The tasks were designed to check if users are able to perform realistic tasks, find facts and navigate the information space.

The focus of the research is on navigation of information space, therefore, as users execute the tasks, it will reveal the navigation behaviour of users on the interface and will be useful in the process of identifying and understanding the navigation challenges users face on the interface and where improvements need to be made. In addition, the tasks execution will illustrate how much information the users can retrieve, the type of tasks they can complete with the system and the extent of support the system offers users to navigate the information space.

The tasks will allow an assessment of the data type displayed when a mouseover is performed on a navigation feature such as a nodelink. In addition it would give an insight to the level of information people can navigate to in order to visualise and retrieve information. Furthermore, the tasks will reveal how people retrieve information from the CMV; do they have to aggregate information themselves? Or go elsewhere to obtain data? It will also assess if the navigation features provided such as nodelinks, pop-up windows, search buttons, and tag cloud by the information space are sufficient to support successful navigation.

### 3.4.2 Task Categorization

The table below gives details of tasks categorized into object identification, exploration and wayfinding. The tasks were categorized to incorporate the three navigation activities to allow the researcher look at the tasks from the navigation perspective.

Task	Category	Description
1. Identify and list two different things on the	Object Identification	The goal is to see how participants can effectively

Task	Category	Description
visualization and write down the information they provide		identify and categorize groups and clusters of objects spread across the given interface. It will also show how well they understand the organization of these objects and if they can retrieve adequate information about them.
2. How many publications has member X made?	Wayfinding and Exploration	The wayfinding aspect of the task demonstrates how well participants are able to interpret the interface, observe events in the different displays and find their way around the interface as they search for specific information. The exploratory part shows their ability to find relationships that exist, in this case, between an individual and their publication.
3. Locate two people in the Centre for Interaction Design (CID) that have	Exploration	This task aims to examine participants' ability to find relationships that exist within

Task	Category	Description
made joint publications with people from other centres'?		the surrounding (for example, centres') and other surrounding (centres') and locate common patterns (publications) and trends that exist.
4. Which centre does member Y belong to?	Wayfinding	The main focus of this task is to observe how participants work out how to reach their destination, the routes they take and if they can tell when they achieve their goal.
5. What subject is the Centre for Emergent Computing (CEC) interested in?	Object Identification, Wayfinding and Exploration	This task will point out participants' ability to retrieve adequate information about a centre. Identify and locate the centre, then discover the relationship between the centre and its subject of interest.
6. How many grants have been awarded to the Centre for Social Informatics (CSI)?	Wayfinding and Exploration.	The wayfinding phase of this task will show participants' ability to reach a destination (find a centre) while the

Task	Category	Description
		exploratory part will demonstrate their understanding of existing relationships.
7. Which individual has made the highest number of publications? Which centre does the person belong to?	Wayfinding and Exploration.	This task will demonstrate participants ability to recognize paths they have taken before to reach required destinations more easily and their understanding of the overview presented by the information space
8. How many mutual publications has member Z made in the last three years?	Object Identification, Wayfinding and Exploration.	This task shows the degree of information participants can find and retrieve about a specific member (relationships and patterns).
9. List the five Centres in the institute and identify the centre with the least number of members.	Object Identification and Wayfinding.	This task shows the level of information participants can retrieve about the categorization and organization of a group of objects.

Task	Category	Description
10. Which centre has the highest number of publications?	Wayfinding and Exploration.	The tasks intent is to see how familiar participants are with the information space and locating relationships within its surrounding.

Table 3. 1: Tasks categorisation

### 3.4.3 IIDi Data as an Information space

The IIDi is a typical research institute that keeps data about the people in the institute, the centre to which they are attached, the publications they have made, and the grants that they have obtained. In other words, the IIDi data can be expressed as a relational model that looks at objects on the information space in terms of the data that is in the entities. The entities (relations) include person, centre, publication and grant. In addition each person, centre, grant and publication is associated with a keyword. A typical relational database implementation of this would consist of the following relations (with the primary key underlined):

Person (PersonID, Name, CentreID)

Centre (CentreID, Name)

Publication (PublicationID, type, location of publication, year)

Grant (GrantID, Amount, Funding Body)

Co-Published (PersonID, PersonID, PublicationID, keyword)

Co-Investigated (PersonID, PersonID, GrantID, keyword)

The relations Person, Centre, Publication and Grant record details of these objects. Co-Published represents the relationship between the people who co-publish a publication. Co-Investigated, records the relationship between people who have collaborated on a grant. A typical CMV representation of the IIDI data is shown in Figure 3.1. The E-R relationship existing in the IIDI CMV interface A is the one-to-one relationship, illustrated in chapter 2, figure 2.7.

Here we propose a many-to-one relationship illustrated in the E-R diagram below (figure 3.10) to improve navigation.

The diagram in figure 3.10 illustrates person as the main entity that is the main focus of the data set. Person has attributes such as keywords, grants, centre and publications. All these attributes have a many to one relationship with person. A summary data is data about these attributes which turns out to be metadata about person. When data about centre is provided along with a summary data about centre, metadata about person has been knowingly provided.

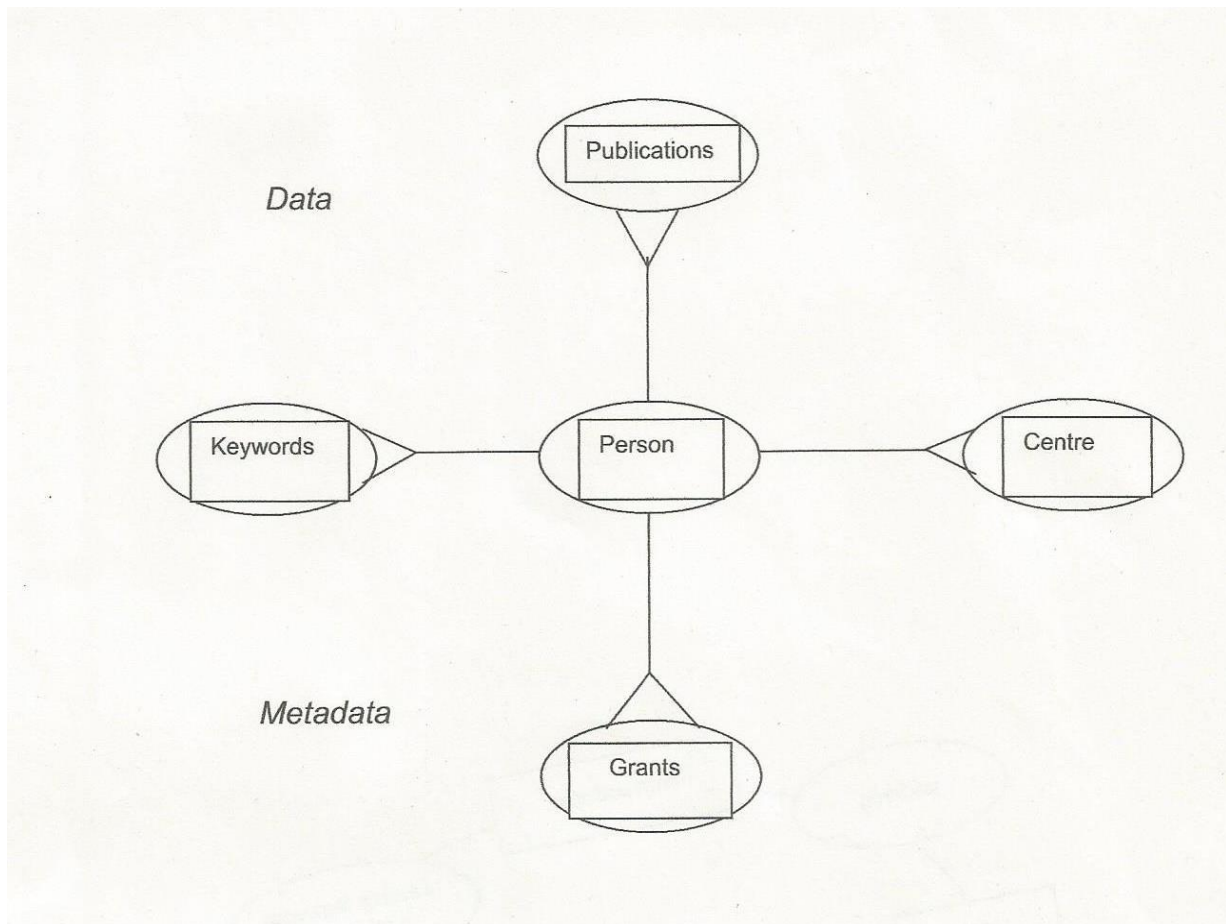


Figure 3. 10: Entity relationship diagram illustrating many-to-one relationships

### 3.4.4 Metadata as Many-to-one (M:1) Relationship

In a many-to-one relationship, a single entity usually a column or a set of columns holds values that make reference to another single entity that has exclusive values. These many-to-one relationships are imposed by a foreign key/primary key relationship and occur in relational databases like CMV's. The primary key is unique in nature for example "MemberID", it recognises a record in the table such as member table. The foreign key can be explained as a relational table that matches the primary key column of another table and may be used to make cross references between tables (Date.1996, Carlos, 2010, Ramez 2011). The table where the foreign key is stored is termed the referencing or child table (Robert, 2005). Both keys make it



possible to link tables thus creating navigation routes and pathways. Many-to-one relationships describe groupings or classifications and illustrate relationships between tables and levels of hierarchy. An example is a university schema which holds tables such as Institute, Centre and Member. Each institute has many centres, each centre has many members however a member can belong to only one centre and a centre can belong to only one institute. The point noteworthy is that a member can belong to only one centre but a centre could have many members described as “many-to-one”. A member can have many publications, co-publications, grants and keywords.

In another perspective the “member” can be termed data and the details about the member termed metadata. The design approach proposed by this research is founded on the hypothesis that when data is presented alongside its metadata, navigation is more successful, visualisation is effective and adequate information is retrieved. The different objects in the database will have a many-to-one relationship allowing data to be presented in different levels. The many-to-one relationship will create a platform for a more detailed description of links between database tables referred to as entities to be provided. This will improve navigation between tables which is the objective of this research. Data about the attributes that have a many-to-one relationship (in this case centres, publication, grants and keywords) will be presented alongside metadata about the main focus of the dataset (in this case member) to allow people navigate easily from one level of information to another level, visualizing the information provided effectively and retrieving required information.

### **3.4.5. Navigating the IIDI Information Space**

A detailed study of the IIDI CMV was carried out by the researcher with the aim of understanding and finding a solution to the navigation issues identified. The main challenge observed was that it was difficult to move between data, retrieve sufficient information and discover relationships. The basic relationship existing between the data was a one-to-one relationship and the type of data presented provided inadequate information. Through research, the concept of navigation of information space was identified as a feasible method to improve navigation which involved harnessing many-to-one relationships and metadata. These many-to-one relationships are enforced through a primary/foreign key relationship.

In the E-R model (figure 3.10), the primary key is the member name and the foreign key is the centre name. The schema proposes a many-to-one relationship. In this case, “people” is the main focus of the database and has a many-to-one relationship with attributes such as, publications, centres, grants and keywords. An institute can have many centres and a centre can have many members but a member can only belong to one centre and a centre can only belong to one institute.

#### **Navigation Steps**

In order to navigate a given information space, the user needs to carry out certain activities to enable them achieve their purpose. These activities are object identification, wayfinding and exploration. Users perform mouseovers as they search and browse and they perform clicks when they need to carry out more explorations or retrieve information. Mouseovers and clicks are described as navigation steps; they define the way people move from one point to another on the interface. The research

proposes that many-to-one relationships will allow data to be displayed along with a summary of its data; this will improve movement between objects within views and from one view to another on the interface. In addition it will reduce the number of mouseovers and clicks people have to perform.

### Navigation Levels

Where data is provided along with the summary of its metadata, information can be displayed in different levels. For instance providing data about centre's along with a summary data of centre provides metadata about people and allows easy movement from a view of data to a view of metadata, in order to visualize and retrieve adequate data at the required level. Features that afford users the ability to access other levels of information are provided. They include labels such as centre data, publication data and grant data are that describe the window the label leads to and act as gateways to access more information. This strategy allows people access information suited to their needs. Metadata contents can be varied owing to the levels of data they represent. Navigating to different levels of information can be carried out depending on what the user requires. At the one level data presented could provide sufficient information to support user navigate between data in order to identify objects of interest on the interface. Users should be able to find answers to questions like Who? What? Where? When?

The labels provided guide the user to other levels of information suitable for instance a member, centre or institute director. The user can explore relationships, and decide which levels of data are suitable for the intended purpose, for example are the data suitable for a planned analysis? For instance in the case of the IIDl data, it should

provide information that allows people draw useful conclusions on grant and publication history of centres.

The lowest, most detailed level is the metadata which are essential for obtaining and using the data. A useful analogy is the labelling in a shopping mall. The information in front of the mall could be likened to object identification: the shop names and brands, as well as what the consumer can expect to see. The customer has to read the directions in more detail to obtain information on how to reach their destination (wayfinding). Finally, the customer has to explore the mall to discover relationships, manage the directions (exploration), and utilize the benefits provided in the shopping mall. It is clear that depending on the purpose of the metadata catalogue, the level of metadata used will vary.

### ***3.5 A Redesigned CMV for the IIDl Database***

The design approach proposed by this research involves implementing the concept of navigation of information space on the IIDl CMV. The concept involves designing for people who live in their information spaces, and navigate through these spaces to visualize and retrieve information in a manner similar to what is obtainable in the physical world. The researcher studied the IIDl data CMV interface A, in detail and identified the navigation issues associated with the interface was redesigned to produce Interface B. Following further evaluation, a more navigable interface C was produced.

The first stage of the redesign involved replacing the upper right view of interface A, (figure 3.1) that was previously a Bar chart with a set of five nodes designed to

represent each of the five centre's (fig. 3.13).matrix representation of centre collaborations in the upper centre view (fig 3.15). Next, a set of five nodes were created to represent each of the five centres in the institute in the URV, figure 3.13. The aim is to develop a design based on understanding data alongside its metadata, and to incorporate common instances where one person's data is another person's metadata. In order to achieve this goal the designer needs to complete the following steps:

- Decide on the main focus of the data set
- Decide on the key relationships based around the focus
- Decide on the visualization type
- Decide on the visualization detail – color, line and so on.

As soon as the designer decides on the main focus of the data set, which in the case of this study is people, the attributes that have a many-to-one relationship with the main focus, such as centre, publication, grants and keywords, for the case in question are identified. Data about these attributes are then provided, and this type of data is expressed by this research as summary data. A summary data is defined as data about the attributes that have a many-to-one relationship with the main focus of any database.

The summary data makes it easier to answer questions at different levels; it makes it possible for people to move from a view of data to a view of metadata. It adds flexibility to the CMV and supports the theory of navigation of information space which moves away from the task focus to embrace a wider range of activities, wayfinding, object identification and exploration, to improve navigation.

### 3.5.1 Interface B

Interface B (Fig 3.11) is the first redesigned version of the IIDI CMV Interface A, with summary data added to aid navigation. It has five views as opposed to the four views in Interface A, (figure 3.1). The matrix representation of collaborations (figure 3.12) known as Upper Centre View (UCV) and the nodelink showing centres in the institute (figure 3.13) named Upper Right View (URV) were added to replace the barchart in interface A which represented publications. The other three views, Lower Left View (LLV), Upper Left View (ULV) and Lower Right View (LRV) were retained and remain similar in structure and function. The purpose of the UCV and URV is to examine the effect of the summary data on navigating the CMV.

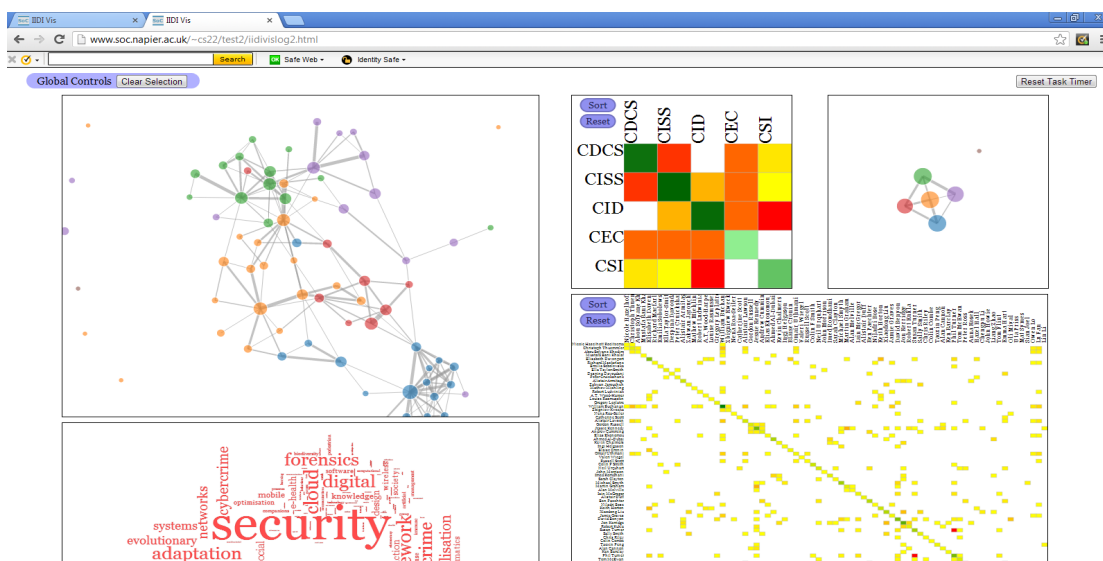


Figure 3. 11: IIDI Data CMV Interface B

Source: [www.soc.napier.ac.uk/~CS22/test2/iidivislog2.html](http://www.soc.napier.ac.uk/~CS22/test2/iidivislog2.html)

The upper right view (URV) of Interface B shows a lone node which symbolizes the *Institute of Informatics and Digital Innovation (IIDI)* and five linked nodes colour coded

for centres. Each of the linked nodes stands for a centre in the institute. Green, represents the *Centre for Interaction Design* (CID), purple signifies the *Centre for Social informatics* (CSI), red denotes the *Centre for Emergent Computing* (CEC), blue stands for the *Centre for Distributed Computing, Networking and Security* (CDCS) and orange indicates the *Centre for Information and Software Systems* (CISS).

In the *upper left view* (ULV), each node on the nodelink represents a member, and the colour of the node points to the centre the member belongs to. The *upper centre view* (fig.3.12) presents data on centre collaborations such as collective grants award between centres through a set of colour coded matrix.

When a mouseover matrix event is performed, an abbreviated data giving centre names is obtained. The matrixes coloured green and aqua provide information about the total number of grants obtained by individual centres. The yellow matrix gives information about mutual grants between the *centre for social informatics* (CSI) and other centres while the orange matrix provides data about joint grants between the *centre for interaction design* (CID) and any other centre. The red matrix and deep orange matrix provide information about shared grants between the *centre for information and software systems* (CISS) and the *centre for distributed computing, networking and security* (CDCS) respectively.

Interface B was enhanced by providing summary data of the metadata by introducing many-to-one relationships in the entity relational model in addition to the one-to-one relationships in existence. The aim of the design is to provide access to information at different levels and create more routes for users to follow on the interface and enhance successful navigation. The summary data was provided along with data to make more

details about the digital asset available and answer a wider range of questions. Take for example a digital asset “publications”, interface A answers questions such as ‘how many publications has person X made’, ‘when has person Y published’, which may be sufficient for a user in a certain category but it does not answer questions like ‘how many conference papers have been published’, how many journals were published by centre X’ or ‘which keywords are associated with centre Z’ which a user in another category such as an institute director may require.. The provision of a summary data is aimed at improving the ability of general and specific users to navigate the information space, carry out various activities and retrieve answers to more questions.

The “Reset” button does not create any change; however when the sort button is selected, it reveals four windows; publications, people, label and grants. When the publication window is selected, and a mouseover matrix is performed, the total number of publications made by a given centre is displayed. This is also obtainable in the case of grants and people. It was hoped that evaluation of Interface B would demonstrate improvements in navigation, but it was anticipated that other issues would emerge and further guidelines would be needed. For example, the issue of access to summary data had not been resolved in interface B, so further development of the upper centre view was performed to produce Interface C.



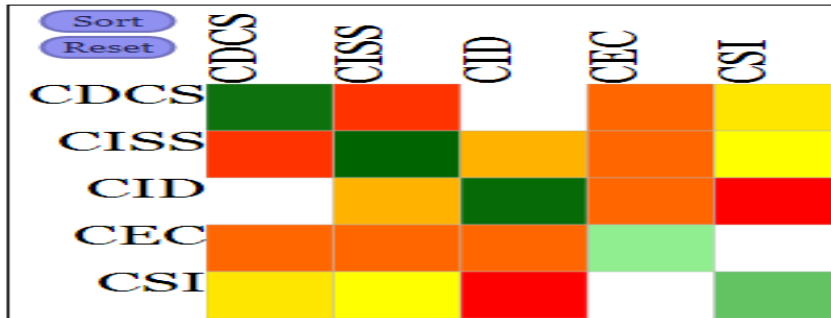


Figure 3. 12: Upper centre view of interface B

### 3.5.2 Interface C

The disparity between the IIDI Data Interface C and interface B is the Upper Centre View. The UCV in interface C provides more data and metadata for the centre attribute than the UCV in interface B. The UCV is a matrix representation of centre collaborations.

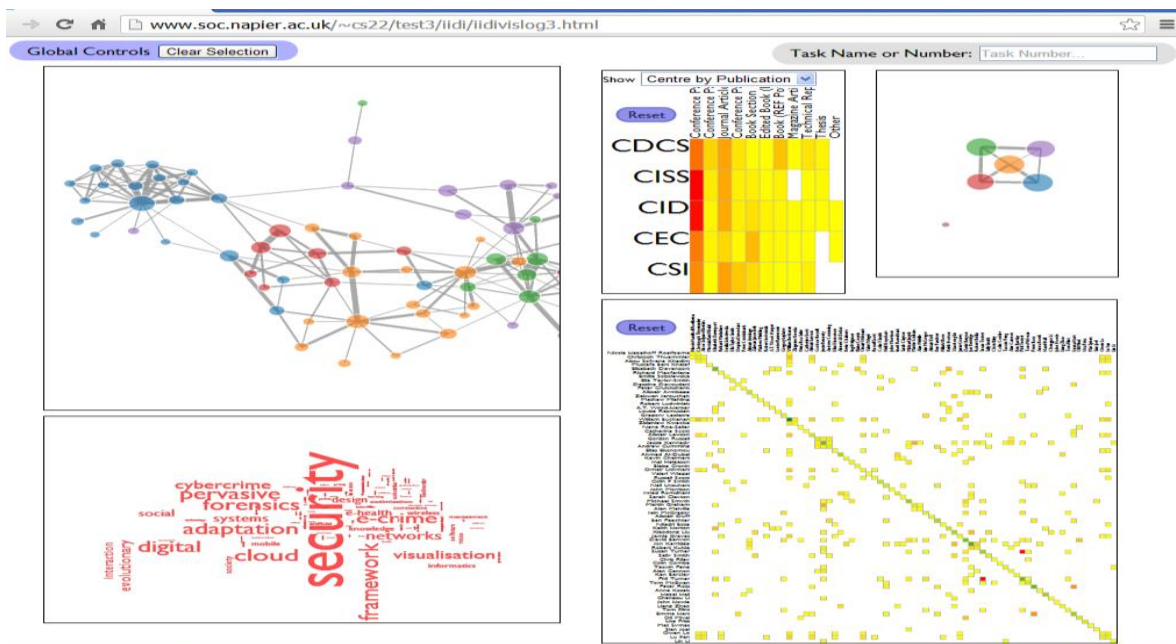


Figure 3. 13: IIDI Data CMV Interface C

Source: [www.soc.napier.ac.uk/~Cs22/test3/iidivislog3.html](http://www.soc.napier.ac.uk/~Cs22/test3/iidivislog3.html)

Interface C (fig. 3.13) illustrates the fully redesigned IIDI data CMV. It presents five views discussed in detail in subsequent paragraphs. All the views except the upper centre view are identical to the corresponding views on interface B.

The upper right view of interface C (fig. 3.14) presents a lone node which yields the acronym IIDI (Institute of Informatics and Digital Innovation) and a set of five linked nodes, colour coded to represent each of the five centres.

The Upper right view (fig. 3.14) of interface C allows users to retrieve information about centres easily. A selected nodelink event gives details such as the number of people, publications and grants in the centre, the name of the head of the centre and the abbreviated name (fig. 3.15) while a mouseover node event performed gives the abbreviated name of the centre (fig. 3.16).

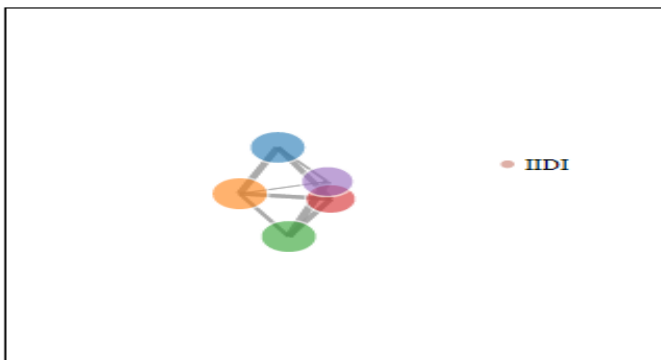


Figure 3. 14: Upper right view of interface C

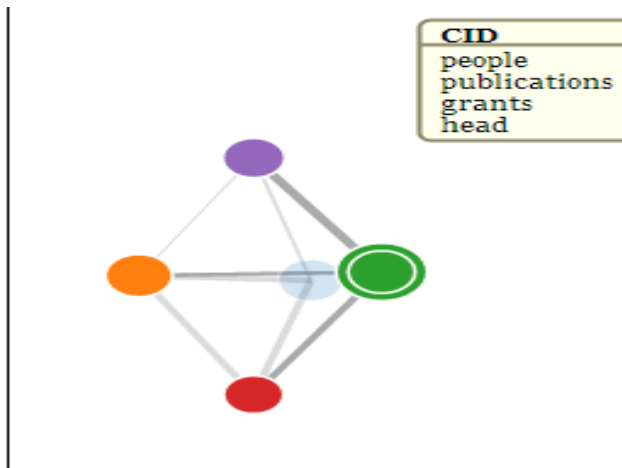


Figure 3. 15: URV of IIDI Data Interface C illustrating selected nodelink event

This view gives the user the required support to navigate more easily through the nodelink in the ULV in order to visualise and retrieve information. For instance, the URV can be compared to an information sign such as the traffic light in the physical world. The user gains knowledge of its meaning by asking someone else or reading the drivers manual, while on the information space, the user performs a mouseover or selects a node. The summary data supports the user's wayfinding need and tells them when they have reached their destination.

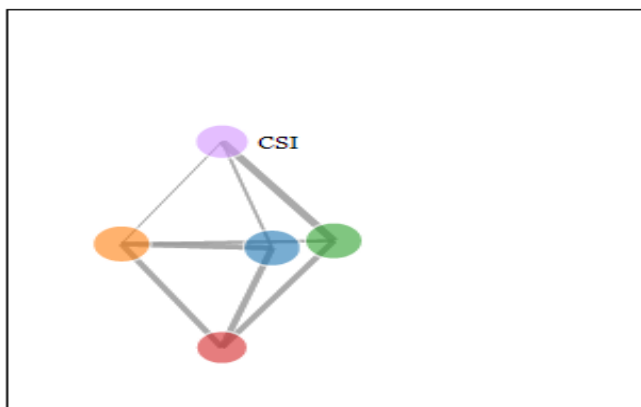


Figure 3. 16: URV of IIDI Data Interface C illustrating mouseover nodelink event

The upper centre view (fig.3.17) was altered. It holds four windows created in order to give users access to summary data about centres. It also permits users navigate from one view of data to another, visualise data and retrieve detailed information about publications and grants at the centre level. Users can navigate to *centre by publication*, *centre by grants*, *cross centre publication* or *cross centre grants* to visualise and retrieve a wide range of information. Users can identify objects, find their way and explore the data set more easily. Navigation, visualization and information retrieval are carried out more effectively.

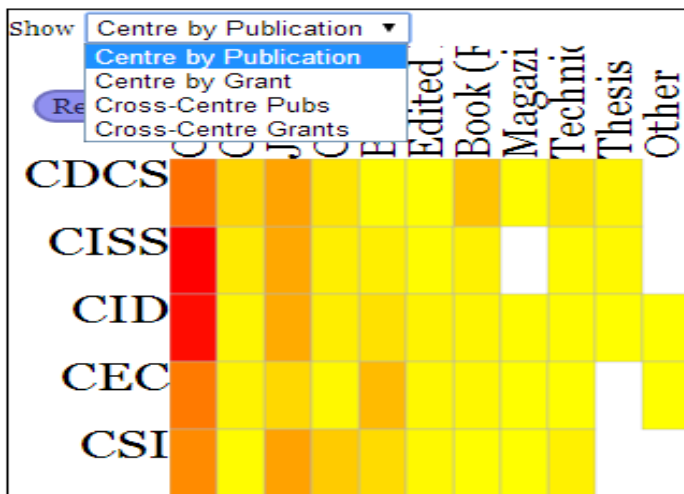


Figure 3. 17: UCV of IIDI CMV Interface C illustrating the labels

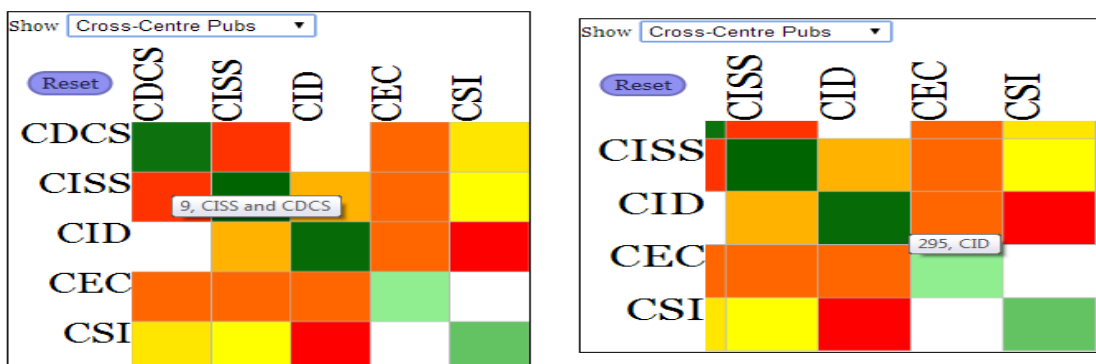


Figure 3. 18: Mouseover matrix event on UCV cross centre publications window.

When the cross centre publication window is selected and a mouseover matrix event is performed on a green matrix, it provides a summary data of the total number of publications made by a given centre (figure 3.18b), on the other hand a MM event performed on any other matrix gives a summary data of total number of mutual publications made by any two given centres (figure 3.18a). Comparable information is obtained for grants as illustrated in figures 3.19a and 3.19b.

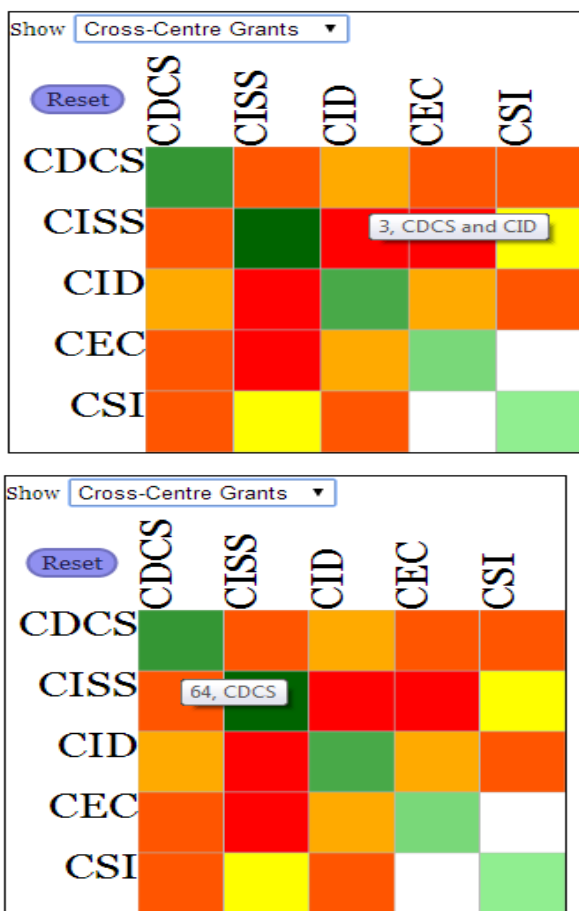


Figure 3. 19: Mouseover matrix on UCV cross –centre grants window.

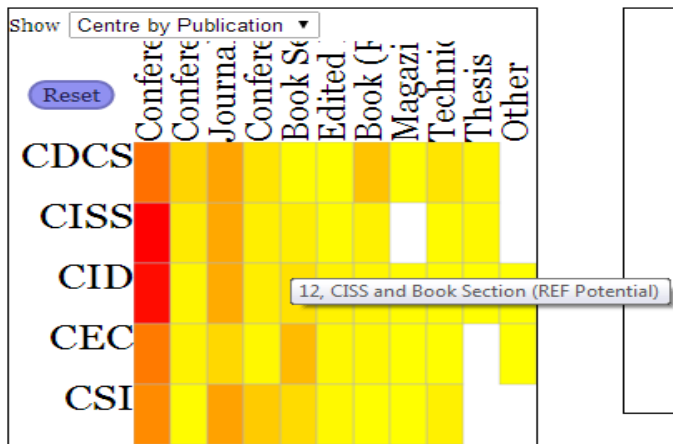


Figure 3. 20: Mouseover matrix event on UCV centre by publication window.

The diagram above (figure 3.20) exemplifies the information obtainable when an MN event is performed. The summary data reveals three details about the publication; the centre that made the publication, in this instance, CISS. The type of publication, in this case a Book Section (REF Potential) and the number of publications made, in this case 12. Similar information is obtained for grants such as the centre that received the grants, the number of grants received and the type of grant (figure 3.21).

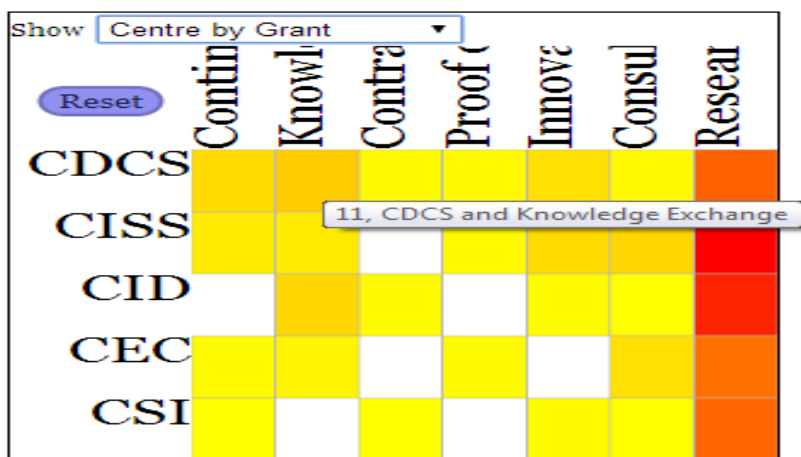


Figure 3. 21: Diagram of UCV showing centre by grants window on Interface C

The redesign phase of interface C brought about the development and implementation of the first three guidelines proposed by this research for designing CMV's. They are outlined as follows:

1. Decide the main focus of the data set.
2. Provide access to the attributes of the main focus of the data set
3. Provide summary data

### **3.6 Conclusion**

This chapter described the different views of the IIDI CMV and explains the redesign of the Interface. Initially, an overview of the IIDI interface was given by describing its structure. The concept of navigation of information space was presented as a method to improve navigation. A many-to-one relationship was incorporated and a summary data provided to create a platform for information to be accessed at different levels.

The IIDI Data Interface C is the proposed design for CMVs developed for MIS. It is designed to provide data about centres, publications, and grants which are attributes that have a many-to-one relationship with the main focus of the database which is metadata about people. The design gives people access to the data about centres directly rather than people having to aggregate these data themselves. The redesigned interface represents an initial attempt to design a navigable CMV with the concept of navigation for information space and summary data.

The forthcoming chapters describe the empirical work undertaken on the interfaces and to assess the usefulness, benefits and challenges of the redesigned Interface C.

## Chapter 4 Empirical Study One

### ***4.1 Introduction***

In the previous chapter, the IIDI CMV was described in detail. It is clear that developing CMVs require a design approach that will make them navigable and bring out their flexible nature in order to allow users accomplish the purpose they set out to achieve on these information spaces. As a principle of interaction, the concept of navigation of information space is not focused exclusively on tasks but incorporates activities such as object identification, wayfinding and exploration. These activities when supported by CMVs allow people to move up and down the levels of data and metadata created by the provision of summary data in a more refined manner.

Even though the development of information visualization has been rapid, insufficient empirical evaluations have been carried out on aspects such as navigation that are clearly important to the subject. The empirical work performed in the subsequent chapters deals with the need for new evaluative methods that can be applied to address challenges specific to information visualization (Chen, 2005).

The design approach proposed by this research is based on people understanding data alongside its summary data. Each piece of data is seen as a separate object with its own summary data, and all the data together is seen as an information space which requires people to navigate through successfully in order to visualise and retrieve adequate information.



This chapter discusses the methodology employed for this study. It describes the research methods chosen and the user based empirical study carried out to evaluate the original IIDI CMV (interface A), in order to assess the degree to which it supports people to navigate through the data set and its effect on visualization and information retrieval.

## ***4.2 Research Methodology***

The experimental method was utilized in this study to test the hypothesis proposed by this research. This method was chosen in view of the fact that it allows a theory to be tested to determine whether it influences a dependent variable, and allows two groups to be compared (Creswell, 2004). Observational studies will be employed carried to enable the researcher observe and record events and acquire knowledge for effective quantitative and qualitative analysis of the study (Albert & Wulf, 2002; Hennesey, Patterson & Lin, 2003). The experiments conducted were organized processes performed in order to refute or confirm the hypothesis proposed by the research (Thomas & Cook, 2005). Laboratory experiments were conducted to illustrate the outcome of altering specific features on the IIDI data visualization. we

### **4.2.1 Experimental Investigations**

The experimental method gives the researcher a degree of control over the environment. It makes it possible to manipulate the independent variable (IV) purposefully, and keep the dependent and random variables constant. The independent variable (IV) is the variable of interest to the researcher, while the dependent variable (DV) is the measure of change observed. Random variables include all other variables that may affect the results. Experimental investigation is a

direct method through which the researcher can reach conclusions confidently knowing that any change in the DV could be as a result of the condition altered in the IV. It has been argued by researchers that the degree to which variables can be managed is directly proportional to the potency of the results. On the other hand researchers agree that it is not possible to manage every variable that exists such as a participant's mental state during the experiment (Lazar, Feng and Hochheiser, 2009).

For this research a laboratory experiment was chosen to allow the researcher study the effect a specific feature could have on the given interface when altered. The activities participants engage in as they navigate the given information space are recorded by means of a log file to facilitate analysis, interpretation and presentation of results. In addition it affords the researcher the opportunity to observe participants and make notes.

#### **4.2.2 Descriptive Investigation**

In this instance, observational studies and interviews were carried out with the objective being to enable the researcher acquire an understanding of participants view on the concept of navigation of information space and navigating CMVs. It also allows the researcher to observe the participants and record descriptive and reflective notes. The descriptive notes help to describe significant events and activities that occur during the experiment, on the other hand reflective notes are individual suggestions from participants based on their insights or general idea (Creswell, 2011). It affords the researcher the opportunity to collect data discreetly in a more natural setting. According to Berg (2001) observational studies can be used to provide a different view

to compare data collected in various ways as part of a quantitative research or to portray general ideas for qualitative research. Furthermore, it provides in-depth knowledge of processes that analysis may lack and help the research make out what work needs to be done in future.

The informal interview was performed as a second descriptive method, participants were asked open ended questions, to enable them express their opinions and talk about their experiences (Creswell, 2004). The data collected from both methods are recorded for analysis as it is a helpful method used to confirm research data.

### ***4.3 The IIDI Data Visualization Interface***

The IIDI data was initially visualised using the Author Network Graph (ANG). It was created by a database designer to allow people visualise data about the institute of Informatics and Digital innovation in the computing department of Edinburgh Napier University. It is a CMV that allows people to see and interpret the IIDI data from different perspectives. It consists of data about the institute, the centres belonging to the institute, members of these centres and collaborations. The empirical study involves a user based study that will demonstrate how users navigate through the information space to visualise and retrieve information. It will give an insight to the actions they take and answer questions such as; does the information space support the navigation needs of users, and to what extent? Are users able to navigate successfully? Visualise effectively and retrieve adequate information?

Two separate information spaces shown below Interface A, (figure 4.1) and Interface B (figure 4.2) were employed for the first experiment.

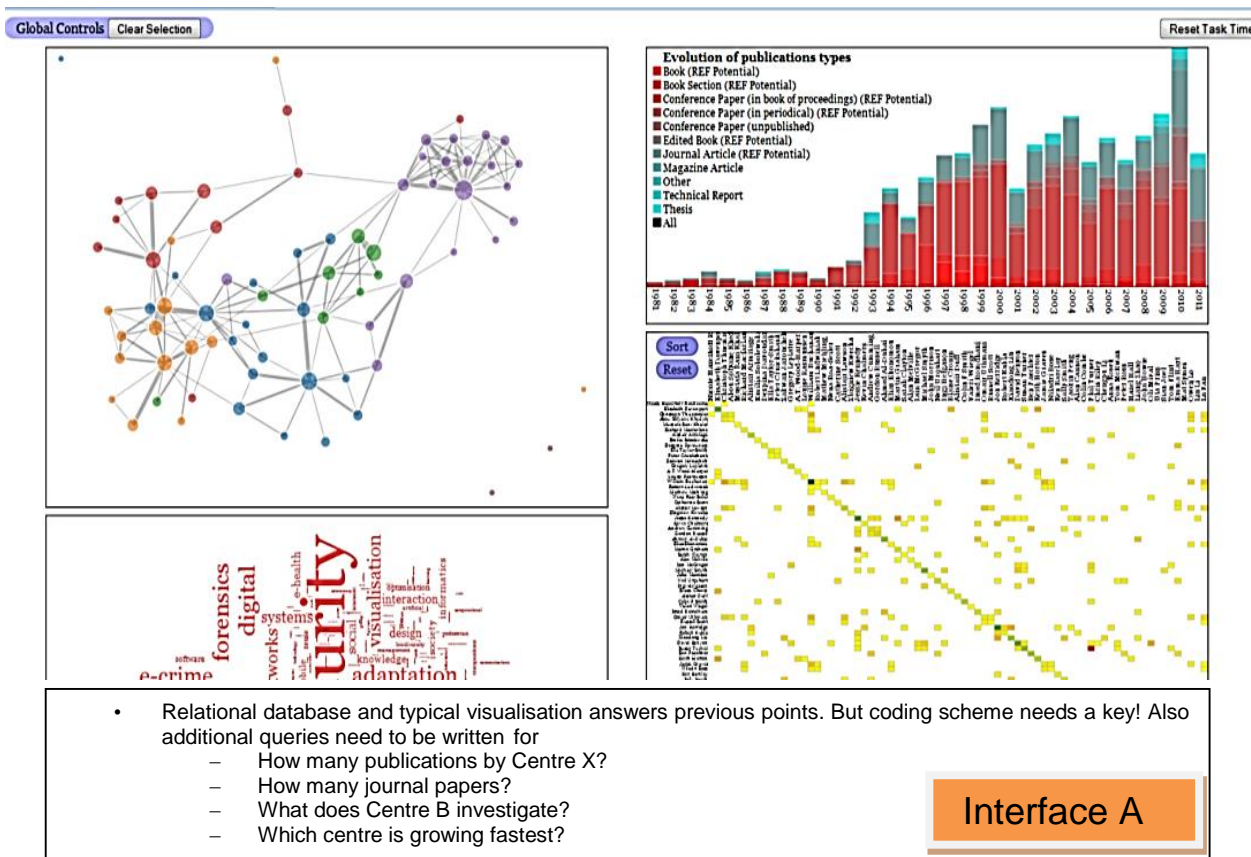


Figure 4. 1: IIDI Data CMV Interface A

Source: [www.soc.napier.ac.uk/~CS22/test\\_2/iidivislog.html](http://www.soc.napier.ac.uk/~CS22/test_2/iidivislog.html)

Interface A, (fig. 4.1) is the initial IIDI CMV while interface B (fig. 4.2) has been altered by implementing the three novel design guidelines developed during this study. First, the main focus of the IIDI data set was identified as people, next access to the attributes that have a many-to-one relationship with people was made available through four windows provided on the upper centre view by introducing a summary data. The redesigned interface B was then assessed through tasks executed by participants.

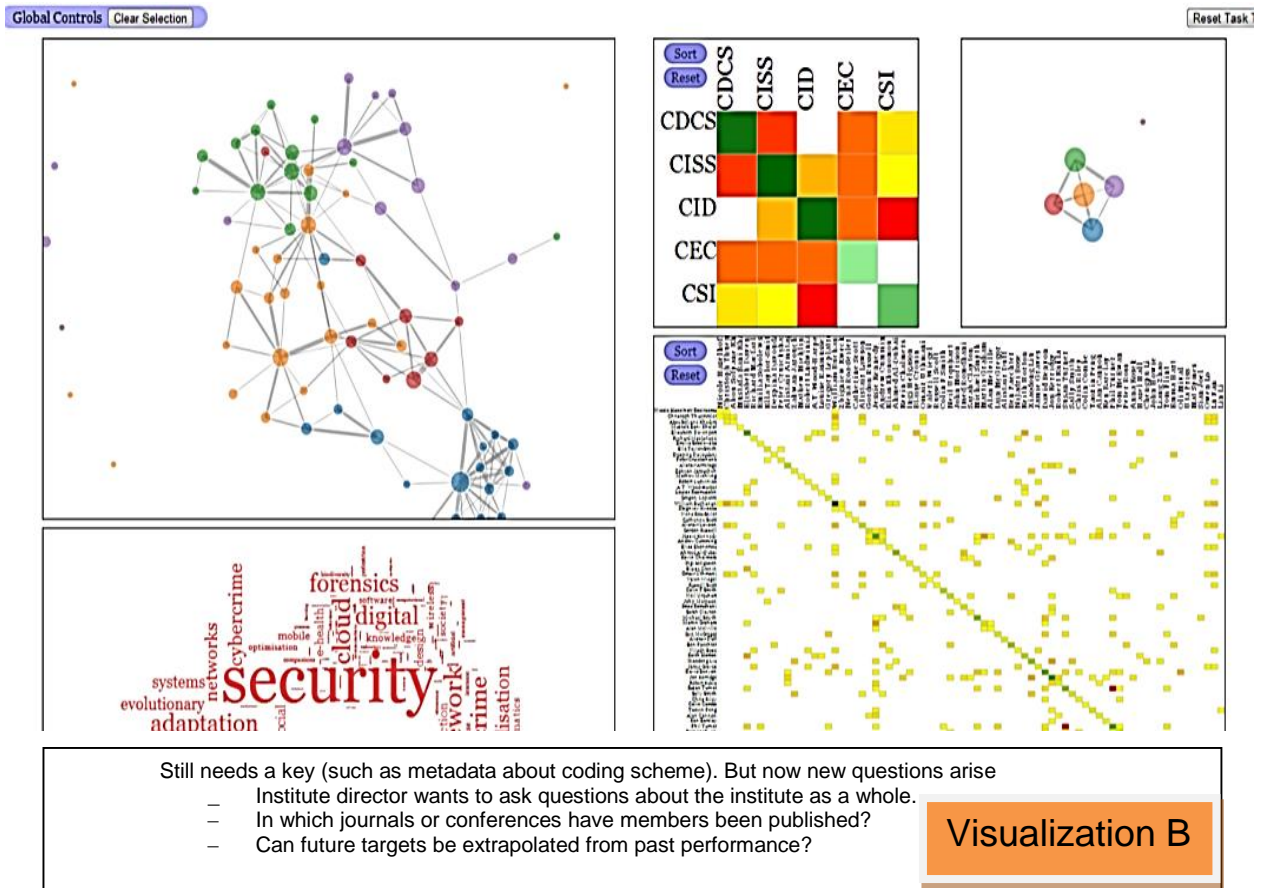


Figure 4. 2: IIDI Data CMV Interface B:

Source: [www.soc.napier.ac.uk/~CS22/test2/iidivislog2.html](http://www.soc.napier.ac.uk/~CS22/test2/iidivislog2.html)

### 4.3.1 The Experiment Design

#### Aim of Experiment

The main contribution of the initial empirical study (study 1) to the research, is that it is a pilot study, carried out in order to provide useful information to enable the researcher to design more rigorous future studies. In addition, it will serve as a platform to carry out specific testing of the Institute of Informatics and Digital Innovation CMV. It will lay the foundation to future studies by examining the extent to which the initial CMV (interface A) and the redesigned CMV (interface B) support users to carry out

successful navigation. The objective is to test the research hypothesis which suggests that providing data along with its 'summary data' in a CMV (figure 4.2) will make it easier and quicker for people to move around the information space thus achieving successful navigation, effective visualization and adequate information retrieval more easily and quickly.

The experiment will go on to examine whether provision of 'summary data' shortens the time taken to retrieve information by improving movement. In addition, the intent is to expose problems people encounter on CMVs and find possible solutions.

#### **4.3.2 Summary Data**

Summary data is explained by this research as data about the attributes that have a many-to-one relationship with the main focus of any database. In the case of the IID database, the main focus is *person*, and the attribute that have a many-to-one relationship with people is centres. Summary data in this case is data about centres. Data about centres is data about the objects in the centre, which include the publications people have made, the grants they have received and the keywords associated with publications and grants. Each of these objects has its own metadata, and together they make up an information space. Summary data is therefore data about centres, data about publications, grants and keywords. Data about centres, publications, keywords and grants is metadata about people. The availability of summary data which allows people to move through data and metadata easily and descriptive metadata which helps people search, identify, recognize and locate objects such as title, author, subjects, keywords and publication type brings about a synergy which enhances navigation.

### **4.3.3. Participants**

All male and female first year students (n=165) of computing across a broad scope of different programmes were invited to take part in the research as part of a timetabled two hour class in HCI referred to as tutorials. The participants had a number of common attributes so as to create a balance and control confounding factors such as age, gender, Internet and computing experience. About 80% of the participants were between the ages of 18 to 23, while the remaining 10 to 15% may have been outside this age range. They were all first year students and first time users of the IIDi visualization. In total 109 students turned up and 100 participated in the experiment. The participants estimated their internet experience verbally and a range of 4 to 10 years was obtained while their daily internet usage ranged from 5 to 15 hours weekly.

### **4.3.4 Experiment Setup**

The participants were placed in two separate groups and carried out the experiment over a period of four days. On the first day of the experiment, the full class was given a PowerPoint lecture, where the research supervisors presented a series of slides explaining the techniques involved and the aims and objectives of the research and an overview of what to expect during the experiment. The first group carried out the experiment on the second and third days respectively while the second group carried out the experiment on the fourth day. At the end of the experiment results were collected from each group for analysis.

The experiment setting was the Interactive Communications Environment (ICE) laboratory and participants were scheduled at specified times over a three day period.

Group one comprising tutorials 1 & 2, was scheduled for day 2 and 3, while group two comprising tutorials 3 & 4 were scheduled for day 4. The apparatus for the experiment were setup and checked prior to each session, they include an interactive table, wall screen 3 and 5 and wireless keyboards.

Four tutorials in all, with 18, 37, 17, and 28 participants in each were involved in the study. The ICE lab used for the study could accommodate up to three 'sets' (a small group of 1-4) to work in parallel. Participants were assigned 8 at a time into 20 minute timeslots, of which 10 minutes was given over to a briefing session about the lab, the nature of the research and to read and sign informed consent forms. In the remaining 10 minutes, students performed a number of predefined tasks using either interface A or B and completed a short feedback sheet. The tasks are listed below:

1. Identify and list two different things on the visualisation and write down the information they provide.
2. How many publications has member X made?
3. Locate two people in the centre for interaction design (CID) that have made joint publications with people from other centres
4. Which centre does member V, belong?
5. What subject is the centre for emergent computing (CEC) interested in?
6. How many grants have been awarded to the Centre for Social Informatics (CSI)?
7. Which individual has made the highest number of publications? Which centre does the person belong to?



8. How many mutual publications has member W made in the last three years and what topic did they address?
9. From the overview, list the five centres in the institute and identify the centre with the least number of members.
10. Which centre has the highest number of publications?

On first day, it was evident that participants found it challenging to complete the ten tasks; this led to establishing a more realistic task list. As a result, six tasks (1, 2, 3, 5, 6 and 9) were selected for tutorials 2, 3 and 4 to carry out. The scaling process involved removing tasks that were not practical, or that required a lengthy period of time to perform and do not reveal the behaviour of participants as they carry out the three navigation activities (object identification, wayfinding and exploration) identified by this research. Although tutorial 1 performed tasks 1 to 6, whilst tutorials 2, 3 and 4 performed tasks 1, 2, 3, 5, 6 and 9, the results were analysed collectively, for the reason that task 4 and task 9 are both wayfinding tasks, and therefore equivalent. In addition, a walkthrough of the tasks was carried out by the researcher and after further detailed discussion with a centre director, two institute members and a database designer, the six tasks outlined above were chosen. The tasks chosen were practical questions that an institute manager, a centre director, members and other users would seek answers to. The tasks were designed to check if users are able to navigate successfully from one view to the other and between the data presented in each view. In addition the selected tasks would reveal if users were able to identify and retrieve the meanings of the objects on the CMV, find their way to required destinations and discover relationships and patterns within the environment.

There was no control for programme of study, as experience shows that students frequently change programmes, due to a shared first year module structure. Additionally, students though assigned to specific times, were inclined to turn up earlier or later and sometimes during a different session. This was expected, and the granularity of analysis is on what the sets were able to achieve in each session. Clearly a set of four students working together on the same interface will have a different experience from a lone student working on an interface; however it was observed that while larger sets benefitted from several viewpoints, they wasted time discussing approaches.

The researcher observed the participants as they worked and made descriptive notes. Informal interviews were conducted after each session with participants who showed further interest, however the same number of people were interviewed in each group. Overall the study hoped to capture the following:

- The task type user was able to complete
- Task completion times
- Navigation pattern (number of clicks, mouseovers and events)
- User satisfaction.

#### **4.3.5. Between Subjects Experiment Design**

The between subjects design chosen for this experiment comprise of two groups of participants who performed tasks (p.131) on the IIDl data CMV. The control group worked with Interface A, a condition where no special treatment was introduced (Bailey, 2008). The treatment group (Hinkelmann and Kempthorne, 2008) worked with Interface B, where a special condition, the summary data was introduced.

The procedure involved a basic design with one independent variable and two test groups which corresponds to the number of experiment conditions (Interface A and B), therefore, the between subjects design was selected. Participants were exposed to only one experiment condition in order to control factors such as the learning effect described by Lazar, Feng and Hochheiser (2010) as a situation where participants learn from separate task condition.

Ten tasks involving the three categories of navigation activities identified by this research; object identification, exploration and wayfinding were selected for the experiment. The intent of the experiment is to evaluate the level of support offered to users by each CMV (interface A & B) to support the execution of different tasks people may want to undertake. The degree of support participants receive will reflect in their ability to navigate the interface successfully, their ability to visualise and retrieve adequate information, these will be measured by the tasks they accomplish, the navigation pattern observed and the time taken. In addition, the type of navigation the information space supports will be measured by the category of tasks more participants complete. The actions participants undertake will be recorded by means of a log file (Log4javascript). The information recorded by the log file will give more insight into the relationship between certain features on CMV information space and the navigation experiences of users.

The log file records the actions taken by participants as they navigate through the IIDI visualization information space. The column on the right of the log file represents the type of information retrieved. The left column illustrates the category of event involved,

and the middle column shows the time taken to retrieve required information. In order to obtain relevant information, the data was looked at from the perspective of the three navigation activities relevant to this research (object identification, wayfinding and exploration) in order to evaluate the time, the events and the number of clicks required to complete a specific task.

#### **4.3.6 The Log File**

The raw data was collected by means of a task sheet and a log file. The task sheet holds records of information retrieved by participants while the log file records the actions participants take as they work through the task. These actions illustrate how they move around the different components on the interface and the length of time they take. It shows the length of navigation steps involved on the IIDI CMV, by providing details of the events executed, the number of clicks and mouseovers made and the navigation pattern (route) of participants. The Log files allow the researcher see navigation between views, and events that take place. It is also useful to help figure out when participants are searching, browsing or selecting. A sample of the actions of some participants retrieved from the log file is illustrated below:

14:55:09.902	161.48	Selected	Matrix CEC and CID
14:55:10.808	162.39	Mouseover	Matrix CID and CID
14:55:11.389	162.97	Mouseover	Matrix CSI and CID
14:55:11.927	163.51	Selected	Matrix CSI and CID
14:55:12.805	164.39	Mouseover	Matrix CID and CID
14:55:13.541	165.12	Selected	Matrix CID
14:55:15.253	166.83	Mouseover	Matrix CISS and CID
14:55:15.714	167.29	Selected	Matrix CISS and CID
14:55:16.550	168.13	Mouseover	Matrix CID and CID
14:55:16.927	168.51	Selected	Matrix CID
14:55:17.538	169.12	Mouseover	Matrix CEC and CID
14:55:18.130	169.71	Selected	Matrix CEC and CID
14:55:18.815	170.40	Mouseover	Matrix CID and CID
14:55:19.017	170.60	Mouseover	Matrix CISS and CID
14:55:19.232	170.81	Mouseover	Matrix CID and CID
14:55:24.389	175.97	Mouseover	Matrix CISS and CID
14:55:25.381	176.96	Mouseover	Matrix CISS and CID
14:55:42.318	193.90	Mouseover	Matrix CISS and CID
14:55:42.987	194.57	Mouseover	Matrix CID and CID
14:55:43.196	194.78	Mouseover	Matrix CISS and CID
14:55:43.591	195.17	Selected	Matrix CISS and CID
14:55:45.689	197.27	Selected	Matrix CISS and CID
14:55:47.777	199.36	Mouseover	Node-link CDCS
14:55:48.198	199.78	Mouseover	Node-link CID
14:55:49.123	200.70	Mouseover	Node-link CSI
14:55:49.989	201.57	Mouseover	Node-link CID
14:55:50.437	202.02	Selected	Node-link CID
14:55:52.547	204.13	Selected	Node-link CID
14:55:52.903	204.48	Selected	Node-link CID
14:55:53.270	204.85	Selected	Node-link CID

Figure 4. 3: Log file extract for Interface A, tutorial 1

#### 4.4 Quantitative Results

The first experiment established the average number of tasks that future sets might realistically achieve in the allotted time as 6 tasks. It looked at the four parameters outlined above. Task completion times were obtained by measuring the time it took participants to complete the given tasks on each interface. The navigation pattern is measured by the type of event undertaken and the number of clicks and mouseovers made for each event. The type of event indicates the route while the clicks and mouseovers determine the length of navigation steps. The task completion is measured by the type of tasks and number of tasks completed.

An informal interview was conducted after each session. The participants were asked six open ended questions which were directed at understanding how much support

they received from the interface, the experiences they had carrying out the tasks, and how satisfied they were with the interface.

#### 4.4.1 Type of tasks completed

Interface A: Tutorial 1

Task number	Task Type	Number of Participants that completed Task
1	Object Identification	14
2	Wayfinding, and Exploration	12
3	Exploration	11
4	Wayfinding / Object Identification	11
5	Object Identification, Wayfinding and Exploration	8
6	Object identification, Exploration	6
7	Wayfinding	0
8	Wayfinding and Exploration	0
9	Object Identification and Wayfinding	0
10	Wayfinding and Exploration	0

Table 4. 1: Task type completed by tutorial 1 on Interface A, (n=18)

Interface A: Tutorial 2

Task Number	Task Type	Number of participants that completed task
1	Object Identification	25
2	Wayfinding and Exploration	23
3	Exploration	19
4	Object Identification / Wayfinding	13
5	Object Identification, Wayfinding and Exploration.	2
6	Wayfinding / Object Identification	2

Table 4. 2: Task type completed by tutorial 2 on Interface A, (n=37)

	Task Type	Number of participants that completed Task.
1	Object Identification	16
2	Wayfinding and Exploration	16
3	Exploration	16
4	Object Identification and Exploration	16
5	Object Identification, Wayfinding and Exploration	13
6	Wayfinding and Object Identification	13

Table 4. 3: Task type completed by tutorial 3 on Interface B, (n=17)

Interface B: Tutorial 4

Task N.	Task type	Number of Participants that completed Tasks
1	Object Identification	22
2	Wayfinding and Exploration	28
3	Exploration	28
4	Object Identification and Exploration	23
5	Object Identification, Wayfinding and Exploration	24
6	Wayfinding and Exploration	22

Table 4. 4: Task type completed by tutorial 4 on Interface B, (n=28)

The results illustrated in Tables 4.1, 4.2, 4.3 and 4.4 suggest that the revised interface B supports participants to carry out the tasks better than the original interface A. The revised interface B resulted in completion of more tasks, whether measuring by individuals or by sets. The smaller tutorial groups (table 4.1 & 4.2) had a higher task completion rate than the larger tutorial groups. Although we ensured that the ten minutes per test session was strictly adhered to, this variation may be as a result of a more hectic through-put of students in the two larger tutorial groups. This could have brought about distractions or disoriented the participants in minor ways.

Table 4.5, below is a comparative analysis of the type of task completed by participants. It gives an idea of the type of navigation activity the interface supports.

Task Number	Task type	Interface A N = 55	Interface B N = 45



1	Object Identification	39	38
2	Wayfinding& Exploration	35	44
3	Exploration	30	44
4	Object Identification & Exploration	24	39
5	Object ID/exploration/wayfinding	10	37
6	Wayfinding	8	33

Table 4. 5: Comparative analysis of type of tasks completed by participants

Task 1 was designed to find out if participants could effectively identify any object on the interface such as the node-link, keywords or scatterplot, and interpret the information it represents. Table 4.5 shows that participants who worked on interface B performed better in the task. Interface A recorded its highest performance in task 1. It can be deduced that both Interfaces support the object identification activity

Task 2 involves two types of navigation, wayfinding and exploration. Participants were asked to find the number of publications made by a member. The wayfinding aspect of this task was to examine whether participants could find out how many publications member X had made. The exploratory aspect was to check whether participants could find the relationship between member X and member X's publication. The results obtained indicate that interface B supported more participants to complete the wayfinding tasks. Participants pointed out that the wayfinding part of the task was more challenging than the exploratory part. According to a female participant 'once you find your way, it is possible to locate relationships'.

Task 3 is an exploration task and its objective is to examine whether the interface supports participants to explore and find existing relationships and patterns within the environment, in this case two members from the centre of interaction design (CID), who have made joint publications with members from other centres. 30 of the 55 participants from Interface A were able to carry out this task with most of them listing names of CID members. During the interview many of the participants pointed out that “retrieving the information about joint publications from the matrix (lower right view) and nodelink (upper left view) required a lot of clicking“. Some admitted they could locate CID members but could not relate them to mutual publications. They explained they performed the mouseover event on a number of nodes and matrixes clicking on some to locate members of the centre for interaction design. The participant’s that were unable to retrieve sufficient information argued that they could not make out any clues to help them find publishing relationships. 44 out of 45 participants on Interface B retrieved names of members of the centre for interaction design (CID). In course of the interview, some participants revealed that information provided on the upper right view on Interface B interface helped them locate the members of the CID.

Task 4/9 involved object identification and exploration. The aim of the task is to see whether participants can identify the different centres in the institute and locate the centre a specific member belongs to. Table 4.5 demonstrates that not many participants who worked on Interface A were able to perform this task; Interface B however showed that participants were able to perform the task. Although the tasks were different, they shared a common objective which is wayfinding and object identification, which is the basis for comparing the data.

Task 5 involved all three types of navigation identified by this research. Object identification, wayfinding and exploration. Participants were required to identify the centre for Emergent Computing (CEC), and examine if they can discover relationships between objects (centres) and their subject of interest (keywords). 10 participants attempted this task on Interface A but no relevant data was recorded as participants were unable to complete the task. Some of them said they could not find the required information while others said they needed more time to learn the interface. On Interface B; 37 participants retrieved relevant information while a few were unable to find the required information.

Task 6, a wayfinding and exploration task asked participants to find the number of grants awarded to the centre of social informatics (CSI). This task would check whether participants could locate the centre and find the number of grants the centre has received. Participants on Interface A were unable to complete this task. The participants commented that the navigation steps involved in retrieving information about grants was tedious. According to a participant who commented during the experiment, "Too many clicks and lots of calculation required to perform this task". This task involved aggregating a lot of data. First, participants have to click on several nodes in the upper left view to find the colour coded for CSI, next they have to click on the nodes with the CSI colour to find members of the CSI. Next, the numbers of grants awarded to individuals in CSI have to be retrieved from the nodes and added up to get the total number of grants. Some participants in Interface B were able to retrieve this information because data was provided along with a summary data on grants on the upper centre display (UCV) and upper right view on Interface B. The number of clicks made on Interface B was less, demonstrating a reduced number of navigation steps.

#### 4.4.2 Task completion times

Results obtained from the log file for task completion times on both Interfaces are illustrated below on table 4.6.

Task	Interface A(tutorial 1)	Interface A (tutorial 2)	Interface B(tutorial 3)	InterfaceB(tutorial 4)
	Time Taken (secs)	Time Taken (secs)	Time Taken (secs)	Time Taken (secs)
1	159	162	72	123
2	339	163	143	145
3	110	122	117	52
4	141	141	57	50
5	91	114	78	84
6	N/A	N/A	N/A	N/A
	<b>840</b>	<b>572</b>	<b>467</b>	<b>454</b>

Table 4. 6: Comparative analysis of task completion times by tutorials

Table 4.6 shows that for task 1, Interface A, tutorial 1 and 2 have task completion times of 159 and 162 seconds. On Interface B, tutorial 3 and 4 had task completion times of 72 and 123 seconds. This suggests that task 1 took less time to perform on Interface B in comparison to Interface A.

The log file shows that for Task 2, task completion times for tutorial 1 and 2 on Interface A, were 339 and 163 seconds, while for tutorial 3 and 4 who worked on interface B, it took 143 and 145 seconds. The results show that less time was required to complete task 2 on interface B.

The results for task 3 show task completion times for tutorial 1 and 2 on interface A were 110 and 122 seconds and on Interface B, tutorials 3 and 4 completed the task in 117 and 52 seconds.

For task 4, Interface A; tutorial 1 and 2 had task completion times of 141 seconds each and on interface B, tutorials 3 and 4 had task completion times of 57 and 50 seconds.

For task 5, Participants in tutorials 3 and 4 who worked on Interface B had task conclusion times of 78 and 84 seconds respectively while tutorials 1 and 2 who worked on Interface A spent 91 and 114 seconds respectively.

#### **4.4.3 Navigation Pattern Analysis**

The navigation pattern encompasses the interface events which give an insight to the actions undertaken by participants and point to the routes followed by each group as they navigate the information space, visualise and retrieve adequate information. Table 4.7 is an example of the data collected for navigation events and the time spent on each event for a given task. The time spent on each event is computed to obtain the task conclusion time for each event. The pattern of events also demonstrates the navigation routes the participant takes on the interface. However, this research will focus more on the number of mouseovers and clicks in order to illustrate the number of steps participants take to complete the tasks, and to determine the interface with faster routes which could reduce disorientation and navigation times. This information is vital in the design of CMVs of visualization when deciding on types of data to provide on information spaces.

A number of participants claimed “it took a while for them to make a connection between the names listed in lower right view and the nodes located in upper left view of the IIDI data visualization”.

[RS] – 4 sec, [TS] – 9 sec, [MN] – 1 sec, [SN] – 5 sec, [MN] – 7 sec, [SN] – 2sec,  
[MN] – 7 sec, [SN] – 2 sec, [MN] – 12 sec, [SN] – 3 sec, [MN] – 1 sec, [SN] – 11 sec,  
[MN] – 2 sec, [SN] – 8 sec, [MN] – 7 sec, [MM] – 2 sec, [SLM] – 17 sec, [MN]- 1 sec,  
[MM] – 3 sec, [SLM] – 3 sec, [MM] – 45 sec

Figure 4. 4: Navigation Pattern illustration for Interface A, Tutorial 2

#### **4.4.4 Clicks Analysis**

This research describes clicks as the steps involved in navigation. It is similar to steps such as movements in different directions made by people in the physical world as they search for their destination. The result obtained from the experiment suggests that the tasks that record more clicks compelled the participants to make a higher number of movements around the information space. Results demonstrate that participants who performed tasks on Interface A retrieved less information, spent more time and made more clicks than participants who performed tasks on Interface B. The highest record of clicks in this experiment was made by Interface A; tutorial 1.

This result suggests that where navigation activities such as object identification, wayfinding and exploration are incorporated by CMVs of visualization, purposeful movement will be supported and successful navigation will be achieved. It will enhance the ability of people to perform tasks. In addition, the provision of summary data will support people retrieve adequate data with faster completion times. This is evident

from the results obtained from participants who carried out tasks on Interface B. Overall; they retrieved more adequate information, made less clicks pointing to successful navigation and had faster completion times. The number of clicks made indicate the number of steps required to navigate to find and retrieve information. A CMV where people require less navigation steps to carry out a task will require less time to navigate through an information space to retrieve information while on the other hand; one that requires more navigation steps will take more time to achieve a task.

Navigation Events	Interface A		Interface B	
	Number of Clicks		Number of Clicks	
	Tutorial 1	Tutorial 2	Tutorial 3	Tutorial 4
Sorted Matrix (SM)	102	114	31	43
Selected Matrix (SLM)	321	489	219	260
Selected Nodelink (SN)	163	161	152	135
Reset Selections (RS)	43	60	25	16
Task Start All (TS)	39	52	16	27
Mouseover Matrix(MM)	812	717	501	542
MouseoverBarchart(MB)	93	117	2	6
MouseoverNodelink(MN)	465	376	257	314
Filtered Year Barchart (FYB)	14	49	N/A	N/A
Filtered Publication Barchart (FPB)	78	105	N/A	N/A
Total Number of Clicks	2130	2240	1203	1343

Table 4. 7: Cross section of number of clicks by event made by tutorial groups

The results in table 4.8 illustrate the total number of clicks made by each tutorial group for each event for all the tasks they carried out. It was derived by computing the number of clicks made for each event from the records obtained from the log file. The selected matrix (SLM) event which occurs in the lower left view featuring the matrix recorded the highest number of clicks amongst other events. This suggests that the lower left view took more time to navigate. This theory was corroborated by a number of participants who pointed out that “they did not understand the information illustrated by the matrix”.

The results indicate that tutorials 3 and 4 made less clicks across all events than tutorials 1 and 2. This points to the probability that the revised Interface B required less navigation steps than the original Interface A. it therefore demonstrates that a reduction in navigation steps could affect navigation behaviour of participants by supporting focused movements and reducing the time it takes to navigate a given information space.

#### **4.4.5 Discussion**

The four separate tutorial groups that performed tasks on both interfaces A and B were made up of irregular sets. Some sets were made of 2 – 4 people and occasionally one individual. However the results show that the sizes of sets appear to have had no significant influence over the type of task completed.



The number of mouseovers obtained from the log file shows that participants had to do a lot of exploration during the experiment.

The task type with the highest completion rate on Interface A was object identification and the least was a wayfinding task (task 6). Interface B, recorded its highest completion rates on wayfinding and exploration tasks, interestingly it also had its lowest completion rate in a wayfinding task (task 6). However Interface B recorded higher completion rates than interface A for both wayfinding tasks (2 and 6). 44 out of 45 and 35 out of 45 participants who worked on interface B completed task 2 and 6 respectively while 35 out of 55 and 8 out of 55 participants on Interface A completed task 2 and 6 respectively, creating the wide variance. This suggests that Interface B offered the participants more support to navigate the interface successfully.

Interface A, supports a number of user tasks but its design lacks a key to the coding method employed to encode centres or to make out what that the lines linking the nodes in the upper left view indicate co-publishing relationships and that the thickness of the line illustrates the number of mutual publications. The upper right view provides a coding scheme which is why many participants named it as the easiest view to navigate.

On interface A, Participants had to make numerous movements, aggregate data themselves and make calculations, whereas on interface B, less exploration, and calculation was required. Participants did not need to aggregate data themselves, which contributed to faster task completion times being obtained. This is credited to the summary data present on interface B which provided a type of metadata absent on interface B.

## **4.5 Qualitative results**

The results obtained from the brief semi structured interview carried out with participants are reported below. The interview was recorded by the researcher on a printed sheet with predetermined answers and spaces provided to write additional answers and observations. Names of interviewees were not recorded as they preferred anonymity.

### **4.5.1 Informal Interview**

A semi structured interview was conducted with about 40 participants. 10 participants were interviewed from each tutorial in order to develop a balanced comparative analysis. A semi structured approach was employed in the interview so as to allow the researcher run a set of six queries comprising open-ended questions. The open-ended questions encouraged a two way conversation and allowed the researcher gain insights and understanding (Gillman, 2000, p.11; Ritchie & Lewis 2003, p.138) on specific issues such as navigation. In addition, interviewees made useful suggestions, both positive and negative, that led to the exploration and development of some aspects of the guidelines. Therefore, the data collected was categorised as positive, negative or neutral for all the questions excluding question 5, whose objective was to find out the interviewees opinion about the tasks, whether the given tasks were difficult or achievable. The data collected from questions was used to measure respondent's evaluation of their actual navigation experience on interface C and helped the researcher understand the general idea they had developed about the CMV. The knowledge acquired hopefully would help in the designing of the CMV interface. The Open-ended Questions are outlined below

1. From your experience using the interface, do you believe that the interface supported you to find your way around the different views efficiently enough for you to visualise data and retrieve adequate information?
2. Can you please tell me about one feature that offered you any support on the interface? Are you saying this feature was relevant to the tasks you undertook?
3. What do you think about the objects spread around the interface? Were you able to identify them? Are you saying you understood what information they represented?
4. The tasks you just completed will evaluate three basic activities: the ability to find your way around, explore the interface and identify objects. Would you say you were able to achieve at least two of these activities?
5. Can you give me an example of any task you recall? Possibly as a result of how easy or challenging it was to complete?
6. Would it be fair to say you had a satisfying experience navigating the IIDI CMV?

#### **4.5.2 Discussion**

The post session informal interview comprised six open-ended questions that sought participant's opinions about both Interfaces. It touched on issues such as navigation activities, gaining an overview, visualization and user experience. The responses obtained from users were broadly categorised into positive, negative and neutral comments and are also reported in detail.

Interface A was rated less favourably by interviewees. In general, participants said "they did not receive enough support from the interface to visualize data or retrieve sufficient information". A few participants noted that they had to carry out calculations

in order to answer question 6. About 7 participants thought they received little support while 3 took a neutral stand. On the other hand, interface B was rated more favourably by participants. 18 interviewees rated it higher for providing more support to facilitate navigation between views in order to visualise and retrieve information. They also noted its support to their being able to complete the tasks.

On interface A, about 14 of the 20 interviewees indicated that the key to the coding system on upper right view of interface A, (fig 4.1) supported their understanding and interpretation of the information represented in the bar charts. Also, 16 out of 20 interviewees on interface B said the upper centre view and upper right view of interface B (fig 4.2) was fundamental to their gaining understanding of the interface. They explained that the upper right view supported them to retrieve information from the upper left view.

The response obtained for question 3 reveals that more interviewees were able to identify and understand the information represented by objects on interface B in contrast to interface A.

Interviewees on interface A agreed they were able to identify objects more easily than find their way around the interface in order to explore data whereas interviewees on interface B said they were able to carry out all three navigation activities reasonably well.

As regards the question about tasks (question 5) most interviewees for interface A, said they understood the tasks but could not execute them as they could not find the

required information and therefore were “unable to proceed beyond task 5” which resulted in the reduction of tasks from 10 to 6.

More interviewees who worked on interface B said they had a more satisfying experience navigating the interface than the interviewees on interface A, who said “it was quite challenging deciding where to go”.

Question	Interface A			Interface B		
	Positive Comments	Neutral Comments	Negative Comments	Positive Comments	Neutral comments	Negative Comments
1	7	3	10	18	N/A	2
2	14	4	2	16	N/A	4
3	11	4	5	17	N/A	3
4	5	3	12	17	N/A	2
6	4	3	13	17	N/A	3

Table 4. 8: Broad categorization of responses from open-ended questions

Participants agree that the easiest view to navigate on interface A was the upper right view. They claim that the key supplied facilitated their ability to understand the information the view represented. The information retrieved by participants who worked on interface A for task 1 is evident that the information displayed on the upper right view of interface A was clear. Table 4.9 represents the broad categorization of responses obtained from the open-ended questions.

## **4.6 Conclusion**

The results suggest that the availability of a coding scheme supported participants to navigate the upper right view of the interface which means that metadata describing the coding system is a significant piece of information which was helpful to participants navigating the interface. Interface B supported the participants to navigate through the data set more easily. It is reasonable to state that design guidelines should include 'first level' metadata about an object in terms of a key.

The subsequent chapter describes a second empirical work which provides a key to the coding scheme and more levels of summary data.

## **Chapter 5 Empirical Study Two**

### ***5.1 Introduction***

The first empirical study discussed in chapter 4 was a preliminary laboratory based experiment which showed that the CMV that supports metadata and encompasses navigation activities such as object identification, wayfinding and exploration appeared to improve task completion times and ability of people to navigate successfully compared to the CMV which does not.

This chapter presents the second empirical work. It investigates the research hypothesis which states that “when data is provided along with its summary data on CMVs, successful navigation can be achieved by people”. On the other hand the null hypothesis states that the summary data would have no effect on successful navigation of CMVs. The research is designed to investigate two independent variables, the summary data and interface key in order to test the hypothesis put forward by this research.

### ***5.2 Research Methodology***

The method chosen for the second empirical work include qualitative and quantitative methods. A controlled experiment involving fewer participants will be performed. Observational studies which will allow the researcher collect descriptive and reflective notes will be undertaken and semi- structured interviews will be conducted.

### **5.2.1 Experimental Investigations**

A controlled experiment was chosen for these empirical work to allow the researcher establish the relation between a set of factors, in this case, navigation activities and the phenomenon in this instance, successful navigation (Kirk, 1982; Oehlert, 2000). The experimental situation will be manipulated to allow the researcher compare two or more conditions whilst other factors are kept constant. The effects on the variables significant to the research will be reduced by controlling the variables that are not of interest.

### **5.2.2 The Variables**

The independent variables in this instance are the summary data and Interface key. The dependent variables in this experiment include the parameters outlined below:

- The task type user was able to complete
- The number of clicks involved in each event
- Time it takes to complete each task.
- User satisfaction

### **5.2.3 Descriptive Investigations**

Interviews will be utilized as descriptive investigation for purposes of this empirical study. The interviews conducted are to allow the participants express their opinions and suggestions about the redesigned interface and enable the researcher determine user satisfaction.



### **5.3 The IIDI Data CMV**

The IIDI data was initially visualised using the Author Network Graph (ANG) in 2011. It was updated and renamed the IIDI data visualization in 2013. It was designed to allow people visualise data about the institute of Informatics and Digital innovation in the computing department of Edinburgh Napier University. It is a CMV (Interface A) designed to allow people see and interpret the IIDI data from different perspectives. It consists of data about the institute, the centres belonging to the institute, members of these centres and publications that have been made. The purpose of the empirical study is to comparatively examine the extent of support offered by the redesigned IIDI CMV, interface C to people to navigate the interface in order to visualize and retrieve information effectively. Interface C is designed to incorporate navigation activities such as object identification, exploration and wayfinding and summary data which is absent on Interface A.

#### **5.3.1 The IIDI Data Interfaces**

The interfaces (figures 5.1, 5.2 and 5.3) as discussed in chapter 3 consist of four (interface A) or five (interface B & C) views. The upper left views of the interface consist of nodelinks colour-coded by centre, and the nodes are linked by lines showing co-published relationships. The thickness of the lines indicates the quantity of co-publications. The lower left views word cloud displays keywords, whose size depends on the frequency of use. The upper right view of Interface A, displays a bar chart which explains publication by year on the X-axis and by type (colour-coded) on the Y-axis whereas the URV of interfaces B and C display a set of five linked nodes colour coded for centres and the Institute. The UCV is a matrix of centres (colour-coded) and the

lower left view of all three interfaces is a matrix of people showing the co-publishing relationships.

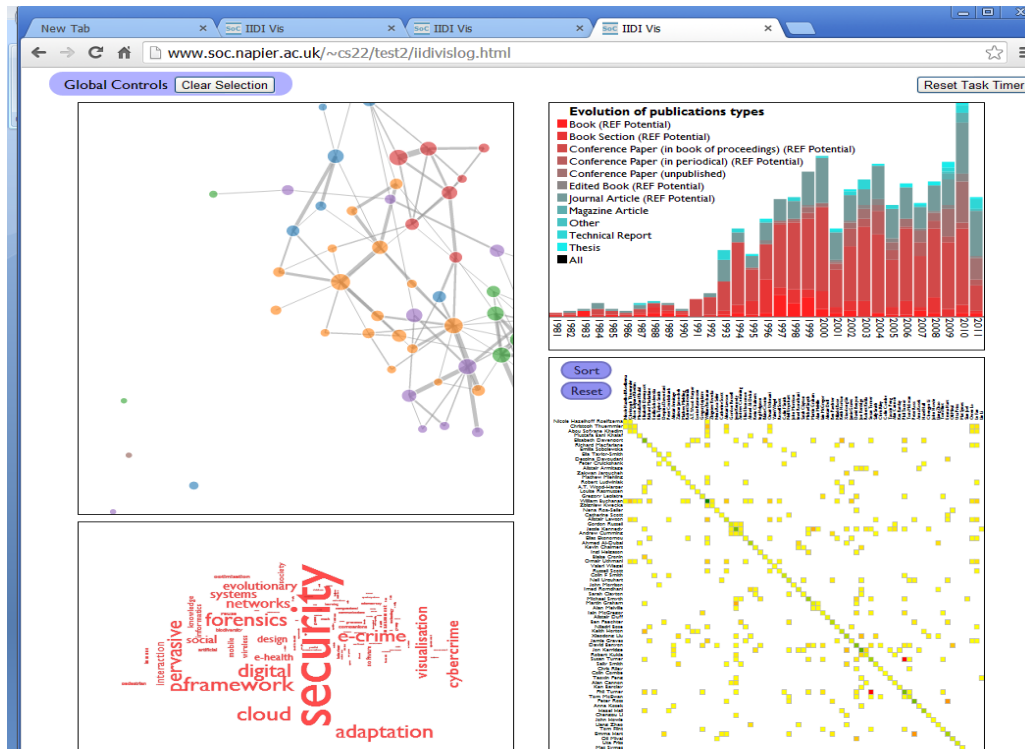


Figure 5. 1: IIDI Data CMV Interface A

Source: [www.soc.napier.ac.uk/~cs22/test2/iidivislog.html](http://www.soc.napier.ac.uk/~cs22/test2/iidivislog.html)

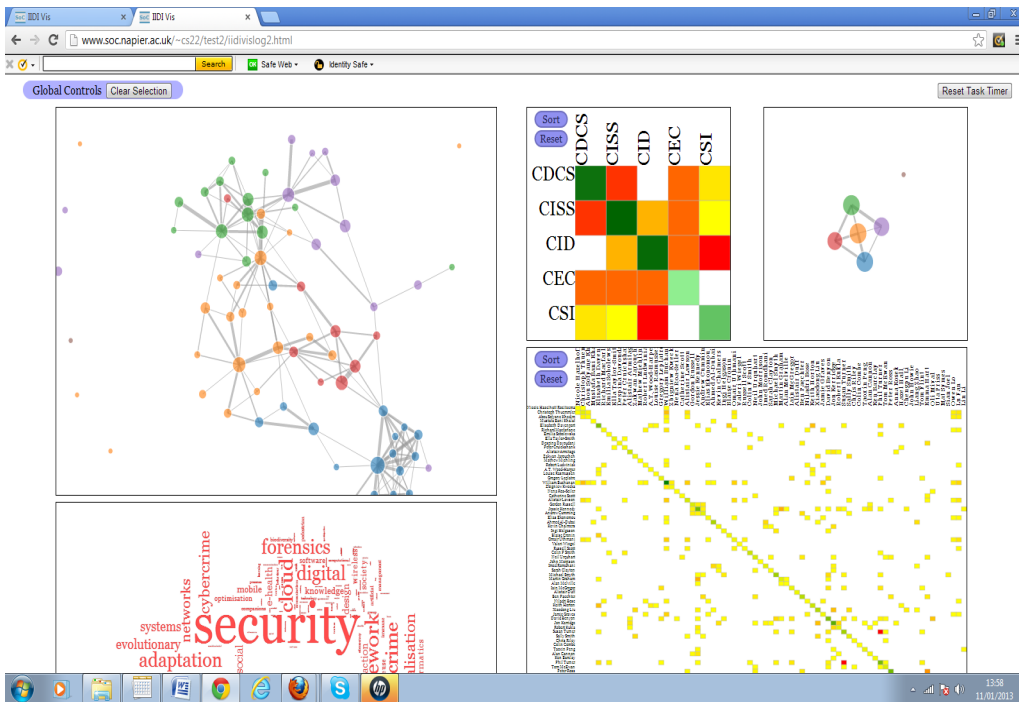


Figure 5. 2: IIDI Data CMV Interface B

Source: [www.soc.napier.ac.uk/~CS22/test2/iidivislog2.html](http://www.soc.napier.ac.uk/~CS22/test2/iidivislog2.html)

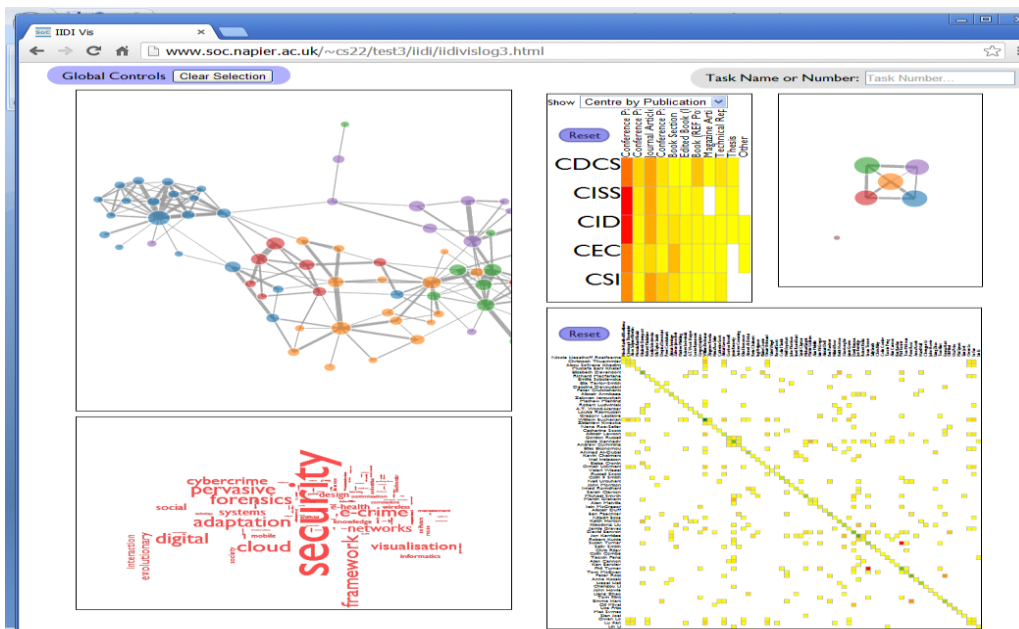


Figure 5. 3: IIDI Data CMV Interface C

Source: [www.soc.napier.ac.uk/~cs22/test3/iidi/iidivislog3.html](http://www.soc.napier.ac.uk/~cs22/test3/iidi/iidivislog3.html)

The variation between interface A and the other two interfaces B and C lies in the Upper Centre View (UCV) and points to an important observation. The upper right view shows evolution of publications types made across the institute from 1981 through 2011, and gives a description of the coding system used for types of publication. Participants in empirical study one said the coding system of publication types supported them to explore and understand the information in the URV. They were mostly of the opinion that it was the easiest view to navigate through and understand. They said the coding system explains the data represented by the bar chart and gives insight into the publication types making navigation easier and swift. This led to the development of the fourth design guideline which is to provide a key to the coding system of the visual aspects of the CMV, known as *interface key* by this research. This was then made available to be evaluated in the second study.

On the other hand, the variation between interfaces B and C is that the UCV in interface C links to metadata about people such as data about centres, publication and grants which improved the navigability aspect of the CMV.

### **5.3.2 The Experiment Design**

The first empirical study reported in Chapter 4 was a preliminary study carried out to lay the foundation for future empirical studies by examining the extent to which the CMVs Interface A (fig 5.1) and B (fig 5.2) support users to carry achieve successful navigation, effective visualization and adequate information retrieval.

The results obtained from this initial study suggest that providing a key to the coding system of the IIDI CMV gave the participants useful information that helped them

navigate through the interface. This gave rise to developing the fourth guideline (provide an interface key) which is metadata about the visual aspects of the interface.

The second empirical study is designed to be a more meticulous experiment to permit evaluation of the IIDI Data CMV under controlled conditions in order to identify how users experience difficulty in navigating information spaces and the reason behind it. It will also examine whether providing summary data on a CMV (figure 5.3) shortens the time it takes to retrieve information by improving movement. Furthermore, the experiment explores the degree to which the summary data provided on Interface C support people to identify objects, explore and find their way in order to achieve successful navigation, visualize data and retrieve information.

### **5.3.3 Participants**

Twelve participants were involved in this experiment, 6 male and 6 female, with an age range of 18 to 38. Eight participants were registered students of the computing department of the Edinburgh Napier University, undergraduates (4) postgraduates (2), research students (2), university staff (2) and administrative staff from other organizations (2). The participants estimated their internet usage at about 3 to 10 hours daily and their internet experience ranged from 4 to 16 years. They were all first time users of the IIDI interface.

### **5.3.4 Experiment Setup**

The participants worked on alternate interfaces and were given 30 minutes break in between to reduce the errors caused by fatigue and practice in a within subjects design. The experiment took place in various laboratories in the Merchiston and

Craiglockhart campuses of Edinburgh Napier University. Participants sat on standard computer tables and worked individually with desk top computers. Each participant was allotted 20 minutes for each session timed from when participant commenced the experiment. The first set of 4 participants worked in Merchiston, room D50 on the first day. A second set of 4 worked in Craiglockhart library on the second day. The third set of 2 participants worked in Merchiston, room C44 on the third day and the last set of 2 participants worked in the Merchiston library on the fourth day. The results were recorded and collected by means of a log file. Each session lasted 20 minutes; participants worked on a set of predefined tasks and completed the feedback forms provided. On the day of the experiment, participants were briefed about the aims and objectives of the research and what the experiment involves with the aid of the same power point slides used to introduce the first empirical study. The briefing lasted 20 minutes including time for questions and answers.

The experiment lasted twenty minutes, during the first 10 minutes, participants read and signed consent forms and familiarized themselves with the interface and in the subsequent 10 minutes, the experiment was performed. Ten specific tasks designed to measure participant's ability to achieve successful navigation were performed. The tasks represent the three navigation activities (object identification, wayfinding and exploration) that support the navigation of information space approach adopted by the research to redesign the IIDl CMV as illustrated by interface C, (fig 5.3). The objective of the tasks is to show if these navigation activities enhance participant's ability to achieve successful navigation, visualize data effectively and retrieve adequate data faster and more easily. The tasks performed are listed below;

1. Identify two objects on the given interface and write down the information they provide.
2. How many publications were made by CID in 2010? What was the subject of the publication?
3. Locate and list two centres that have made joint publications. Write the names of the people involved in this publication.
4. Which centre has been awarded the highest number of grants? Name the member who received it?
5. Identify the centre which focuses on networking. In which year did it make its highest number of publications?
6. Locate member Y. Which centre does the member belong to? How many publications did the member make in 2009? How many grants has the member received?
7. From the overview, list the five centres in the institute.
8. Who is the head of the CEC?
9. List four types of publications made by the institute.
10. Name the centre with the least number of grants.

### **5.3.5 The Interface key**

Following the experience of participants in empirical study one with the URV of Interface A (figure 5.1) whose coded system had a key that helped participants understand the information, an interface key explaining the coded system of Interface C was created. The interface key explained the coded system of the IIDI CMV to

facilitate participants understanding. The Interface key supplied to participants for interface C is given below:

- The upper left view (ULV)
- Colour coded nodes – portray members of various centres
- Size of nodes – illustrate the number of publications.
- Node links – show co-publishing relationships.
- Thickness of links – indicates volume of publications
- The lower left view (LLV)
- Word cloud – keywords indicating subjects and area of interest of centres and people
- Size of keywords – illustrates frequency of use
- The upper right view (URV)
- Colour coded linked nodes – represent centres
- Colour – depicts specific centre
- Solitary node – represents the institute
- The lower right view (LRV)
- Matrix of people – shows co-publishing relationship
- The upper centre view (UCV)
- Matrix of centres – collaborations between centres

### **5.3.6 Within Subjects Design**

The within subjects design allows the same group of subjects to undergo more than one treatment. It was used for the second experiment which had fewer volunteers than the first to ensure that although the number of participants were reduced, the total number was sufficient to bring about statistical significance and reduce errors.



According to Richard Hall (1998) a fundamental inferential statistics principle is that, as the number of subjects increase, statistical power increases, and the probability of beta error which is the probability of not finding an effect when one is in existence decreases. In addition when the same people work under two separate conditions, factors like memory skill, meticulous nature and ability to concentrate that bring about individual differences which affect dependant variables in the between subjects design are reduced as the situations concerned with individual difference variable is almost the same. A participant who is meticulous will probably demonstrate this nature in both situations.

Six participants worked first on task sheet 2 and then on task sheet 1, this was termed the C- A order. The other six participants worked first on task sheet 1 and next on task sheet 2, termed the A-C order. A set of 10 tasks with questions originating from the three categories of navigation activities were given to the participants and they were given the option to opt out of the experiment at any time. All participants completed the experiment. Data was gathered through the log file and the post interview session.

### **5.3.7 Task development**

The initial tasks were developed through a focus group where users came up with queries they wanted to answer on the IIDI visualization. The tasks were split into 3 categories: object identification (tasks that represented identification such as what? Where?; wayfinding (tasks that involved location) and exploration (tasks that required comparison) to enable the research identify to what extent the CMVs support people to navigate in order to visualize and retrieve adequate information. The tasks were grouped based on the type of navigation activity required to complete the task. The

tasks selected from the pool of questions obtained from the focus met the following conditions:

1. To represent typical questions an institute manager, centre director, members and other people interested in the IIDl Data would want to answer on a CMV of a management information system.
2. To demonstrate that the type of data provided on an information space can improve navigation of CMVs.
3. To illustrate the relevance of navigation activities in achieving successful navigation, effective visualization and adequate information retrieval on CMVs.

The objective of the task is to allow the researcher compare the effect of the summary data and interface key on both interfaces. In addition, it will make it possible for the researcher to identify trends in the participants' navigation behaviour. For instance, an improvement in the time it takes a participant to navigate the IIDl Data CMV. Furthermore, it will enable the researcher confirm dependent relationships within the data set. For instance, is achieving successful navigation reliant on navigation activities? Is effective visualization and adequate information retrieval dependent on successful navigation?

### 5.3.8 Tasks Categorization

Task	Category	Subject
1.	Object Identification	Can user spot groups of objects, categorize these objects and list the information they provide?

Task	Category	Subject
2a	Wayfinding	Can user navigate to a given Centre and retrieve information about it?
2b	Exploration	Can a user connect the publication to the centre(s) and keywords?
3a.	Wayfinding	Can user navigate through the given information space, locate centres and relevant information about them?
3b.	Exploration	Can user find relationships that exist between centres?
4a.	Wayfinding	Can user navigate to a given centre to retrieve specific information about it?
4b.	Exploration	Can user find relationships that exist within centres?
5a.	Object Identification	Can a user identify a given centre?  Can a user find the subject of interest of a centre?
5b.	Wayfinding	Can a user find relationships that exist during specified periods within a centre?

Task	Category	Subject
5c.	Exploration	
6.	Exploration	Can a user discover relationships that exist between a person and their environment?
7.	Object Identification	Can user gain an overview of categories of objects on the given information space.
8.	Wayfinding	Can user find detailed information about a centre?
9.	Object Identification	Can user identify various publication types on the information space?
10.	Exploratory task	Can user find detailed information relationships about relationships that exist within the institute?

Table 5. 1: Tasks categorization for empirical work 2

## 5.4 Quantitative Results

The navigation tasks carried out by participants on the two separate IIDI interfaces A and C, are reported in this section. The analysis focuses on the category of task participants were able to achieve, which is measured by the nature of navigation activities involved in performing the task while the navigation steps taken is measured by the number of clicks and mouseovers made. These parameters demonstrate the degree of support described as user support (Weinschenk and Baker, 2000) offered to people by interfaces as they navigate through the IIDI CMV information space. User support affords people the support required to learn and use a system.

Table 5.2 illustrates the tasks left uncompleted by each participant on both interfaces. It shows that participants carried out more tasks on interface C than on interface A.

### 5.4.1 Task Analysis of individual Participants

Participants (A – C Order)	Uncompleted Tasks (Interface A)	Uncompleted Tasks (Interface C)
A	2, 3, 4, 5, 9 10	2, 4, 5, 9
B	2, 4, 5, 9, 10	2, 4, 5, 8, 9, 10
C	2-5, 7, 8, 9, 10	2, 3, 4, 5, 9
D	2,4,5,8,10	2, 5
H	2, 3, 4, 5, 6, 8, 10	2, 5, 7, 8, 9
L	2, 3, 4, 8, 10	2,

(C –A Order)		
E	2, 4, 10	2
F	2, 3, 4, 10	2, 4
G	2, 3, 4, 5, 6, 7, 8, 10	2, 5
I	2, 4, 5, 8, 10	2, 4
J	2, 3, 4, 5, 7, 8, 10	2
K	2, 3, 4, 5, 8, 10	2, 4, 5, 6, 7, 8, 10

Table 5. 2: Cross section of uncompleted tasks on interfaces A and C.

#### 5.4.2 Task Category Analysis of Individual Participants

Task Category	Interface A (Number of participants who completed tasks)	Interface C (Number of participants who completed tasks)
1.Object identification	12	12
2a.Wayfinding 2b.Exploration	0	0
3a.Wayfinding 3b.Exploration	4	11
4a.Wayfinding 4b.Exploration	0	7
5a.Wayfinding 5b. Object Identification 5c.Exploration	3	6

Task Category	Interface A (Number of participants who completed tasks)	Interface C (Number of participants who completed tasks)
6.Exploration	8	12
7.Object Identification	9	12
8.Wayfinding	2	11
9.Object Identification	9	9
10.Exploration	0	10

Table 5. 3: Cross section of task category completed by individual participants

The results in table 5.3 indicate that Interface C supported more participants to perform the given tasks than interface A. More participants completed a higher number of tasks involving all three navigations activities on Interface C.

The highest number of tasks completed by participants on interface C was object identification and exploration. Next were the tasks requiring both wayfinding and exploration and the sole task that required all three navigation activities. The least task completed was the task that required only wayfinding. Comparatively, the highest number of tasks supported by both Interfaces was object identification and the least task supported involved wayfinding. However participants on Interface C showed an improvement in completing more tasks in all categories in less time (Table 5.3). This indicates that the provision of an interface key and summary data supported the tasks and improved the ability of participant's to achieve successful navigation. In general more information was retrieved on interface C than on A, this suggests that summary data facilitates visualization.

### 5.4.3 Analysis of Task Completion by Interface

The participants were placed randomly in two groups. Participants A, B, C, D, H and L carried out their tasks first on interface A and subsequently on interface C. They belong to the A – C order. Participants E, F, G, I, J and K performed the 10 ten tasks initially on interface C and then interface A. They belong to the C –A order. The different order was created to control the learning effect, one of the drawbacks of the within subjects design (Lazar, Feng and Hochheiser, 2010). The results in table 5.4 demonstrate that out of the 12 participants, 10 obtained higher scores from interface C, irrespective of the order they performed the task. This is an indication that participants retrieved more information from interface C. The scores participants obtained from performing the 10 tasks on both interfaces are summarised below according to the order they belong to:

A-C Order	Interface A		Interface C		C – A Order	Interface A		Interface C	
Participant	Task score	Time (min)	Task score	Time (min)	Participant	Task score	Time (min)	Task Score	Time (min)
A	4.0	18	<b>7.0</b>	13	E	6.0	14	<b>8.5</b>	14
B	<b>5.5</b>	13	5.0	14	F	5.5	18	<b>7.5</b>	14
C	2.0	15	<b>6.5</b>	12	G	2.0	18	<b>8.0</b>	15
D	<b>6.0</b>	14	5.0	14	I	6.0	14	<b>8.0</b>	12
H	3.0	17	<b>8.0</b>	13	J	3.0	13	<b>8.5</b>	8
L	4.5	12	<b>8.0</b>	13	K	4.0	18	<b>8.0</b>	9
Sum	<b>25</b>	<b>89</b>	<b>39.5</b>	<b>79</b>	Sum	<b>26.5</b>	<b>87</b>	<b>48.5</b>	<b>72</b>



A-C Order	Interface A		Interface C		C – A Order	Interface A		Interface C	
Participant	Task score	Time (min)	Task score	Time (min)	Participant	Task score	Time (min)	Task Score	Time (min)
Ave:	<b>4.16</b>	<b>14.8</b>	<b>6.58</b>	<b>13.1</b>	Ave:	<b>4.41</b>	<b>14.5</b>	<b>8.08</b>	<b>12</b>

Table 5. 4: Comparative analysis of task scores and task completion times of participants on interfaces A and C

#### 5.4.4 Analysis of Task completion Times

The records obtained from the log file show that on the average participants on interface C obtained faster task completion times than participants on interface A. Two of the participants were slow in completing the tasks on interface A and one of the two participants performed fewer tasks. However, table 5.4 illustrates that participants who worked in the C-A order on interface C completed their tasks faster than those of the A - C origin. This shows that irrespective of the order, participants completed tasks faster on interface C than on A. These results suggest that the changes obtained in task completion times appear to originate from the summary data provided by interface C.

#### A-C Order (Interface A)

Participant	Task score	Time (minutes)	Number of Event	Clicks	Mouseover	Navigation pattern
A	4.0	18	4	24	29	MN-MB- MM-MN- SN

B	5.5	13	9	161	251	MN-MM- MN-SN
C	2.0	15	7	74	120	MN-SN-- MM-SLM
D	6.0	14	5	79	152	MN-SN- MN-MN
H	3.0	17	9	58	134	MN-MB- MM-MN- SN
L	4.5	12	7	70	112	MN-MB- MM-MN- SN
<b>SUM</b>	<b>25</b>	<b>89</b>	<b>41</b>	<b>446</b>	<b>798</b>	

Table 5. 5: Summary of results of participants (A – C order) on Interface A

#### A-C Order Interface C

Participant	Task score	Time (minutes)	Number of Event	Clicks	Mouseover	Navigation pattern
A	7.0	13	6	52	87	MM-MM- SN-SLM
B	5.0	14	5	59	94	MM-MN- SN-MN

C	6.5	12	5	52	53	MM-MN- MM-SLM
D	5.0	14	7	108	121	MM-CMA- MM-SLM
H	8.0	13	6	51	173	MM-MN- MM-MN- SN
L	8.0	13	6	71	102	MM-MN- SN-MM- SLM
<b>SUM</b>	<b>39.5</b>	<b>79</b>	<b>29</b>	<b>393</b>	<b>677</b>	

Table 5. 6: Summary of results of participants on Interface C (A-C order)

C-A Order Interface A

Participant	Task score	Time (minutes)	Number of Event	Clicks	Mouseover	Navigation pattern
E	6.0	14	7	74	200	MM-SLM- SM-MN- SN
F	5.5	18	4	55	181	MN-MM- MB-MN- SN
G	2.0	18	7	221	365	MN-SN- MN-SN

Participant	Task score	Time (minutes)	Number of Event	Clicks	Mouseover	Navigation pattern
I	6.0	14	4	124	102	MN-MM- MN-SN
J	3.0	13	6	100	235	MN-MB- MN-SN
K	4.0	18	7	95	139	MB-SN- MN-SN- MM-SLM
<b>SUM</b>	<b>26.5</b>	<b>87</b>	<b>35</b>	<b>455</b>	<b>985</b>	

Table 5. 7: Summary of results of participants (C – A order) on Interface A

### C-A Order Interface C

Participant	Task score	Time (minutes)	Number of Event	Clicks	Mouseover	Navigation pattern
E	8.5	14	6	93	203	MN-MM- MN-SN- MM
F	7.5	14	6	58	110	MM-MN- MM-MN- SN
G	8.0	15	6	53	105	MN-MM- MN-SN

Participant	Task score	Time (minutes)	Number of Event	Clicks	Mouseover	Navigation pattern
I	8.0	12	6	52	121	MN-SN- SLM-SN
J	8.5	8	5	65	129	MM-MN- MM-SLM- MN-SN
K	8.0	9	6	68	136	MM-MN- SLM-MM- SN
<b>SUM</b>	<b>48.5</b>	<b>72</b>	<b>23</b>	<b>389</b>	<b>804</b>	

Table 5. 8: Summary results of participants (C – A order) on Interface C

Tables 5.5 and 5.6 demonstrate that participants belonging to the A-C order performed less clicks and mouseovers (navigation steps) on interface C than on A.

A similar result was obtained from the C-A order who also obtained less mouseovers and clicks on interface C than on A (tables 5.7 & 5.8). Participants F and G obtained the fewest mouseovers and clicks (table 5.8) and the participants with the fastest task completion times were E & J on interface C (table 5.8). In general, participants obtained faster task completion times, completed more tasks and obtained less clicks and mouseovers on interface C than on A.

A summary of navigation steps of individual participants is given in table 5.9 below.

Participant	Interface A		Interface C	
	Clicks	Mouseovers	Clicks	Mouseovers
A	24	29	52	87
B	161	251	59	94
C	74	120	52	53
D	79	152	108	121
E	74	200	93	293
F	55	181	38	210
G	221	365	53	205
H	58	134	51	173
I	124	102	52	121
J	100	235	65	189
K	95	139	48	196
L	70	112	71	102
Total	<b>1134</b>	<b>2020</b>	742	1844

Table 5. 9: Comparative analysis of navigation steps on interfaces A and C

Table 5.9 illustrates that the highest number of mouseovers and clicks which point to navigation steps was made on interface A. The reduced navigation steps on interface C could be as a result of the support offered by Interface C to participants to move around in a more focused manner in order to find and retrieve adequate information. The results on the task sheet confirm that more adequate information was retrieved from interface C with faster task completion times. Fewer clicks were also made on interface C indicating that fewer navigation steps were required. The empirical work demonstrates that the changes obtained in category of task completed, task

completion times and navigation steps could have originated from the summary data and interface key provided by interface C.

## **5.5 Qualitative Results**

A semi – structured interview was conducted after the experiment and although similar questions were asked during the interview sessions in empirical study one and two, the responses were more closely categorised to a five point scale. The interviewees were asked two specific open ended questions, the aim was to allow features that provided navigation support to be discussed (question 2) and about tasks depending on how challenging or easy they were to achieve (question 5). For the other four queries, interviewees were gently guided to locate their views on a 5- point likert scale in order to obtain less biased measurements. All twelve participants willingly took part in the interview and made similar observations about the interface key when compared to participants in the previous experiment. They agreed that the interface key and the upper centre view offered a lot of support to their ability to navigate to data in order to retrieve information. The information obtained was recorded and transcribed for analysis.

### **5.5.1 The Interview Questions**

1. From your experience using the interface, do you believe that the interface supported you to find your way around the different views efficiently enough for you to visualise data and retrieve adequate information?
2. Can you please tell me about one feature that offered you any support on the interface?

3. Were you able to interpret the information represented by the objects on the Interface?
4. The tasks you just completed evaluate three basic activities: the ability to find your way around, explore the interface and identify objects. Would you say you were able to achieve at least two of these activities?
5. Can you give me an example of any task you recall? Possibly as a result of how easy or challenging it was to complete?
6. Would it be fair to say you had a satisfying experience navigating the ANG?

In the second empirical study, the same participants worked on both interface A, and C. Their response to the open-ended questions gave detailed insight to the features that supported them to navigate either interface. On interface A, most interviewees agreed that the upper right view was very useful. They said it was the easiest view to understand as a result of the key provided to explain the coding system. They pointed out that they were able to navigate between data easily in order to visualize and retrieve information; however they noted that the other views were quite challenging.

On interface C, most interviewees said the upper centre view and the upper right view were very valuable as it gave them access to more information. They explained that the interface key served as a guide, steering them in the right direction.

### **5.5.2 Interface Analysis**

The five likert scale questions used in the interview were summed up into direct statements for easy reference. They are outlined below;

Totally Agree (TA); Mildly Agree (MA); Neutral (N);



Mildly Disagree (MD); Totally Disagree (TD)

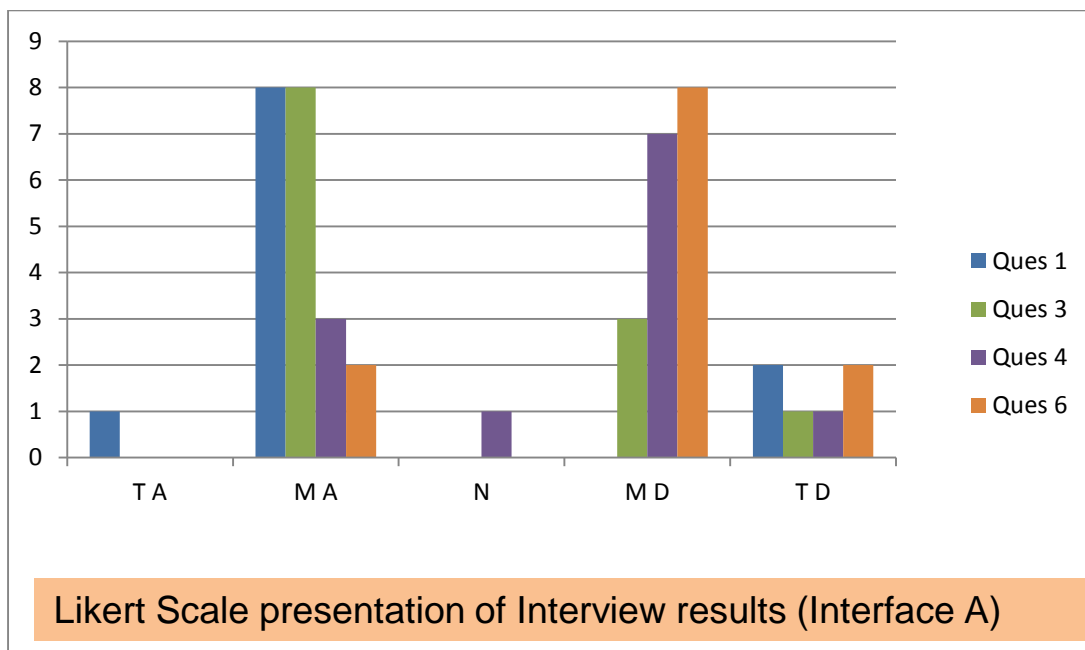


Figure 5. 4: Interview results for 12 respondents for Interface A

On interface C, no 'neutral', 'mildly disagree' or 'totally disagree' responses were recorded while interface A, recorded some 'mildly disagree' and 'totally disagree' responses. The results obtained from the interview indicate that interviewees were able to navigate interface C more easily than A. Interviewees agreed that the interface key played a major role in their ability to interpret the objects on the interface.

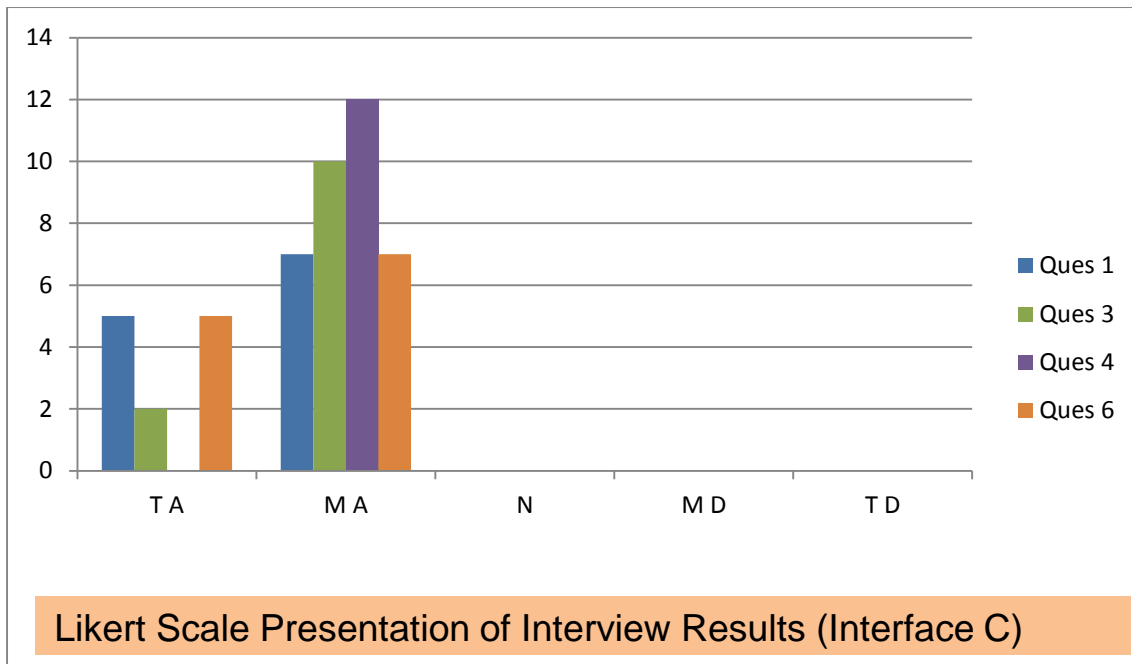


Figure 5. 5: Interview results for 12 respondents, Interface C

## 5.6 Statistical significance

Statistical significance is a useful technique to determine that the outcome of a study is not by chance but is attributed to a specific reason. It is analysed as the probability that a change observed in an empirical study occurs due to chance. It is expressed as a P-value. A smaller P-value indicates that the results obtained are likely to be true. It helps researchers decide when to reject a null hypothesis. When the statistical significance calculated is a P-value less than 0.05 ( $p < 0.05$ ) it is generally acceptable by researchers (Earl, 2013) and the null hypothesis is rejected.

Statistical analysis is a useful technique to help researchers find important trends and variations in a data set. It also highlights relationships between variables. In order to choose the correct significance test, the type of data collected and the design of the empirical study is taken into consideration. Parametric tests are employed when the data collected are scaled in intervals or normally distributed. When the assumptions

mentioned below are not met, non-parametric tests are employed. A study adopting a within groups design that does not meet the assumptions.

The Wilcoxin signed ranks test is employed to validate the results of the data in this research. It is appropriate for this experiment because the data collected for the experiment was obtained by influencing the independent variable so as to observe the outcome on the dependent variables. In addition, it allows the researcher compare data collected from the same user group such as data collected from a within group design where the same participants performed the experiment under both conditions (Lazar, Feng & Hochheiser, 2010).

Experimental Hypothesis:

The provision of data along with its summary data will improve people's ability to navigate coordinate multiple view information spaces. In addition it will enhance the effectiveness of data visualization, reduce visualization times and enhance adequate data retrieval.

Null Hypothesis:

The provision of data along with its summary data creates no significant change on people's ability to navigate an information space, visualize effectively or retrieve adequate information.

The Wilcoxon signed rank test:

The Wilcoxon signed rank test is a non-parametric test employed to validate data that is not normally distributed and is collected from a repeated measures design. The data is usually paired and collected before and after a specific treatment from the same user group. The test determines whether the difference between the conditions provided in interface A and B are significant. It will also show if a noteworthy variation exists in the performance of the two groups of participants.

Wilcoxon signed rank test analyses the difference obtained from scores and takes into account the magnitude of observed differences. The study assessed the effectiveness of data provided along with its summary data on IID1 CMV interface C, designed to improve the ability of people to navigate successfully through a CMV information space in order to visualize data effectively and retrieve adequate information.

The study is interested in participants' performance on interface A and C. Participants will complete tasks faster on interface C than on A if the data provided along with its summary data has an effect on their ability to achieve successful navigation, visualize effectively and retrieve adequate information. A total of 12 people enrolled in the study and the length of time each participant takes to complete the tasks within a 10 minute period is measured on interface A (no summary data) and on interface C (summary data present). In order to determine the statistical significance, the mean difference is computed, the test statistic is determined and the critical value.

Participant	Interface C (Summary data)	Interface A (No summary data)	Difference	Ordered Difference	Ranks	Signed Ranks
A	7.0	4.0	3	-0.5	1	-1
B	5.0	5.5	-0.5	-1.0	2	-2
C	6.5	2.0	4.5	2.0	3.5	3.5
D	5.0	6.0	-1.0	2.0	3.5	3.5
E	8.5	6.0	2.5	2.5	5	5
F	7.5	5.5	2.0	3.0	6	6
G	8.0	2.0	6.0	3.5	7	7
H	8.0	3.0	5.0	4.0	8	8
I	8.0	6.0	2.0	4.5	9	9
J	8.5	3.0	5.5	5	10	10
K	8.0	4.0	4.0	5.5	11	11
L	8.0	4.5	3.5	6	12	12

Table 5. 10: Table illustrating computation for Wilcoxon signed rank test

Table 5.10 shows that the probability ratings on the average end up higher on interface C than on A. In order to determine if the observed difference is statistically significant, the Wilcoxon test statistic ( $w$ ) is obtained.

Number of positive signs ( $w^+$ ) = 10)

Number of negative signs ( $w^-$ ) = 2)

Observed test statistic ( $w$ ) = 3

$n = 12$

Critical value from critical value table = 13

$\alpha = 0.05$

H0 = null hypothesis

H1 = research hypothesis

Decision rule = reject H0 if  $w \leq 13$

In this experiment obtained value of  $w = 3 \leq 13$

	$\alpha = (0.05)$	$\alpha = (0.01)$
N=12	13	7

Table 5. 11: Critical table values for Wilcoxon signed rank test

Ranks

	N	Mean of Rank	Sum of Ranks
Interface C – A Negative ranks	2	7.5	75
Positive ranks	10	-1.5	-3
Ties	0		
Total	12		

Test Statistics

	Interface C - A
	W = 3
Assumption sig. (2 – tailed test)	$\alpha = 0.05$

A Wilcoxon signed ranks tests indicated that data provided along with its summary data improved navigation, visualization and information retrieval. Therefore we reject H0. The result is statistically significant at  $\alpha = 0.05$ . This confirms that the median difference is positive, which means that providing data along with its summary data improves navigation, brings about a more effective visualization and adequate information retrieval.

## **5.7 Conclusion:**

The results obtained from the second empirical study are encouraging and indicate that providing data along with its summary data improves the ability of people to achieve successful navigation, visualize data effectively and retrieve adequate information.

The Wilcoxon signed rank test shows that the results obtained are statistically significant, meaning that summary data facilitates successful navigation and in general improves visualization and information retrieval. It is reasonable to say that CMVs could be made more navigable if people are able to manage the different aspects of an interactive information space such as navigation, visualization and information retrieval. This will facilitate people's ability to understand an information space and then utilize the knowledge acquired to retrieve information (Spence, 2002). In addition it demonstrates providing data along with its summary data is a useful method to make information available to a broad range of people. The next chapter will discuss the set of design principles drawn from the previous experiments and a third empirical study that evaluates these designs.

## **Chapter 6 Empirical Study Three**

### ***6.1 Introduction***

The results obtained from the second empirical study in chapter five illustrate that summary data which is explained as data about centres, publication and grants and metadata about people was useful to people navigating the IIDl Data CMV. A review of literature has drawn attention to the subject that even though metadata has been used in various ways to improve interfaces, empirical evaluations carried out on these systems are few (McCormac, Parsons and Butavicius, 2007).

Many CMVs have a task based origin as designers of these CMVs start by establishing a good understanding of the tasks users will want to perform. Then they build a system that is expected to allow the user accomplish these tasks (Baldonado, Woodruff and Kutchinsky, 2000). However, members of the workshop held during a working conference on advanced visual interfaces (AVI) agreed that the design process for CMVs could be advanced by evaluation (Baldonado, Woodruff and Kutchinsky, 2000).

From the empirical studies carried out in chapters 4 and 5, the research has drawn up a set of guidelines for designers of CMVs, specifically for management information systems in order to move the CMV from being fundamentally task oriented to the flexibility provided by the concept of navigation of information space. The first three guidelines were developed and implemented on interface B before the first empirical study discussed in chapter four while the fourth guideline was developed during the study. At the end of the second empirical study, the fifth guideline was developed. The flexible nature of the CMV allows a broad range of tasks to be performed indicating



that the characteristic of the redesigned CMV affords users the opportunity to appropriate the information space.

In this chapter, this research will put forward and evaluate the set of guidelines in order to validate their contribution to successful navigation.

## **6.2 User Interface**

User interfaces are commonly evaluated by usability tests and empirical studies. Usability studies are usually employed to measure a tool while empirical studies evaluate interfaces. The IIDI visualization interface C will be evaluated by an observational user study using empirical methods. Other methods for user interface evaluation are automatic, formal and informal ways. In the recent state of development of interfaces automatic methods are challenging, while applying formal methods are complex (Nielsen and Mark, 1994) and handle large user interfaces poorly. Automatic methods for user interface evaluation involve usability measures computed by running a user interface requirement using an appropriate program. However, Nielsen and Mark, (1994), point out that automatic methods do not work under the present state of art.

Formal methods involve using mock-ups or prototypes and procedures to calculate usability measures. A paper mock-up can be described as a prototype if it provides the minimum functionality of a system and enables testing of a design while a prototype is an initial illustration or model of a design built to test a concept. Mock-ups are used by designers for purposes of acquiring feedback from users and prototypes are created in order to test and try out a new design and improve its precision by system analysts and users. Prototyping makes it possible to provide specifications for

a real, working system rather than a theoretical one. It is sometimes the step between formalization and evaluation of an idea. Usability studies can be carried out using paper mock ups or prototyping, as these methods are inexpensive, rapid and very productive. Still, these formal methods are challenging to apply and are not effective enough to handle large user interfaces.

### **6.2.1 Design Guidelines**

Design guidelines for interface are groups of recommendations taken into account by designers and developers when creating user interfaces. They include general design principles derived through research such as design methodology, expression of an important design viewpoint and assumptions about human behaviour (Jonathan, 1989). A number of design guidelines such as the information seeking mantra (Shneiderman, 1987, 1994), the ten heuristics for user interface design (Nielson, 1993) for general user interfaces and guidelines for using multiple views in information visualization (Baldonado, Kutchinsky and Woodruff, (2000) have been proposed and after appropriate upgrading, extending and interpreting have been applied in many interactive systems. Design guidelines created for general user interfaces are of importance (Nielson, 1994 and Shneiderman, 1987). However the prospect that specific guidelines tailored towards specific user interfaces may improve results. Baldonado, Woodruff and Kutchinsky (2000) claim there is little specific guidance presently available to designers of CMVs. They argue that designers may find it difficult to choose the most relevant set of guidelines to enable them create multiple views out of the pool of general design guidelines in existence.

The development of specific principles for designing multiple views will advance designing CMVs that support successful navigation, effective visualization and adequate information retrieval. Baldonado and the members of his group have developed guidelines and customized versions of general guidelines that address problems affecting multiple view systems (Baldonado, Woodruff and Kutchinsky, 2000). These guidelines were presented in two sets. The first set of guidelines address how to help designers choose when to use multiple views while the second sets of guidelines discuss how to use these multiple views.

The guidelines presented in this study are results obtained from separate sets of user studies. They were put together from the concept of navigation of information space (Benyon, 1998, 2006, 2007), by observing navigation behaviour of participants in the empirical studies in chapter 4 and 5 and from important viewpoints expressed by participants. These guidelines are to be used in addition to existing guidelines in order to produce navigable CMVs. Multiple view systems offer the user a range of benefits such as improving user performance and enhancing discovery of novel relationships (North and Shneiderman, 1997). Although navigation of CMVs has not been studied sufficiently, it plays a significant role in information space interaction, visualization and data retrieval. Benyon and Hook (1997) predict that creating large information spaces and paying little attention to navigational aids, will bring about significant differences in performance between individuals. These differences will be seen in information retrieval times, user satisfaction in using the system, their ability to learn the system and rely on the results they obtain. In general the usability of the system will be largely different for different users.

### **6.2.2 Proposed Guidelines for Designing CMVs.**

The guidelines proposed by this research are concerned with making information spaces effectively navigable. The set of guidelines specifically address CMVs created to visualise MIS. Navigation involves the successful movement of people within the information space from a current position to a destination, even when the location of the destination is not known.

Research shows that most user studies simply report whether their technique is useful or not but do not give reasons for the usefulness (Kosara, Healey, Interrante, Laidlaw, and Ware, 2003). In addition to illustrating the usefulness of this research, the study will explain why the concept of navigation information technique is effective. The empirical studies discussed in chapter 4 and 5 were designed to focus on the concept of navigation of information space in order to assess different aspects (Shneiderman, 1994) of the interface. The design guidelines the research employed to make Interface C more effective than Interface A are presented in this study as short statements and form the set of guidelines proposed by this study.

#### **Decide main focus of data set**

When designing a CMV, the designer should decide the story the CMV wants to tell. The main focus of the data set and the key relationships based around that focus should be clearly defined. Defining the centre of attention of the data set helps the designer decide the attributes that require provision of summary data. This principle can be applied on interactive visualization information spaces such as CMVs where data is presented from different perspectives. An example of its application can be seen on the IIDl data interface B and C.

### **Provide access to attributes of the main focus of data set.**

In designing a CMV, access to the main focus of the data set positively impacts on navigation, visualization and information retrieval. This access can be provided through summary data which makes the data about attributes that have a many-to-one relationship with the main focus of the data set available. People need to access objects for a set of metadata values in order to analyse and retrieve information from these objects. The designer may need to choose the appropriate levels of data required in order to provide access to the metadata of the focal point of the data set. Access to metadata provides more details, answers more questions and facilitates movement as data is made available on various levels. An application of this method can be seen on interface C of the IIDI data, participants were able to move from a view of data to a view of metadata and back again, depending on the focus of their queries. They gain a general idea of the interface and filter the data to access other views to obtain more details (Shneiderman, 1998) . This method makes it possible for the needs of a broad range of users to be met. The story the CMV wants to tell will determine the choice of which metadata to expose. The type of visualization to use in a CMV is a crucial aspect but choosing the appropriate levels of data is also an important aspect.

### **Provide Summary Data**

Summary data is described by this research as data provided alongside its metadata. It is data about content. In the context of navigation of information spaces we define it as data about the attributes that have a many to one relationship (M: 1) with the main focus of any database. For instance, in the case of an information management

system such as the IIDl data, the main focus is people. Centres is a key relationship based around this main focus and has a many to one relationship with people so in this instance, data about centres, publications and grants is summary data. Data about these attributes is metadata about people. The concept of navigation of information space directly impacts on navigation by applying some knowledge of the way physical spaces are designed to the CMV information space. It incorporates navigation activities such as object identification, wayfinding and exploration and provides summary data. This brings about a flexible information space and makes it easier for people to navigate through. Currently, the CMV designer takes a task based approach (Baldonado, Woodruff and Kutchinsky, 2000) creating a CMV that answers specific questions. Nonetheless, the summary data takes a flexible approach incorporating task and navigation activities in order to answer general questions. The research proposes that if CMVs are designed to incorporate navigation activities and data is provided alongside its metadata (summary data) it will bridge the gap created by a task based information space and produce a balanced information space with the potential to support users answer specific and general questions and in some cases appropriate the space. Furthermore, the possibility of achieving successful navigation, visualization and information retrieval is heightened. The results from the empirical work indicate that the provision of summary data on interface B and C positively impacted on the task completion times of users.

### **Provide an Interface Key**

A key can be explained as an object that literally opens a door and affords people the chance to execute an action. An interface key can be described as metadata. It is data about the visual design aspects of the CMV which interprets the objects on the

interface to facilitate understanding. In empirical study one, participants pointed to it as a feature that offered them an opportunity to interact with the data and virtually opened up their understanding of the Interface. It could facilitate understanding of maps, landmarks, labels, signs and objects scattered around the information space. It is not a new concept in itself and is applicable in various information spaces however it has not been integrated as a guideline. Research on how to utilize knowledge of maps, landmarks and signs to design information spaces is in progress. A key can provide personalized support to users of information spaces by making available to them interpretations that will enable them understand the information presented in order to make meaningful conclusions. The initial empirical study showed that people spent a lot of time exploring the interface, this is indicated by the higher number of mouseovers obtained (chapter 5, Table 5.9). Participants pointed out that the easiest view to navigate on interface A of the IIDl visualization data was the upper right view which supplied an interface key. This research suggests that an interface key should be integrated as the first level metadata of an object in order to describe the coding system in any visualization.

### **Design for Appropriation**

Designing for appropriation is described by Alan Dix (2007) as designing to allow the unexpected. Appropriation can be explained as something intended for a purpose made suitable for the present circumstance, for instance emails intended for distant communication with friends is now used by people to email themselves web links while browsing thus appropriating emails. Appropriation incorporates change and flexibility. It allows users carry out tasks the designer may not have envisaged. A CMV created for a specific group of people to perform specific tasks may become outdated in a few

months if new people with new ideas join the group. On the other hand users may accept the technology, coming up with several strategies and become accustomed to it and use it for purposes other than what the designer intended. Users understand technology and use it better if they can use it in their own way. In the words of Shneiderman (2002), 'the old computing studies what computers can do, the new computing is about what people can do'. A CMV designed to be flexible will be able to answer other sets of questions in addition to the specific set of tasks it was designed to suit. A CMV designed to incorporate summary data will allow for the flexibility required by users to meet practical needs and appropriate CMV information spaces where the need arises.

### ***6.3 Evaluating the IIDl Data Interface C***

In HCI researchers have deliberated on the types of evaluation techniques most suitable for different studies. To appraise the general usefulness of a tool, a broad range of evaluation methods such as observational user studies are used. It has been suggested that in evaluating exploratory interfaces, emphasis should be put on information seekers and the tasks they need to perform so as to obtain richer feedback on the competence of a method (Kraaij and Post, 2006).

#### **6.3.1 The value of summary data and Interface key**

The purpose of this empirical study is to measure the added value of summary data and interface key in terms of navigation. It will test the effectiveness, efficiency, and satisfaction of the interface designed with the proposed guidelines in relation to its navigability. For example does the improved navigability impact on the interface usefulness in achieving tasks (information retrieval) and task completion times? If



there is a benefit, what is the importance? It will evaluate whether users can access different levels of data and appropriate the interface.

### **6.3.2 Method**

The approach employed to evaluate the IIDi interface C is one of the most frequently used empirical methods, known as user testing. User testing is a reliable way to evaluate an interface with real users. In addition aspects of the think-aloud protocol of the instant data analysis (Kjeldskov, Skov and Stage, 2004) may be used to identify the strengths and weaknesses of the design. This study will focus on the navigability of the interface. The participants will navigate interface C in order to retrieve the required information, and the data obtained will be used to assess the set of guidelines drawn from the initial studies carried out by this research. The user study will demonstrate whether each principle contributes to successful navigation. Qualitative data will be obtained from responses given by participants during the semi-structured interview for purposes of analysis.

### 6.3.3 IIDI Visualization Interface C

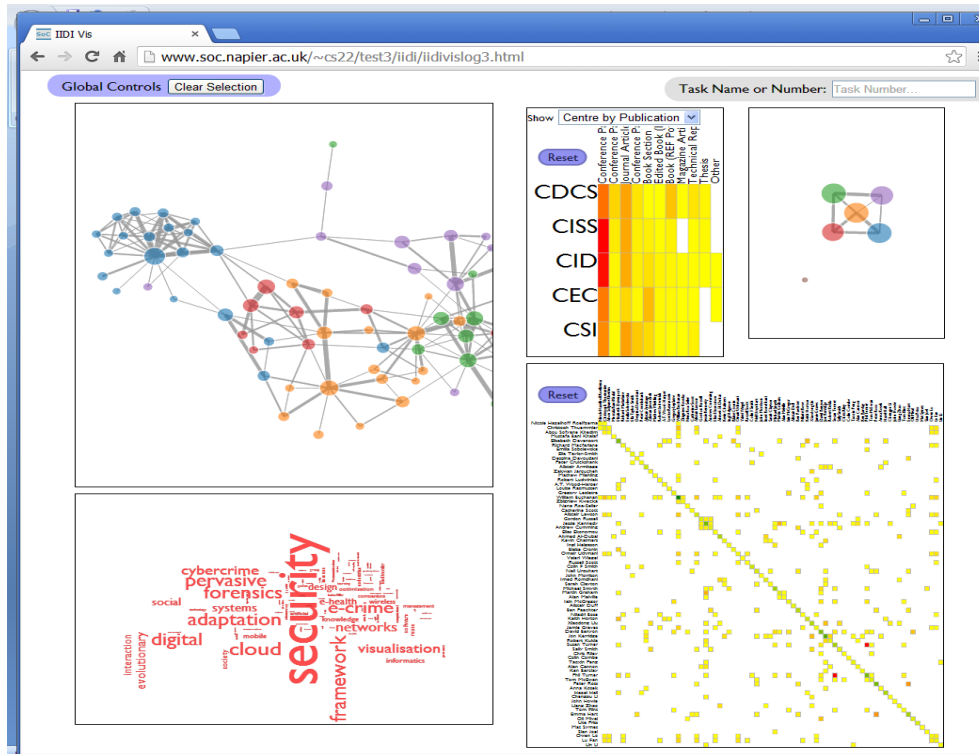


Figure 6. 1: IIDI Data CMV Interface C:

Source: [www.soc.napier.ac.uk/~cs22/test3/iidi/iidivislog3.html](http://www.soc.napier.ac.uk/~cs22/test3/iidi/iidivislog3.html)

### 6.3.4 Participants

Six users (three males and three females) between the ages of 28 and 50 took part in testing the IIDI Data interface C. Two database designers, one interaction designer, a research student and two managerial staff were recruited for the evaluation by word of mouth. They were all experienced users of computers and information technologies. The researcher was assisted by a postgraduate student who had participated in six user studies and was assigned the role of test monitor. The researcher acted as the data logger during the sessions, observing and taking down notes.

### 6.3.5 Experiment Setup

The six user studies were carried out in the Merchiston library of the Edinburgh Napier University and participants were monitored closely to observe how they interact with the interface. A regular desktop computer, mouse and key board were used for the evaluation. The tasks selected for this study were designed to validate the contribution the guidelines proposed by the research make to developing navigable information spaces. The information retrieved by executing the tasks indicate whether user's have sufficient access to information about objects represented in the interface. It is designed to confirm whether the metadata about content (summary data) and metadata about visual aspects (interface key) improves participant's ability to navigate through the CMV. A case in point is task 6, where participants were asked to develop individual tasks, execute the task and state if they were able to achieve it. This task demonstrates that interface C affords user's the opportunity to appropriate the interface. They can carry out tasks that the designer did not think about, which confirm that the interface has moved from being task centred to being more flexible.

1. Identify and list two objects on the IIDl visualization and write down the information they provide.
2. Locate two people in the centre for interaction design (CID) that have made joint publications with people from other centres.
3. Write down the name of the person who has received the highest number of grants in each centre.
4. Find the member with the highest number of publications? How many grants has the person received?

5. From the overview, list the five centres in the institute. Which centre has made the highest number of publications? How many grants has the centre been awarded?
6. How many conference papers have been published by CDCS?
7. a). Think of a task you would like to carry out and write it down.  
b). Perform the task.  
c). Please state if you were able to perform the task. d). If your answer is No. Please state why?

### **6.3.6 Procedure**

The six user studies were carried out in one day. Three users participated in the morning session and the other three in the afternoon session. Each participant was assigned thirty minutes.

#### **Briefing**

The participants were handed consent forms to read and fill and understood that they could cease to participate at any time. The test monitor introduced the participants to the procedure. Following this they were handed the task sheet and interface key.

#### **Task solving**

The interface test was structured by seven representative tasks, at least one task to specifically evaluate each of the guidelines proposed. The purpose of each task was to replicate practical information seeking tasks. The participants worked on answering

the questions until they were told to stop by the test monitor when it was time or when they had completed the task. They were encouraged to think aloud during the task solving session explaining their interaction with the interface and expressing their understanding of the design or any difficulties encountered. The test monitor asked the participants questions for clarification, however no help was offered to the participants to prevent ruining the results.

### Interview

Following each study, the participants were asked about the experience they had working on the interface. They were told to rate the interface briefly and offer any comments they had.

### Data analysis

The researcher took notes and recorded interesting observations such as the actions users took, where they succeeded and the challenges they faced while they carried out tasks. Think –aloud thoughts were also recorded on post- it notes with each participants name indicated for purposes of differentiation. Observations and comments from each participant were recorded. The notes were categorized into themes and significant information collected.

### The Variables

The independent variable, Interface C

The dependant variables measured in this study are the parameters that are used to assess user interfaces. They can be used as tools to evaluate whether the proposed

principles implemented on the IIDI data interface C improve navigation, visualization and information retrieval.

The parameters are outlined below:

Success rate (the number of participants who complete the task successfully)

Task completion times

User satisfaction

Utility

## **6.4 Results**

We present results that relate to the use of the set of principles proposed by this research to improve navigation. During the data analysis the significance of summary data and interface key in relation to navigation was examined. The analysis looked at the usefulness, effectiveness, efficiency and user satisfaction of interface C. The parameters measured include:

### **6.4.1 Success rate**

The success rate measures how effective a design is. It is defined as the percentage of tasks that users complete correctly (Nielsen, 2001). Partial score can be granted to users who complete much of the task depending on the extent of error. Partial score estimates are useful as they provide a more practical idea of the quality of the design than an unconditional approach to success and failure. According to Nielsen (2001) 50% credit works well when there is no basis compelling researchers to give alternate type of errors such as high or low scores.

Success rates clearly indicate how much users can accomplish on an information space. Nielsen (2001) describes user success as an unavoidable factor in measuring how useful and effective a design is.

	<b>Task 1</b>	<b>Task 2</b>	<b>Task 3</b>	<b>Task 4</b>	<b>Task 5</b>	<b>Task 6</b>	<b>Task 7</b>
<b>User T</b>	S	S	S	S	S	S	S
<b>User U</b>	S	S	S	S	S	F	F
<b>User V</b>	S	S	S	S	S	S	S
<b>User W</b>	S	S	P	P	S	F	S
<b>User X</b>	S	S	P	S	S	S	S
<b>User Y</b>	S	S	S	S	S	P	F

Table 6. 1: Success rates of users on interface C

Key: S = Success, P = Partial success, F = Failure

The results show that 34 out of the 42 attempts to perform the tasks were successful, 4 were partially successful while 4 failed. Success rates provide a method to measure progress made by a new design and how effective or useful it is.

$$\text{Success rate} = (34 + (4 * 0.5)) / 42 = 86\%.$$

#### 6.4.2 Task completion times

The time users take to complete a task can be used to measure how efficient an information space is. It usually has a wider distribution as some individuals take longer or shorter time to complete tasks. The distribution of tasks times will be used to present

the results obtained from this experiment. The core measure used for this parameter is the average time taken to complete each task.

<b>Tasks</b>	<b>T</b>	<b>U</b>	<b>V</b>	<b>W</b>	<b>X</b>	<b>Y</b>	<b>Task Average</b>
<b>1</b>	60sec.	66sec.	138sec.	62sec.	37sec.	25sec.	65sec
<b>2</b>	67sec.	71sec.	65sec.	143sec.	67sec.	61sec.	79sec.
<b>3</b>	182sec.	139sec.	282sec.	151sec.	125sec.	186sec.	176sec.
<b>4</b>	50sec.	91sec.	39sec.	123sec.	131sec.	74sec.	85sec.
<b>5</b>	60sec.	123sec.	85sec.	39sec.	66sec.	60sec.	73sec.
<b>6</b>	62sec.	120sec.	60sec.	60sec.	77sec.	72sec.	76sec.
<b>7</b>	184sec.	123sec.	61sec.	61sec.	204sec.	123sec.	126sec.
<b>Total time</b>	665sec.	733sec.	730sec	637sec.	707sec.	602sec.	

Table 6. 2: Task completion times (seconds) of users on Interface C

The average time taken to complete the entire task = 679seconds (11m 32s).

### 6.4.3 User Satisfaction

This affords users the opportunity to rate the interface. The interface was rated from 'difficult' to 'easy'. The participants were probed further and interesting responses recorded. They were asked specifically to rate the interface in terms of the given task and their ability to execute them. One participant found the interface easy to use, two found it fairly difficult and three found it fairly easy to use.

Comments made by participants during the user study were interpreted and representative comments outlined below:



I found the *UCV* and the *URV* quite useful. They were very helpful in tackling task 5, 6 and 7.

The '*Interface key*' was valuable in understanding and interpreting the interface in general.

The *matrix* was the most challenging view to understand.

The *URV* supported the execution of tasks related to the *ULV*. It reduced the number of clicks performed on the nodes.

The labels provided in the *UCV* provide links to detailed information about centres, grants and publications.

The *UCV* helped to find the relationship between people centres, publications and grants.

The *UCV* provided detailed information about the central focus of the data

#### **6.4.4 Utility**

Utility refers to the functionality of the design, like how permeable (Benyon, 2005) the environment is and the extent to which it provides features required to support users navigate the interface. In this evaluation appropriation was seen as an aspect of utility.

In order to measure this parameter, the participants were asked to name features on the interface C that they found quite useful from very useful through least useful. The features on the *UCV* such as the labels that link to centres, publications and grants windows and the nodes on the *URV* were identified as very useful. Four participants said the features on the *UCV* and *URV* were easy and clear to use. The participants agreed that the interface key helped them understand the interface and they were able to appropriate the interface for their own purpose. This is expressed by the ability of

four out of six participants to tackle question 7, a specific question designed to evaluate whether participants can evaluate the interface.

## **6.5 Discussion**

This chapter demonstrates how the set of guidelines drawn from previous empirical studies was evaluated through a user based testing on the IIDI Data Interface C. The interest of the research was centred on evaluating how useful the IIDI interface C, implemented with the proposed set of guidelines was to the users.

The evaluation involved 6 participants who performed various tasks. Success rates, task completion times, the functionality and subjective rating of the interface were obtained. The results obtained from the third empirical work demonstrate that when the additional guidelines proposed by this research were implemented on the IIDI data interface C, users were able to navigate the interface successfully as illustrated by the success rate (86%). The subjective rating by users was positive and the UCV and URV that provided additional summary data were rated as very useful view, easy and clear to use. In addition, the users agreed that the interface key provided invaluable support.

The concept of navigation of information space is an effective approach. One reason is because it enables the developer of the IIDI CMV interface C design the interface in a way similar to real world information spaces. Activities that people undertake in the real world when executing a task were incorporated in Interface C by providing data along with its summary data. Information was made available in different levels to cater for a wide range of people. In addition, an interface key that can be compared to

information signs in real life was provided to enable people interpret and understand the information space. This approach facilitated peoples' ability to achieve successful navigation, effective visualization and adequate information retrieval. Specifically, participants said the Interface key was helpful in gaining understanding of the interface and interacting with the data set. Evaluators suggested that the lower right view should be improved so as to facilitate people's ability to find meaningful relationships.

The evaluation undertaken as part of this thesis project is an initial attempt to study how the concept of navigation of information space may support navigating CMVs. The insights suggest that the provision of summary data and an interface key is a feasible approach to improve navigation, visualization and information retrieval on CMVs. Nevertheless, further studies on the concept of navigation of information space are required to advance the potential role of navigation in visualization and information retrieval.

## **6.6 Conclusion**

The results have shown that when the set of guidelines proposed by this research were implemented on the IIDI Data CMV for an MIS system, it offered significant support to people to navigate in order to visualise, discover and retrieve adequate information. It can be inferred from these results that these additional guidelines when applied in addition to the existing guidelines will advance CMVs and make them more navigable. Furthermore, CMVs will have the added advantage of people being able to appropriate the interface. In the next chapter a brief case study bordering on summary data as metadata will be discussed.

## **Chapter 7 Visualizing the SIMD Data: A Case Study**

### ***7.1 Introduction***

The previous chapter evaluated an interface that provides summary data (data and metadata) can ease some navigation challenges encountered by people on information spaces. The purpose of this chapter is to generalize the idea further to a different domain. It presents a brief case study of the Scottish Index of Multiple Deprivation (SIMD) data in order to discuss the difference availability of summary data can bring about on navigation, visualization and information retrieval.

### ***7.2 The Scottish Index of Multiple Deprivation (SIMD) Data***

The Scottish Index of Multiple Deprivation (SIMD) is a technique used by local authorities, the Scottish Government, the National Health Scheme and other governmental bodies in Scotland to identify deprived places in Scotland. It embraces several aspects of deprivation and puts them together into a single index. It splits Scotland into 6,505 small consistent areas containing about 350 households called datazones. The index provides a comparative ranking for each datazone from 1, most deprived to 6,505, least deprived. The concept of the SIMD is to identify datazones with concentrations of multiple deprivations and then direct policies and resources to the places that have the greatest need.

Deprivation is the challenge that is created due to a lack of resources that should cover health, education, employment, housing, access to services and finance. The SIMD collects data related to several aspects of life to obtain a clearer picture of the deprivation across Scotland. The aspects of deprivation it incorporates makes

available an extensive range of information and statistics on health, unemployment, education, crime, poverty, housing, population, and social or community issues in Scotland and works by merging 38 indicators across these 7 domains. The general index is a sum of the seven domain scores.

The SIMD allows people using it to focus on the datazones below a certain rank, for example, the 5%, 10%, 15% or 20% most deprived datazones in Scotland. This can be achieved by the SIMD interactive mapping which permits people to visualise interactive maps of the SIMD outcome for different areas in Scotland. This technique enables people search by postcode, datazone or the name of the area and access results for the 7 domains that make up the SIMD or the overall SIMD. Additionally, people can make comparisons between datazones, SIMDs, results for different areas and view the population of each area. People can access data on crime, income, employment, education, housing, access and health for the different areas of Scotland and compare this data in order to find the percentage deprivation of the different areas.

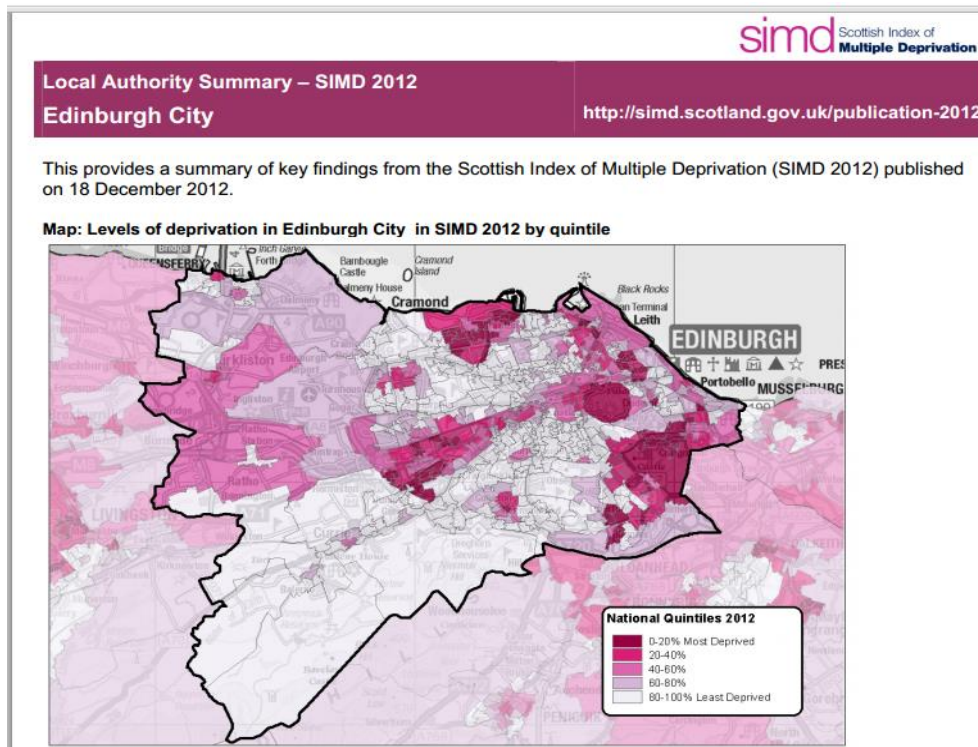


Figure 7. 1: Diagram illustrating levels of deprivation in Edinburgh

Source: Adapted from SIMD 2012

### **7.3 SIMD Data Visualization as a CMV**

A brief study involving visualizing the SIMD data as a CMV was carried out in order to illustrate the influence of summary data in solving real life problems. The approach to this case study was qualitative. A focus group was chosen to allow participants discuss their perceptions and opinions towards a concept and highlight insights usually accessible when people interact with one another (Lindlof and Taylor, 2002). It is also an effective communication method where user stories written by participants will be used to develop personas for purposes of this study. However focus groups have some basic limitations such as observer dependency where the results obtained are subject to the researcher's interpretation or where group members would rather reach an agreement without critical evaluation in order to reduce conflict (Douglas, 2005). In

order to control the limitations, all participants were encouraged to write down their enquiries and each one was discussed as a topic with everyone participating.

### **7.3.1 Focus Group**

Ten participants who would directly benefit from the SIMD data were invited to an informal focus group to discuss the type of information they require and how it could be obtained. Seven participants, all parents of children in P6 with the common issue of making a choice of which high school their children will attend in the next few months turned up willingly. The SIMD data was briefly introduced to participants and everyone had a chance to browse through the website. The participants wrote down the nature of information they would like to obtain from the SIMD data on post-it notes. Most of the information as concerned high school catchment areas for their children and a discussion with the researcher took place where the participants contributed their ideas and critiques. The focus group lasted about 2 hours with a 10 minute coffee break in between. Data was collected for analysis.

### **7.3.2 User Stories**

User stories define requirements and provide adequate information to help designers draft realistic plans and implement them (Cohn, 2004). The data collected from the focus group is presented below as user stories:

“As a parent I want to search for the areas in the neighbourhood where the high school my child will attend is located and the background of the people my child will be interacting with”.

“As a parent I want to find details like the social and political setting of the neighbourhood where the school is located”. “As a parent I would like to search for the crime rate of different environments where the schools my child is likely to attend are located”. “I want to know if it is safe for my child to walk to school”.

“As a parent I want to find the average performance of children in their GCSE’s in the catchment area school”. “As a parent I want to search for the number of staff and students in the school and the teacher/student ratio”. “As a parent I would like to find details about the percentage of graduates from the catchment school that attend university”. “As a parent I want to know the schools in the catchment area that encourage sports, I want my child to learn a sport so he can get a scholarship to go to University in future”.

### **7.3.3 Personas**

Seven personas representing the user stories were created but only one was developed through the user stories documented for purposes of this study. One representative task was performed by this persona and a walkthrough of the steps taken by the representative persona was recorded.

Personas are fictional characters (Jenkinson, 2009) used to represent typical users of the SIMD interface. They are based on knowledge of real users and help the study identify the expectations and motivations of users interacting with the SIMD visualization.

Persona 1: Education

User: Nickel Wilson, female, 37, Single mother, Midwife.



Background: Nickel is a single mother of two daughters aged 11 and 10, presently living in Inverness has just been offered a position as a senior midwife at Edinburgh royal infirmary hospital. She is looking to move to Edinburgh before Christmas.

Key Goals: As a working single mum, she will have to do obtain basic information online about family life in Edinburgh.

Usage Scenario: Nickel would like to find a Roman Catholic high school by catchment area easily accessible to her kids by foot, bus or car from their home.

*Walk through of Persona task.*

The walk through was performed to illustrate the steps the parent of a P6 child would take while searching for a high school for their child. As a parent of a P6 child it was easier to make contact with other parents of children in Primary 6 to discuss mutual concerns about finding suitable high schools for our children within and outside the catchment areas provided by the government.

Representative Persona: Education

User: Nickel Wilson, female, 37, Single mother, Midwife.

*Background:* Nickel is a single mother of two daughters aged 11 and 10, presently living in Inverness and has just been offered a position as a senior midwife at Edinburgh royal infirmary Edinburgh. She is looking to move to Edinburgh before Christmas.

*Key Goals:* To find a Roman Catholic high school with a high percentage of its students attending university after graduation so her children will have adequate support to obtain GCSE grades and qualify to attend university

*Usage Scenario:* Nickel would like to find all Roman Catholic high schools near her home within 10 to 15 minutes drive from her house so her kids can get to school easily.

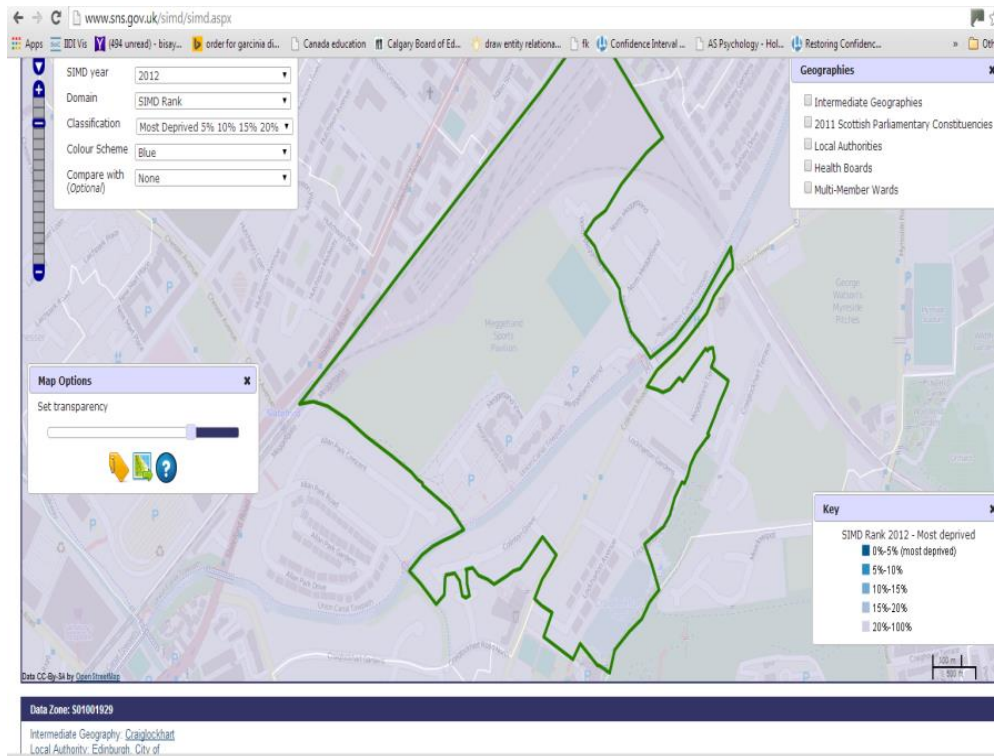


Figure 7. 2: Map showing data zone S01001929 Craiglockhart, Edinburgh

Source: Adapted from [www.sns.gov.uk](http://www.sns.gov.uk)

The diagram in figure 7.2 is an interactive map showing areas within a datazone. This map allows people locate and view areas within datazone and find the percentage deprivation of the area. It provides people with some of the information they require to choose places to live, work and school. The steps the persona, Nickel took to navigate through, visualise and retrieve information from the SIMD data set are outlined below:

Step 1: Nickel Logs onto the [www.sns.gov.uk/simd/simd.aspx](http://www.sns.gov.uk/simd/simd.aspx) interface (figure 7.2)

Entered “Edinburgh” in the search box provided.

22 pages giving the list of datazones in Edinburgh are displayed. It is ordered according to ranks, the least being 40 and highest 6505

The ranking of location indicate that an area with a higher ranking in education means better education is provided in these datazones. She clicks on page 22 and then on datazone S01001929, Craiglockhart, the location with the highest rank of 6505.

The map below is displayed showing areas including Meggetland, Lockharton and Colinton. She clicks on the map but it does not provide further information. She clicks outside the map and information on the datazone and its rank is displayed above the map.

Nickel wants to compare education ranks of different datazones in Edinburgh. She clicks 'compare with' on the "filter results" button which allows a user compare the ranks of different years. Nickel clicks on other areas such as Morningside (SIMD rank 6452) to compare to Craiglockhart. Nickel goes back to "filter results" and clicks on "education" to compare "Craiglockhart" and "Colinton" but obtains the same figure. The user is unable to complete the task.

Findings:

Nickel is unable to compare education ranks of different areas in Edinburgh.

The filter results button gives the same results for all the domains irrespective of which is selected (income, employment, health, education, geographic access, housing and crime).

Nickel could navigate to and visualise domain ranks, SIMD ranks and datazone information (data) for places in Edinburgh but she was unable to navigate to, visualise and compare domain ranks for different areas. She could not complete her task or make decisions regarding her children’s education based on data the SIMD provided on educational ranks for the datazones.

The information Nickel sought could not be obtained from the SIMD website and she was unable to retrieve the required information. The CMV provided to visualise the SIMD data (figure 7.3) was utilized however the information retrieved was also insufficient. The argument of this research is that the challenge could be completed if the CMV provides summary data in order to allow people access data and metadata required to support people make informed choices.

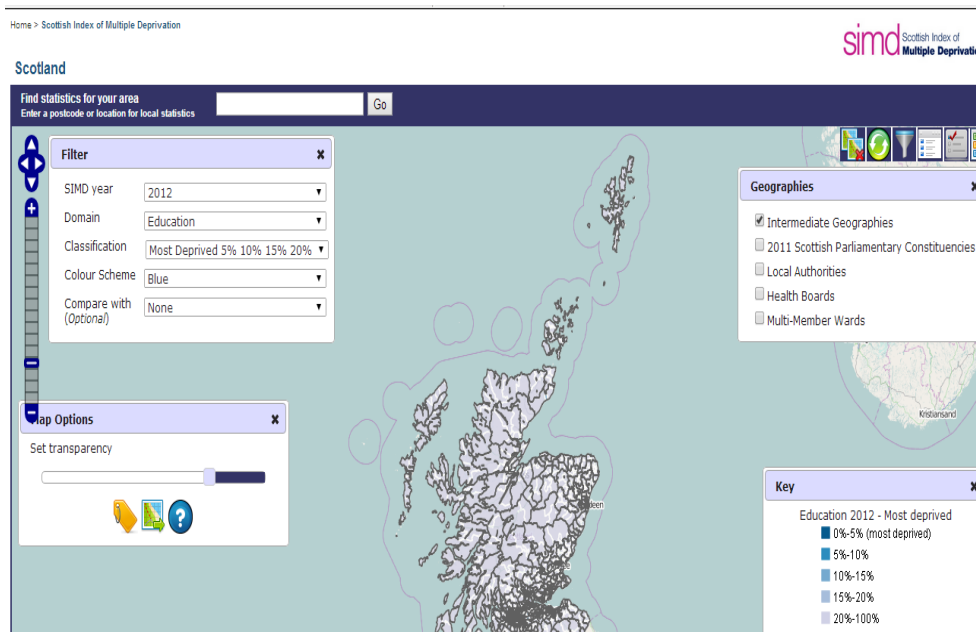


Figure 7. 3: A CMV for SIMD data

Source: Adapted from [www.sns.gov.uk](http://www.sns.gov.uk)

During the focus group, a participant suggested that parents could search for information on the map for catchment areas, which is one of the standard ways the government makes data available. Nickel narrowed the information she wanted to retrieve to one single school, Holyrood and worked with the map in figure 7.4 but was unable to navigate through successfully to retrieve adequate data and make informed decisions. Data such as names of areas in catchment schools and feeder primaries are provided on the map nevertheless more detailed information like the number of people that come from out of catchment areas to attend schools in specific areas or the number of people who leave their catchment boundaries to attend other schools is not provided. This type of information could help the parent who requires information such as the background of the children attending the same school with her child make informed decisions.

First, the areas within the secondary catchment area are concentrated and it would be challenging to successfully navigate through the information space with the data provided. A CMV designed with summary data could make the design more navigable, provide effective visualization and allow people retrieve adequate information on the catchment area map.

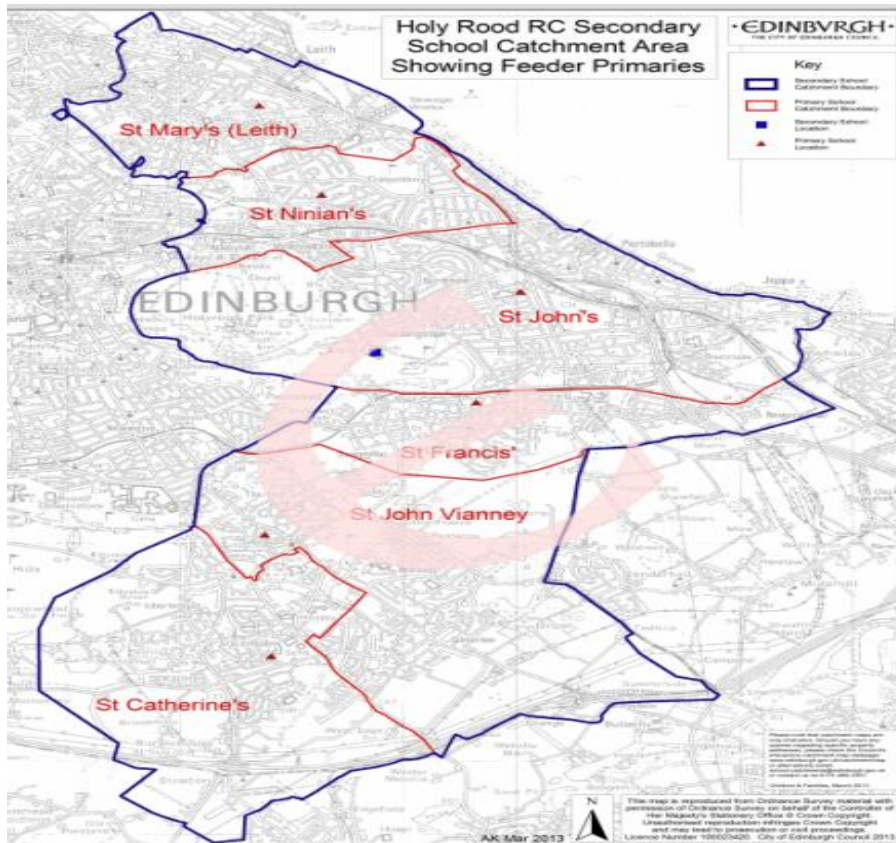


Figure 7. 4: Holy Rood RC secondary school catchment area showing feeder primaries

Source: Adapted from [www.edinburgh.gov.uk](http://www.edinburgh.gov.uk)

In 2012, Alasdair Rae redesigned the map of most deprived areas in Scotland by implementing a method similar to some aspects of the guidelines proposed by this research (page 282). He provided metadata which improved people's ability to navigate through the dataset. This is an additional illustration of the change incorporating summary data can bring about in people's ability to navigate, visualize and retrieve adequate data.

The research proposes that a partial redesign of the SIMD CMV based on the guidelines proposed by this research will improve people's ability to navigate the CMV in order to visualize and retrieve adequate information.

*Deciding the main focus* of the SIMD data will give a clear understanding of the *attributes* the designer is required *to provide access to*. A *summary data* (metadata about content) of these attributes can then be made available, in addition to an *interface key* (metadata about visual aspects). Implementing these design guidelines will make the SIMD CMV more flexible, navigable and allow people *appropriate* the interface. Based on this, the research makes the following proposal.

Title: Proposed recommendations for redesigning Navigable CMV's.

Proposal: Partial Redesign of SIMD CMV

Proposal details

The Scottish Index of Multiple Deprivation (SIMD) is a rich set of data that gives information about datazones in Scotland. It makes available information about health, unemployment, education, crime, poverty, housing, population, and the community. The information it holds has the potential to guide people make decisions concerning where to live, work or which school their children should attend. However, one challenge encountered by people who require information from the SIMD CMV is navigation; hence the rich data available on the interface is difficult to access. The research proposes that implementing the set of guidelines developed through its empirical studies will bring about a more navigable SIMD CMV which will allow people visualize and retrieve required information from the CMV.

Proposal objectives

The main objective of this proposal is to develop an SIMD CMV that will allow people achieve successful navigation between data and views on the interface, visualize data

effectively, interpret these data correctly and retrieve adequate information. In order to redesign the SIMD CMV, the following guidelines should be implemented:

1. The main focus of the SIMD data should be decided.
2. Access to the attributes that have a many-to-one relationship with the main focus should be provided.
3. A summary data (metadata about content) should be provided
4. An interface key (metadata about visual aspects) should be made available
5. The SIMD CMV should be designed for appropriation

#### Proposal benefits

The estimated benefits of the redesign proposed include making the SIMD interface flexible in order to cater for a broad range of tasks different people may require. In addition, it will create a more navigable interface, improve people's ability to navigate between data and retrieve information quicker and easily. Furthermore, it will afford people the opportunity to appropriate the information space.

### ***7.4 Chapter Discussion***

The main objective of this brief case study is to evaluate a different domain to find evidence that supports the findings made by this research that summary data improves navigation. Three separate maps (figure 7.2, 7.3 and 7.4) were used however, the persona was unable to retrieve adequate information. It is evident that the SIMD data is a rich set of data that can provide a wide range of information however, it does not offer the user sufficient support to navigate through the information space in order to visualise the data set and retrieve adequate information. One of the reasons is because people do not have access to data about the attributes



of the main focus of the data set. First, the focal point of the data set needs to be determined. Next, the attributes that have a many -to-one relationship with the focal point of the data set needs to be determined and summary data provided accordingly in order to create access to the different levels of information available.

The example illustrated above support the findings made by this research. The main difference between the maps, figure A and B (appendix 2, page) is that the map in figure B, provides data and when an area is selected it provides metadata about that specific area as illustrated in fig C. This is an example of summary data, it supports navigation, makes it easier to visualize and to retrieve information. When summary data is available (Fig B, appendix 2, page), the data set is easier to navigate, visualize and retrieve information as a result of the metadata made available. People can navigate to and access different levels of data in order to compare SIMD ranks more easily when summary data is available (figure B, appendix 2, page) than when it is absent (fig A, appendix 2, page ).

It is important to note that the type of data provided on an information space influences the navigation behaviour and ability of people. One of the issues raised by this research is that CMVs should be designed to allow people move from a view of data to a view of metadata and back again. The earlier studies carried out have identified that summary data, which is data and metadata about certain attributes of a data set support successful navigation and advance effective visualization.

The user stories developed from the focus group give an insight to the type of information parents would like to obtain as regards choosing a high school for their

children. In order for parents to obtain this information, a CMV that lets them navigate successfully and visualise the necessary data would be useful.

Extra data or controls supported by the metadata can be provided on the CMV to help people navigate to and choose data where required. In addition it will support people who seek answers to more general questions and bridge the gap between tasks that designers imagine that people will want to do and the real life questions people need to answer.

### ***7.5 Conclusion:***

Looking at the result of the brief case study, it can be inferred that, though the SIMD is a rich data set, people are unable retrieve adequate information. It is plausible to say that If the SIMD is visualized through a CMV designed to display a summary data of the data zone codes and data zone names in addition to the ranks it would support successful navigation, effective visualization and information retrieval. Furthermore, if the interactive map is designed to display the data zone names when a mouse over is performed it will enhance navigation and allow people select the data zones they want to compare more easily and quickly. People can navigate to and visualise the data more effectively and relate easily with names rather than codes.

## **Chapter 8 Discussion**

### ***8.1 Introduction***

The primary objective of this research is to identify new methods of navigation in order to improve navigation on information spaces such as CMVs for management information systems. The study started by examining the concept of navigation of information spaces with attention given to several navigation problems user's encounter on information spaces, which is summed up as their inability to achieve successful navigation. The challenge is expressed by user's failure to identify objects, find their way and discover existing and new relationships easily and quickly in order to visualise and retrieve adequate information. This affects their aptitude to visualise effectively and retrieve required information. Subsequently, based on empirical evidence, we introduce and implement a novel method; the concept of navigation of information space to support people to navigate CMVs in order to achieve successful navigation. Finally, a set of design guidelines tailored to CMVs following evaluation is proposed as additional design guidelines.

This chapter discusses the approach developed by this research to address the navigation challenges identified, the usefulness of the technique and the design adjustments made on the current CMV to provide adequate support for users.

### ***8.2 Information Spaces and Navigation***

This research has identified and introduced the concept of navigation of information spaces (Benyon, 1992) as an approach to improve navigation of information spaces with particular reference to CMVs.

The initiative of the concept is to design for people living in their information spaces, which involves designing CMVs in a way similar to how physical spaces are designed. The concept of navigation of information spaces sees people as existing in their information spaces. The World is currently described as a global village; people shop, acquire an education, work and communicate within various information spaces regularly. Many people carry out their daily activities within these spaces; hence people live in their information spaces. Section 2.4.3.3 explains that people who live in their information spaces want to carry out activities and not just tasks, unlike when people exist outside their information spaces and just access the information space to carry out a task. In order to achieve a task people need to carry out various activities. Rather than focus on user tasks, the concept of navigation as a theory of interaction encompasses a broad scope of activities such as object identification, wayfinding and exploration in order to create the flexibility required for people to carry out tasks successfully. A CMV designed to encompass these navigation activities offer the flexibility required for a CMV to be navigable. For instance, to locate a place, people undertake various activities; first, they identify the place with the aid of various objects, then they find a route to take them to their destination with the support offered by landmarks and signs and finally they explore the environment to get the information they require. As they go through these different phases, they perform various tasks depending on whether the task requires an object identification, a wayfinding or exploration activity.

The flexibility provided by navigation activities creates the ease of navigation required for data to be visualized effectively and information retrieved efficiently.

In chapter 2 information spaces is explained as a space where people carry out various activities and need to move from one point to another to achieve their purpose. Navigation is therefore an essential aspect of any information space.

The IIDI CMV Interface C was designed to afford the different aspects of navigation characterised as object identification, wayfinding and exploration in order to provide the opportunity for users to carry out their tasks flexibly and in a manner similar to the real world they are familiar with. The interface was designed to incorporate summary data (data provided along with its metadata), to make data available on different levels. The availability of different levels of data made it possible for people to navigate easily from one view to the other and from a level of data to a level of metadata. An interface key (data about the visual aspects of the information space) explaining the coding scheme of interface C was also made available. The key gave the information seekers cues; it facilitated their ability to gain knowledge of the interface and interpret the knowledge effectively and retrieve adequate information. The results presented in table 5.9 illustrates that users were able to retrieve more information on interface C than on interface A which lacks summary data and metadata about the coding system.

According to Andrienko (2006), a task consists of two aspects; a target and a constraint. The target is the information the user wants to retrieve while the constraint is the condition that needs to be fulfilled before the target is realized. The navigation activities can be likened to these constraints, a user who needs to perform identification tasks, or tasks that need people to compare objects or find locations (Andrienko, 2003) will need to fulfil conditions that involve object identification, exploration and wayfinding activities.

The concept of navigation of information space affords users the opportunity to navigate easily in order to interact with a dataset on different levels, it affords them the opportunity to navigate through the dataset and fulfil the conditions required to achieve their tasks and it supports them to gain knowledge to interpret and understand a data set.

The IIDI data CMV, interface C was implemented with a summary data and an interface Key describing the coding scheme was provided. Users who performed tasks on this interface completed a wider range of tasks (Table 5.3), achieved faster task completion times and retrieved more information (table 5.4) and performed less clicks and mouseovers (table 5.9). The summary data supported users to navigate to and interact with information on different levels, while the interface key supported their interpretation and understanding of the data.

Looking at the case study discussed in chapter 7, it is evident that the provision of metadata improved the usefulness of the interactive map (Fig 7.6 & 7.7). It was easier to navigate to the required information and compare SIMD ranks with the map in figure 7.6 rather than 7.5. The main difference between the maps is the availability of metadata and the recognition that people need to move around and carry out activities in order to achieve their purpose on information spaces.

### **8.3 Metadata**

In order to implement the concept of navigation of information space, metadata about content (summary data) and metadata about visual aspects of the interface (Interface

key) was developed. The concept makes it possible to provide a level of flexibility that is absent in a task based design. It requires the designer to decide on the main focus of the data set and then on the key relationships based around that focus. Following this, the designer should provide access to the metadata of the main focus, both in terms of an interface key for the visualization and in terms of a summary data for the attributes that have a many-to-one relationship with the main focus of the data set.

Metadata is discussed in detail in chapter 2 as descriptive information which allows data to be retrieved, used or managed (Tweedie, 1997). It has been used in various ways to complement visualization and facilitate information retrieval. Some examples developed earlier include the Filmfinder (Ahlberg and Shneiderman, 1994), snap together visualization (North and Shneiderman, 2000) Medioviz (Grun, Gerken, Jetter, Konig and Reiterer (2005) and the Use of Metadata Visualization to Assist Information retrieval (McCormac, Parsons and Butavicius, 2007).

Other methods explored to improve navigation include spatial ability (Ahmed and Blunstein, 2005), lessons from the built environment (Benyon, 2005) and supporting users with appropriate metaphors (Benyon and Hook, 1997).

Many of these methods achieved an improvement in the speed at which tasks were performed by utilizing the usual ideas of metadata which include descriptive, intrinsic, administrative, structural metadata. However, these metadata types appear not to be sufficient to meet the current demands of information retrieval in information spaces. There is a need to provide details about other aspects of an information space to facilitate navigation, visualization and information retrieval. For instance, metadata about the visual aspects of the data, and metadata about the objects in the database

should be provided. Roberts (2007) suggests that CMVs could provide helpful hints to on how to use the system and provide additional information about the data to ease the difficulty users sometimes encounter. He explains that users find it difficult to understand how the system is structured such as which views are coordinated and what information each view communicates.

Section 6.2.2 discusses summary data which is defined as data about the attributes that have a many-to-one relationship with the main focus of any data set, while the interface key is metadata about the visual aspects or the coding scheme of the design. The availability of these types of metadata contributes significantly to people achieving successful navigation on the IIDI CMV interface C.

In the case of the IIDI CMV, the main focus of the data set is people. The attributes that have a many-to-one relationship with people are centres, publications, grants, and keywords. Summary data which is data about these attributes was made available on the redesigned interface C (figure 6.1) which made it possible for people to gain access to the metadata of the main focus (people). Data about the attributes is metadata about the main focus. With summary data made available, people are able to navigate to and visualize information on different levels on a many-to-one basis rather than on the same level between relations on a one-to-one basis. They are able to move between data and from a view of data to a view of metadata and back again based on the focus of their query. The objective of the summary data is to capture many-to-one relationships on the CMV to improve navigation as the existing one-to-one relationship on the current CMV do not offer users sufficient support to achieve successful navigation.



The design approach brings concepts and methods from navigation in the real world to navigation in information spaces. An example is this is the second technique employed in designing interface C, the interface key. It is data about the visual aspects of the design. For instance, interface C provides an interface key which can be compared to information signs provided in the real world to support people understand the data set, interpret it and recognise when they have reached their destination.

The interface key is not a novel design however, its capacity has not been maximised, and many designs do not provide this valuable tool. The participants that undertook the second empirical study described in chapter 5 of this study confirm that the interface key, played a significant role in their ability to complete the given tasks. They claim that it helped them to understand and interpret interface C. The interface key provides information about the objects that are not the main focus of the database. It explains the visual aspects of the interface such as the images displayed which facilitates understanding and gives people direction on where to go in order to obtain the required information.

The success of the approach is seen in the overall picture of the IIDI CMV in Interface C. It is flexible enough to allow people carry out tasks, provides metadata about the main focus of the database in terms of a summary data and metadata about the visual aspects of the database in terms of a key in order to support people navigate, visualise and retrieve information effectively.

The availability of metadata about content and metadata about visual aspects of the data summary data, have brought about fundamental benefits to the IIDI CMV. According to the users who evaluated Interface C in chapter 6, the IIDI CMV provides a rich description of data which helps people understand the general idea it presents. It also makes it possible for people to aggregate and compare data about several pieces of information. The description of the visual aspects and the detailed information the summary data provide improved their ability to understand and interpret the interface. A user who understands the interface will be able to interpret the information space in order to identify objects, find their way to required destinations and explore the data set.

#### ***8.4 Evaluating Interface C***

Results from the empirical study carried out in chapter 4 and 5 show an improvement in people's ability to achieve successful navigation. Though CMVs are powerful interactive visualization interfaces with more than one view which enables people visualise data from different perspectives, people still face challenges navigating these CMVs as they do not provide a platform for task to be carried out flexibly. The results obtained from this research demonstrate that CMVs that incorporate navigation activities are flexible and enable people navigate information spaces in a slightly different way.

The findings obtained from the empirical study demonstrate the following: The original IIDI CMV Interface A, (figure 5.1) is designed to focus more on user tasks. It supports a number of user tasks that require identification which require object identification activities. However it is difficult to carry out tasks that require exploration and

wayfinding. One main problem of the design is that it lacks a key to the coding system used to encode centres or to make out that the lines linking the nodes in the upper left view indicate the Co-published relationship and that the thickness of the lines represents the volume of Co-publications.

Consequently, participants who undertook tasks on Interface A, found it challenging to achieve their purpose. They had to do a lot of explorations which involved a large number of clicks and mouseovers. In addition, they had to aggregate a lot of data themselves which resulted in inadequate information retrieval, longer task conclusion times and inability to perform a wide variety of tasks.

It is challenging to find data about centres, publications and grants or answer questions relating to metadata when Interface A (fig 5.1) is used. In order to resolve this, a visualization of metadata about centres is introduced. The bar chart in the URV of interface A is replaced with a matrix representation of centre collaborations (UCV) and a network representation of centres in the URV to produce Interface C.

Interface C incorporates navigation activities which moves the CMV from a task based focus to a more flexible interface. It provides the first piece of metadata which is metadata about the coding system required by any visualization. Then a summary data which is a more detailed metadata is provided. Access to metadata about people, which is the main focus of the IIDI dataset, is available on the UCV. Labels which replicate information signs in the real world are provided on the UCV; they determine what content each view will present and guide the users to locate their required destination.

Interface C provides different levels of information and allows users switch from a view of data to a view of metadata and back again. The major improvement of Interface C is seen in the results obtained. Participants obtained faster task conclusion times, retrieved more correct data and were able to carry out a broad category of tasks. The amount of exploration they had to undertake was also reduced as seen in the reduced number of mouseovers and clicks.

During the second empirical study, as participants performed tasks on Interface C, the researcher observed that they were more focused. They were less confused and able to make meaningful movements. In addition, participants pointed out that they made less effort to navigate on interface C than interface A, they claim they could easily make out what information the different views held. They went further to explain that they did not need to aggregate as much data or make calculations on interface C as compared to interface A, and they were able to carry out two tasks concurrently. This suggests that interface C has the potential to reduce navigation challenges like lost in hyperspace and reduce cognitive overhead discussed in chapter 2.

Looking at the results obtained from the empirical studies undertaken and comparing the interface A and C, it is apparent that the research hypothesis stated in section 1.5 is acceptable. The results of the first empirical study reported in chapter 4, demonstrate that participants who worked on interface B obtained faster conclusion times (section 4.4.1.2), completed a wider category of tasks (table 4.5) and carried out less explorations in comparison to Interface A, this can be seen in a reduced number of clicks and mouseovers (table 4.7).

The result obtained from the second empirical study was similar. Participants on interface A had longer tasks completion times (table 5.4), carried out a smaller category of tasks (table 5.3) and obtained a higher number of clicks and mouseovers (table 5.9) which shows that they had to perform lengthy explorations. In contrast on interface C, participants obtained faster task conclusion times, carried out a wider range of tasks and performed less clicks and mouseovers irrespective of the order (A-C order or C-A order) in which participants undertook tasks.

Access to metadata of the main focus of the data set on interface C brought about an improvement on participant's ability to navigate between data, visualise data effectively and retrieve adequate information. Although users are interested in different aspects of information and follow different routes to similar locations on information spaces, their ability to access metadata supported them to meet their general needs. The result obtained demonstrates that users retrieved more information from interface C than from A (fig 5.4). This correlates the initial findings from empirical work one, that provision of summary data has a positive influence on user's navigation abilities.

An evaluation of interface C implemented with the guidelines proposed by this research in chapter 6 was rated favourably by the users. Visible improvements were seen on user's ability to navigate the CMV as high success rates and task completion times were obtained. The subjective rating of the interface by users was positive. They also identified the additional views (UCV and URV) as useful views. The findings suggest that these additional guidelines support users to achieve successful navigation, effective visualization and adequate information retrieval.

Some participants demonstrated their ability to appropriate the information space by designing specific tasks for themselves and retrieving the required information from interface C. This is an indication that designing for people who live in their information spaces and providing sufficient support for navigation activities to be performed could open up the possibility of people being able to appropriate CMVs.

One person's data is likely to be another person's metadata, for instance, an institute director would want to look at data at the institute level and may need to extrapolate future goals from past performance. A centre director requires data at the centre level and on the other hand someone else may be interested in the subjects that the keywords points to. Furthermore, research students may need to analyse publications. These different levels of information can be provided by making content metadata and metadata about visual aspects of the information space available.

In the brief case study carried out with the SIMD data, it is evident that navigation of the SIMD data in figure 7.5 improved when metadata was made available as seen in figures 7.6 & 7.7.

General guidelines exist for user interfaces but the design process of CMVs could be advanced by identifying and evaluating specific guidelines tailored to suit these information spaces. This will ensure that CMV designers have the support they require when creating CMVs. It will also ensure the development of navigable CMVs that will bring about user satisfaction.

The empirical findings to emerge from this research suggest a role for a summary data and metadata about other aspects of the interface such as interface key in creating

navigable CMVs which will directly influence visualization and information retrieval. It contributes to several aspects of existing data visualization methods by providing additional principles to accomplish effective visualization. A rundown of the outcome is given below.

### Model

Investigations carried out on the IIDl data CMV (interface A) brought about the development of a different model (interface B). Further design adjustments were made on interface B to create interface C. The proposed model (Interface C) incorporates navigation activities, provides summary data and an interface key, which interface A, the original model lacks.

### Summary

The strong point of this theory is that it would reduce navigation problems faced by users while visualising data on CMVs. The research survey shows that more people were able to visualise the IIDl data and retrieve adequate information when they used interface C in comparison with interface A.

## **8.5 Conclusion**

The issue of navigation in information space is central in interactive visualization design. The research carried out suggests that if designers create CMV information spaces without integrating navigation support, then there will be significant difference in performance between individuals as regards their ability to achieve successful navigation, effective visualization and retrieve adequate information.

The big issue is that data is used for different purposes and one person's data is another person's metadata. We do not know what tasks people will want to perform on any datasets. However people should be able to navigate any given data set. Providing data and metadata is useful to support people to interpret and manage information. Understanding the metadata goes on further to enable adequate information retrieval.

There is a need for the development of navigable information spaces. This research has identified that if CMVs provide access to metadata; metadata about design aspects (visual data) and metadata about content (summary data) it will incorporate more tasks, reduce navigation steps and improve user effectiveness.



## Chapter 9 Conclusion

### ***9.1 Introduction***

Chapter nine sums up the general contributions this work will make to data visualization. It points out the limitations of the research and outlines further work to be carried out in future. This research focuses on the concept of navigation relating to the domain of CMV information spaces. The research objective is to create flexible CMVs to allow people perform the task they require. The design is based on understanding data alongside its metadata and gaining knowledge of how humans navigate their information space. The design approach brings concepts from navigation (object identification, wayfinding and exploration) in the real world to navigation of data sets in information spaces. This CMV design has been developed to allow people move from the view of data to a view of metadata and back again, depending on the focus of their queries. This involves a move from providing the CMV of a relation (or interrelated relations at the same level) to a visualization that provides ways to move between data and metadata. The design has been evaluated by users and generally it has been found to offer essential benefits over the current CMV.

The evaluation demonstrates that the navigation of information space approach matches well with user needs on CMVs and users easily identify with the methodology. The CMV was designed to support users move easily from one level of information to another within the information space, thus achieving successful navigation, effective visualization and adequate information retrieval.

## **9.2 Main Contributions**

The major contributions of this work are:

The research has undertaken empirical studies to develop new methods that address challenges specific to data visualization on a CMV of a management information system.

### Concept

The research developed the concept of navigation of information spaces as an approach to improve navigation on CMVs particularly management information systems. The concept when looked at as a theory of interaction provides a platform for CMVs to be designed for people living in their information spaces. This makes it possible to develop a CMV that does not focus on user tasks but encompasses a wide range of activities; object identification, wayfinding and exploration. The resulting CMV is flexible and allows users navigate easily and answer more queries in less time.

### Design

The initial IIDi CMV (Interface A) was redesigned by providing metadata about content in terms of a summary data of the attributes that have a many-to-one relationship with the main focus of the database in order to create different levels of information, provide a view of data and metadata and meet a broad range of user needs. The resulting interface C is easier to navigate, visualize and retrieve information.

Next, metadata about the visual aspects in terms of an interface key was made available to users to give them useful cues about what each view holds and help them interpret and understand the interface.

## Evaluation

The redesigned Interface of the IIDI CMV was evaluated with the objective of rating its performance, benefits and acceptance. The user based study undertaken resulted in a positive response from participants. The redesigned IIDI CMV is currently in use in the IIDI institute.

## Design Guidelines

The research has examined navigation based activities on information spaces to propose a simple but strong design procedure to support future development of CMVs. The proposed CMV design will act as a framework to support successful navigation by users to visualise a management information system data and advance the utilization of CMVs for data visualization. The research has proposed and performed empirical evaluations on the proposed guidelines for designing CMVs. These guidelines include:

7. Decide the main focus of the data set
8. Provide access to attributes of the main focus of data set
9. Provide metadata for content in terms of a summary data of the attributes that have a many-to-one relationship with the main focus of the data set.
10. Provide metadata for the visual aspects of the information space such as an interface key

## Design for appropriation

The research has utilized the approach of navigation of information spaces to address issues of navigation on CMVs by providing access to metadata.

The approach illustrates how a CMV which incorporates activities like object identification, wayfinding and exploration facilitate movement, visualization and information retrieval on CMVs.

It demonstrates how task completion times, task types, navigation steps, success rates and the features provided on the CMV information space are influenced by navigation. The design focuses more on the user and navigation rather than on tasks and the system.

The design is developed to improve human performance in the area of navigation. The concept of navigation of information space applied gives the design the flexibility required to allow users appropriate the information space. The research has emphasized the support users can obtain from an information space that provides an interface Key. It is of benefit to users to have an interface key provided as a principle in creating information spaces.

#### Limitations

The work presented in this thesis is an initial attempt to improve navigation, visualization and information retrieval on information spaces such as CMVs, in particular management information systems. with the concept of navigation of information space approach.

The design is in its preliminary stage, as a result the Interface key has not been integrated into the system and no comparative study of the evaluation was carried out.

### **9.3 Future Work**

During the thesis, areas for future work have been identified

Utilizing the CMV for predictive purposes

The CMV design could be augmented to allow for predictive purposes. For instance, a centre director of the IIDI institute is asked to supply targets of publications and grants for the year 2020. How does he extrapolate? How does a director tell that he can improve the number of papers written in 2014 in 2020? How will summary data enable the director tell the number of papers a specific author has written, has left unpublished and so on. How can the CMV be designed to allow for extrapolation?

Reducing cost of querying huge databases

The additional guidelines apply to administrative and structural metadata which is the type of metadata used in MIS systems of large organisations. The desire to query the huge databases owned by these organisations exist but the official procedure is lengthy and it is expensive to the fund. How will these CMVs designed with summary data allow people prototype things that may become queries? At the same time shortening the procedure and reducing the cost involved?

Appropriation of information space

Continued research is required to explore CMVs in more detail in order to produce a more formal CMV that allows for appropriation of space. Appropriation of information space allows users to see the space as their own which agrees with the concept of navigation space that sees people as living in their information spaces.

Interface Key

An interface key is like the door of the information space. It is a survey representation of the interface. Novel research is required to establish its relationship to affordances. Interface key is not a novel invention; however its potential is not fully utilized. The ability of the interface key to afford opportunities for interaction could yield new forms of interaction.

#### **9.4 Conclusion**

Although coordinate multiple views have maintained a platform for understanding and interpreting data, navigating these views has remained arduous. The existing CMVs have not sufficiently recognized navigation as an aspect of data visualization that could enhance interaction with data and possibly facilitate understanding and interpretation of data. This research has looked at the main significance of navigation to data visualization and the limitations of CMVs in the area of navigation.

Going back to the hypothesis proposed by this research it is now feasible to state that designing a CMV to provide metadata about content and metadata about the visual aspects supports successful navigation. This novel method of navigation identified aims at advancing the way people move through CMV information spaces in order to achieve successful navigation which supports effective data visualization and adequate information retrieval.

## References

Alberts, B., & Wulf, W. A. (2002) Making the Nation Safer: The Role of Science and Technology in Countering Terrorism, Nat'l Academies Press.  
<http://www.nap.edu/openbook/0309084814/html>

Anderson, R.J. (2001) Engineering: A Guide to building dependable distributed systems. 2nd edition. Wiley Computer Publishing. ISBN 0-471-38922-6

Andreano, J.M., & Cahill, L. (2009) Sex influences on the neurobiology of learning and memory. *Learn Mem.* 16(4):248-66.

Andrienko, G., & Andrienko, N. (1999) Making a GIS Intelligent: CommonGIS project view. In AGILE99, pages 19–24, Rome.

Andrienko, N., & Andrienko, G. (2003) Coordinated Views for Informed Spatial Decision-making. In CMV '03: Proceedings of the conference on Coordinated and Multiple Views in Exploratory Visualization, page 44, Washington, DC, USA. IEEE Computer Society.

Andrienko, N., Andrienko, G., & Gatalsky P. (2003) Exploratory Spatio – temporal visualization: an analytical view, *Journal of visual languages and computing* 14 (2003) 503-504.

Andrienko, N., & Andrienko, G. (2006). Exploratory Analysis of Spatial and Temporal Data: a systematic approach. New York: Springer.

Antonio, M., & Rada, R. (1997) Activities and communication modes. *Int. J. Hum.-Comput. Stud.* 46(2-3): 469-483

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

Arora, S. (2013) Ubiquitous Coordinated and Multiple Views, Msc. Dissertation, Dalhousie University, Halifax, Nova Scotia.

Arthur, P., & Passini, R. (1992) *Wayfinding-People, Signs, and Architecture*. McGraw-Hill, New York

Bachman, C.W. (1975) Trends in database management. *Proc., AFIPS 1975 NCC*, Vol. 44, AFIPS Press, Montvale, N.J., pp. 569-576.

Bagley, P. (1968) *Extension of programming language concepts*, Philadelphia: University City Science Centre

Bailey, R. A. (2008) *Design of comparative experiments*. Cambridge University Press ISBN 978-0-521-68357-9. MR 2422352.



Baldonado, M.Q.W., Woodruff, A., & Kuchinsky, A. (2000) Guidelines for Using Multiple Views in Information Visualization. In *Advanced Visual Interfaces*, pages 110–119.

Barker, I. (2005) Information Scent: helping people find the content they want. *Usability and information architecture* article.

Baudisch, P., Lee, B., & Hanna, L. (2004) Fishnet, a fisheye web browser with search term popouts: a comparative evaluation with overview and linear view. *Proceedings of the working conference on Advanced Visual Interfaces (AVI'04)*, pages 133–140.

Becker, R. A., & Cleveland, W.S. (1987) Brushing scatterplots, *Technometrics*.

Benjamin B. B., & Shneiderman, B. (2003) *The craft of Information visualization: Readings and Reflections*, Morgan Kaufmann ISBN 1-55860-915-6.

Benyon, D. R. (1992) Task Analysis and Systems Design: The Discipline of Data. In *Interacting with Computers*, 4(2) 246 – 259.

Benyon, D., & Murray, D. (1993) Developing Adaptive Systems to Fit Individual Aptitudes, *Proc. IWIUI*.

Benyon, D., & Hook, K. (1997) Navigation in information space: supporting the individual. In: *INTERACT'97. IFIP Conference Proceedings*, 96.

Benyon, D. R. (1998a) Beyond Navigation as Metaphor . In Proceedings of 2nd European Conference on Digital Libraries, Crete, Sept.

Benyon, D. R. (1998b) Cognitive Ergonomics as Navigation in Information Space  
Ergonomics 41 (2) Feb. 153 - 156

Benyon D. R., Green, T. R. G., & Bental, D. (1999) Conceptual Modelling for Human Computer Interaction, Using ERMIA, Springer, London.

Benyon, D. R., & Imaz, M. (1999) Metaphors and Models: Conceptual Foundations of Representations for Interactive Systems Design. Human Computer Interaction, 14 159-189.

Benyon, D. (2001) Special Theme: Human computer Interaction. ERCIM news No. 46, July 2001 [contents] Navigating Information Spaces.

Benyon, D. (2001) The new HCI? Navigation of information space, Knowledge-Based Systems 14(8), 425-430.

Benyon, D. (2003) Human-Centred Interaction Design In: Cairncross, S., Varey, A., McEwan, T. (Eds.) Proceedings of 6th HCI Educators Workshop,.

Benyon, D., Turner, P., & Turner, S. (2005) Designing Interactive Systems. Maidenhead: Pearson Education.

Benyon, D. (2006) Navigating Information Space: Web site design and lessons from the built environment. *PsychNology Journal*, 4(1), 7 - 24.

Benyon, D. (2007) *Designing with Blends*. Cambridge, MA: MIT Press.

Benyon, D., Hook, K., & Nigay, L. (2010) Spaces of Interaction. In: *ACM/BCS Visions Conference*.

Berg, B.L. (2001) *Qualitative Research Methods for the Social Sciences*. 4th Edition. Ally & Bacon, 2001.

Blunstein, J., & Ahmed, I. (2005) Navigation in Information Space. In proceedings of the second international conference on web based communities (WBC'05). Pp 281-286. Algarve-Portugal.

Boukhelifa , N., Roberts, J.C., Roberts, J.P., & Rodgers, J. P. (2003) A coordination model for exploratory multi-view visualization 2003, *Coordinated and Multiple Views in Exploratory Visualization*, International Conference. Pages76-76 IEEE Computer Society

Bretherton, F. P., & Singley, P.T. (1994). "Metadata: A User's View, Proceedings of the International Conference on Very Large Data Bases (VLDB)". pp. 1091–1094.

Card, S.K., Robertson, G.G., & Mackinlay, J. D. (1991) The information visualizer, an information workspace. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'91), pages 181–188.

Card, S.K., Mackinlay, J. D., & Shneiderman. B. (1999) Readings in information visualization: using vision to think. Morgan Kaufmann Publishers Inc. San Francisco, CA, USA.

Card, S.K., Pirolli, P., Van Der Wege, M., Morrison, J. B., Reeder, R. W., Schraedley, P.K., & Boshart, J. (2001) Information Scent as a Driver of Web Behavior Graphs: Results of a Protocol Analysis Method for Web Usability. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'01), pages 498–505.

Cardwell, M., Clark, L., & Meldrum, C. (1996) Psychology for A level London Harper Collins.

Carmel, E., Crawford, S., & Chen, H. (1992) Browsing in Hypertext: a cognitive study. IEEE transactions on Systems, Man and Cybernetics, 22, 865}884.

Carr, D. B., Littlefield, R. J., & Nicholson, W. L. (1986) Scatterplot matrix techniques for large n. In Proceedings of the Seventeenth Symposium on the interface of Computer Sciences and Statistics, pages 297–306, New York, NY, USA. Elsevier North-Holland, Inc.

Cathro, W. (1997) "Metadata: an overview". Retrieved 6 January 2010.

Chalmers, M. (1999) Comparing Information Access Approaches. Journal of the Association for Science direct & Technology. Doi 10.1002.(SICI) 1097-4571 50: 12<1108. Copyright @ 1999. John Wiley & sons, inc.

Chen, P. (1976) "The Entity-Relationship Model--Toward a Unified View of Data". In: ACM Transactions on Database Systems 1/1/1976 ACM-Press ISSN 0362-5915, S. 9–36

Chen, P. (2002) "Entity-Relationship Modelling Historical Events, Future Trends, and Lessons Learned". In: Software Pioneers: Contributions to Software Engineering.

Chen, C. (2005) Top 10 Unsolved Information Visualization Problems. In visualization Viewpoints. IEEE Computer Graphics and Applications, July/August, 12-16.

Chen, J. L., & Stanney, K. M (1999). A theoretical model of wayfinding in virtual environments: Proposed strategies for navigational aiding. Presence: Tele-operators and Virtual Environments, 8(6), 671-685

Cohn, M. (2004) User stories Applied: For Agile Software Development. Addison-Wesley.

Chuah, M.C., Roth, S. F., Mattis, J., & Kolojechick, J. (1995) SDM: Selective dynamic manipulation of visualizations. In Proceedings of the ACM Symposium on User Interface Software and Technology, 3D User Interfaces, pages 61–70.

Codd, E. F. (1970) The Relational Model for Database management, Addison-Wesley Publishing Company, 1990.

Collins, C., & Carpendale, S. (2007) “VisLink: Revealing Relationships Amongst Visualizations,” IEEE Trans. on Visualization and Computer Graphics (Proc. of the IEEE Conf. on Information Visualization (InfoVis)), vol. 13, issue. 6, 2007.

Convertino, G., Chen, J., Yost, B., Ryu, Y-S., & North, C. (2003) Exploring context switching and cognition in dual-view coordinated visualizations. In Coordinated and Multiple Views in Exploratory Visualization (CMV03), volume 00, page 55, Los Alamitos, CA, USA. IEEE Computer Society

Coronel, C. (2010) Database Systems: Design, Implementation, and Management. Independence KY: South-Western/Cengage Learning. p. 65. ISBN 978-0-538-74884-1

Coutaz, J., Nigay, L., Salber, D., Blandford, A., May, J., & Young, R. M. (1995) Four easy Pieces for assessing the Usability of Multidimensional interaction: the CARE Properties. InterAct 1995.

Cooper-Kuhlen, E. (1992) Contextualising discourse: The prosody of interactive repair, the contextualization of language.

Creswell, J.W. (2011) Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research Hardcover – 28 Feb 2011 Los Angeles: SAGE publications.

Creswell, J.W. (2004) Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research. Pearson 2nd Edition

Dahlback, N., Hook, K., & Sjolinder, M. (1996) Spatial cognition in the mind and in the world-the case of hypermedia navigation. The eighteenth annual meeting of the cognitive science society. San Diego.

Dahlbäck, N., & Lonqvist, P. (2000) Navigation and Learning - A Cognitive Analysis of User Tasks in Electronic Information Spaces, Short paper in CHI '2000,, Amsterdam, Netherlands

Darken, R. P., & Sibert, J. L. (1996b) Wayfinding strategies and Behaviors in large virtual worlds. Proceedings of CHI+96, pp. 142}149. New York: ACM.

Darken, R.P., & Peterson, B. (2001). Spatial Orientation, Wayfinding, and Representation. Handbook of Virtual Environment Technology. Stanney, K. Ed.

Date, C. J. (1996) A guide to the SQL standard. Addison-Wesley. p. 206. ISBN 978-0201964264.

Davenport, E., Buckner, K., & Spence R. (1998) Making interactions visible: tools for social browsing, Workshop on Memory

Dennis, A., Wixom, B.H., & Roth, R.M. (2012) Systems Analysis and Design. Wiley.

Dieberger, A. (1997) Navigation metaphors and social navigation in information spaces, Position paper for the Conference on Computer Human Interaction (CHI)'97.

Diehl, S. (2002) Software Visualization. International Seminar. Revised Papers (LNCS Vol. 2269), Dagstuhl Castle, Germany, 20-25 May 2001 (Dagstuhl Seminar Proceedings).

Diehl, S. (2007) Software Visualization — Visualizing the Structure, Behaviour, and Evolution of Software. Springer, 2007, ISBN 978-3-540-46504-1

Dillion, A., McKnight, C., & Richardson, J. (1993) Space – the final chapter or why physical representation are not semantic intentions. In C. McKnight, Dillion, A. & Richardson, J. (Eds.); Hypertext: A psychological perspective (pp.169-191). New York: Ellis Horwood.

Dillon, A., & Vaughan, M. (1997) It's the journey and the destination: Shape and the emergent property of genre in evaluating digital documents, *New Review of Multimedia and Hypermedia*, 3, 91–106.



Dix, Alan J., Andrew, H., & Dongbo, X. (2003) "Post-web cognition: evolving knowledge strategies for global information environments." *International Journal of Web Engineering Technology* :1:1 (2003): 112.

Dix, A. (2007). *Designing for appropriation*. BCS-HCI'07 proceedings of the 21st British HCI group Annual Conference on People and computers: HCI...but not as we know it – Vol 2, pg 27 – 30. British computer Society, swinton, UK @ UK 2007. : ISBN:978-1-902505-95-4

Douglas, R. (2005). *Get back in the box: innovation from the inside out*, New York: Collins.

Downs, R., & Stea, D. (1973) "Cognitive Representations". In R. Downs, & D. Stea, (Eds.), *Image and Environment* (pp. 79-86). Chicago, Aldine.

Downs, R. M., & Stea, D. (1977) *Maps in Mind. Reflections on Cognitive mapping*. New York: Harper & Row.

Dourish, P., & Chalmers, M. (1994) *Running Out ofSpace: Models of Information Navigation*. Short paper presented at HCI'94 (Glasgow, UK).

Earl, R.B. (2013) "The Logic of Sampling". *The Practice of Social Research* (13th Ed.). Belmont, CA. Pp185-226. ISBN 1-133-14979-6

Edwards, D. M., & Hardman, L. (1999) 'Lost in Hyperspace'. Cognitive mapping and navigation in a hypertext environment. In Mc Allese, R., editor, Hypertext: theory into practice, pages 90 – 105. Intellect, Oxford.

Educause, (2010) Annual conference. Anaheim, California.  
[www.educause.edu/annual-conferenc/2010](http://www.educause.edu/annual-conferenc/2010)

Elmasri, R. (2011) Fundamentals of Database Systems. Addison-Wesley. pp. 73–74.

Elvins, T.T. (1997) Wayfinding 2: the lost world. November (1997) SIGGRAPH Computer Graphics, volume 31, issue 4. ACM.

Ericson, D., Johansson, J., & Cooper, M. (2005) Visual data analysis using tracked statistical measures within parallel coordinate representations. In CMV '05: Proceedings of the Coordinated and Multiple Views in Exploratory Visualization (CMV'05), pages 42–53, Washington, DC, USA, 2005. IEEE Computer Society.

Foltz, M. A. (1998) Designing Navigable Information Spaces. B.S. Computer Science and B.S. Electrical Engineering Washington University in St. Louis

Friendly, M. (2008) "Milestones in the history of Thematic Cartography, statistical graphics and data visualization".

Friedman, B., & Hendry, D.G. (2008) Theories and practice of design for interactive systems: Eight design perspectives in ten short weeks. Proceedings of ACM

Conference on Designing Interactive Systems (DIS'08), 435-444. New York: ACM Press.

Friedman, V. (2007) Modern Approaches.

[www.smashingmagazine.com/2007/08/02/data/visualization-modern-approaches/](http://www.smashingmagazine.com/2007/08/02/data/visualization-modern-approaches/)

Friedman, V. (2008) Data Visualization and Inforgraphics in: Graphics, Monday inspiration, January 14th, 2008.

Garling, T., Book, A., & Ergesen, N. (1982) Memory for the Spatial Layout of the Everyday Physical Environment: Different Rates of Acquisition of Different Types of Information. *Scandinavian Journal of Psychology*, 23, 23-35.

Geary, D. C., & Lin, J. (1998) Numerical cognition: Age-related differences in the speed of executing biologically-primary and biologically-secondary processes. *Experimental Aging Research*, 24, 101- 137.

Gillman, B. (2000) *The Research Interview*. London: Continuum.

Gottdiener, M., & Lagopoulos, A. Ph. (Eds.) (1986) *The City and the Sign*. Columbia University Press, New York  
Massey, D. (1996) *Space, Place and Gender* Blackwell, Oxford

Green, T. R. G., & Benyon, D. R. (1996) The skull beneath the skin; Entity-relationship modelling of Information Artefacts. *International Journal of Human-*

Computer Studies 44(6) 801-828

Grun, C., Gerken, J., Jetter, H., Konig, W., & Reiterer, H. (2005) MedioVis – A user centred library metadata browser. Proceedings in the ninth European Conference on Research and advanced Technology for Digital Libraries (ECDL 2005). Vienna, Australia.

Guenther, R., & Radebaugh J. (2014) National Information Standards Organization; (2004). Understanding Metadata. Bethesda, MD: NISO Press. ISBN 1-880124-62-9. Retrieved 2 April 2014.

Hall R.H. (1998) Psychology world creative commons. Web.mst.edu Copyright © 2009.

Halpern, D. F. (2000) Sex Differences in Cognitive Abilities (3rd Edition). Mahwah, NJ: Lawrence Erlbaum, Associates, Inc. Publishers.

Hearst, M. (2007) Search user Interfaces Cambridge University Press.

Heer, J., & Robertson, G.G. (2007) Animated Transitions in Statistical Data Graphics. Proceedings of the IEEE Symposium on Information Visualization (INFOVIS'07), 13(6):1240–1247.

Hinkelman, K., & Kempthorne, O. (2008) design and analysis of experiments: Introduction to experimental design. Wiley, New York, vol. 1

Hornbaek, K., & Frokjaer, E. (2001) Reading of electronic documents: the usability of linear, fisheye, and overview+ detail interfaces. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'01), pages 293–300.

Hornbaek, K., & Hertzum, M. (2007) Untangling the usability of fisheye menus. *ACM Transactions on Computer-Human Interaction (ToCHI)*, 14(2), 2007.

Hook, K., & Dahlbäck, N. (1997) Designing navigational aids for individuals. In *CHI97 Workshop on Navigation in Electronic Worlds*.

Ivan, V., & Meister, E. G. (2005) “Smart Visibility in Visualization”. In: *Computational Aesthetics in Graphics, Visualizations and Imaging*. L. Neumann et al. (Ed).

James, J. T., & Kristin, A. Cook. (Ed.) (2005). *Illuminating the Path: The R & D Agenda for visual Analytics*. National visualization and Analytics centre. P.30.

Javed, W., & Elmqvist, N. (2012) Exploring the design space of composite visualization. *Pacific Vis 2012*: 1-8.

Jenkinson, A. (2009) What happened to strategic segmentation? *Journal of direct, data and digital Marketing Practise*. 11:2, 124-139. Doi: 10.1057/dddmp.2009.27. Palgrave Macmillan, Basingstoke, UK.

Jul, S., & Furnas, G.W. (1997) Navigation in electronic worlds. A CHI'97 Workshop. New York: ACM. SIGCHI Bulletin, 29, 44 }49.

Keim, D. A. (2002) Information visualization and visual data mining. IEEE Transactions on Visualization and Computer Graphics, USA vol. 8 (Jan. March 2002), no 1, p 1 8, 67 refs.

Keim, D. A., Mansmann, F., Schneidewind, J., & Ziegler, H. (2006) Challenges in Visual Data Analysis. In Proceedings of the Tenth International Conference on Information Visualization, 9-16, 2006.

Keim, D. A., Kohlhammer, G., Ellis & Mannsmann F. (2010) Mastering The Information Age – Solving problems with Visual Analytics. Eurographics, 2010.

Kirk, R. (1982) Experimental Design: Procedures for the behavioural sciences, 2nd edition. Pacific Grove, California: Brooks/Cole Publishing Company.

Kimball, R. (2008) The Data Warehouse Lifecycle Toolkit (Second Edition ed.) New York: Wiley.

Kimura, D. (1999) Sex and Cognition, MIT Press, Cambridge, MA, USA.

Navigation in Hypermedia, 2002, Technical Report (Available at: <http://aetos.it.teithe.gr/~cs1msa/docs/navigation>)

Knight, C. (2002) System and Software Visualization. In Handbook of software Engineering & knowledge engineering. Vol. 2, Emerging technologies (Vol. 2): World Scientific Publishing Company.

Kjeldskov, J., Skov, M. B., & Stage J. (2004) Instant Data Analysis: Evaluating Usability in a Day. Proceedings of NordiCHI 2004. Tampere, Finland. ACM, pp. 233-240. DOI=10.1145/1028014.1028050

Kohlhammer, J., Keim, D. Margit, P. Giuseppe, S., & Andrienko, G. (2011) Solving problems with visual Analytics. Edited by Giacobino, E. & Pfeifer, R. In proceedings of the 2nd European Future Technologies Conference and Exhibition 2011 (FET 11).

Kosara, R., Healey, C.G., Interrante, V., Laidlaw, D.H., & Ware, C. (2003) Thoughts on user studies: Why, how and when. IEEE Computer graphics and Applications, 23(4), 20-25.

Kraaij, W., & Post, W. (2006) Task-based Evaluation of Exploratory Search Systems. (TNO, The Netherlands) Report on ACM SIGIR Workshop on Evaluating Exploratory Search Systems

Kuipers, B. (1982) The 'Map in the Head' Metaphor, Environment and Behaviour, 14, 202-220.

Landau, S. (2014) Highlights from Making sense of Snowden, Part II: Whats Significant in the NSA Revelations. *IEEE Security & Privacy*, 12(1), 62-64. Doi: 10.1109/MSP.2013.161

Lawrence, M., Lee, E.K., Cook, D., Hofmann, H., & Wurtele, E., (2006) Explorase: Exploratory data analysis of systems biology data. In *CMV '06: Proceedings of the conference on Coordinated & Multiple Views In Exploratory Visualization*, pages 14–20, Washington, DC, USA, 2006. IEEE Computer Society.

Lazar, J., Feng, J.H., & Hochheiser, H. (2010) *Research methods in Human computer Interaction*. John Wiley West sussex, United Kingdom.

Lindlof, T.R., & Taylor, B.C. (2002) *Qualitative Communications Research Methods*; 2nd Edition. Thousand Oaks, CA: Sage.

Lynch, K. (1960). *The Image of the City*. Cambridge, USA, MIT Press

Macaulay, C., Benyon, D. R., & Crerar, A. (2000) *Ethnography in Design: Uncovering New Artefacts* *International Journal of Human Computer Studies*

McCall, R., & Benyon, D. R. (2000) *Using ENISpace for designing 3D Virtual Environments*. In *International Workshop on Virtual Reality Systems*, Brest, France, July



Marcus, A., Feng, L., & Maletic, J. I. (2003) 3D representations for software visualization. Paper presented at the Proceedings of the 2003 ACM symposium on Software visualization, San Diego, California.

Mackinlay, Jock D. (1999) Readings in Information Visualization: using vision to think. Card, S.K., Shneiderman, B (eds.). Morgan Kaufmann publishers inc. Pp.686. ISBN 1-55860-533-9.

McCaleb, M. R. (1999) A Conceptual Data Model of Datum Systems. National Institute of Standards and Technology.

McCormac, A., Parsons, K., & Butavicus, M. (2007) The Use of Metadata Visualisation to Assist Information Retrieval. Australia.

Meiguins, S. B., Meiguins, G.A.S., Leandro, H.A., Augusto de Moraes, R.L., & Sergio Clayton, V.P. (2010) Coordinated and Multiple data Views in Augmented reality Environment, INTECH, Croatia, pp.230.

Morgan, G. (2004) SPSS for Introductory Statistics. Hillsdale: Lawrence Erlbaum. ISBN 0-8058-4789 B

Munro, A., Hook, K., & Benyon, D. R. (1999) Footprints In the Snow. In Munro, A., Hook, K. and Benyon, D. R. (eds.) Social Navigation of Information Space. Springer, London

Nielsen, J., & Mack, R. L. (Eds.) (1994). Usability Inspection Methods. John Wiley & Sons, New York, NY, ISBN 0-471-01877-5.

Nielsen, J. (1994) Usability engineering. Morgan Kaufman publishers. ISBN 0-12-518406-9

Nielsen, J. (1994) Usability Engineering. AP Professional Press, Boston.

Nielsen, J. (2001) E-Commerce user experience. ISBN 0-970-60720-2

Neisser, U. (1976). Cognition and reality: Principles and implications of cognitive psychology. New York: Freeman. ISBN 978-0716704775

North, C., & Shneiderman, B. (1997) "A Taxonomy of Multiple Window Coordinations." Univ. Maryland Computer Science Dept. Technical Report #CS-TR-3854, 1997

North, C., & Shneiderman, B. (1999a). Snap-together visualization: Coordinating multiple views to explore information (technical Report No. #CS-TR-4020). University of Maryland Computer Science Department, Maryland, USA.

North, C. & Shneiderman, B. (2000) Snap-together visualization: A user interface for coordinating visualizations via relational schemata. In Proceedings of Advanced Visual Interfaces, pages 128–135, Italy.

Oehlert, G. (2000) A first course in design and Analysis of Experiments. New York: Freeman and company

Olsen, G. (2004) Making Personas More Powerful: Details to Drive strategic and Tactical Design.

Passini, R. (1994) Wayfinding in Architecture. Berlin, Germany, Van Nostrand.

Passini, R. (2000) Sign-posting information design. In R. Jacobson (Ed.), Information Design (pp. 83-98). Cambridge, MA: MIT Press.

Pirolli, P., & Card, S. (1999) Information foraging Psychological Review, Vol 106(4), Oct 1999, 643-675. doi: 10.1037/0033-295X.106.4.643.

Plumlee, M., & Ware, C. (2006) Zooming versus multiple window interfaces: Cognitive costs of visual comparisons. ACM Transactions on Computer-Human Interaction (TOCHI), 13(2):179–209, 2006.

Raubal, M. (n.d.) Spatial cognition — VO introduction to digital cartography. Lecture note at Institute of Geoinformatics, University of Munster, Germany.

Ritchie, J., & Lewis, J. (2003) Qualitative research practice: a guide for social science students and researchers. London: SAGE

Robertson, G. (1997) Navigation in information spaces. CHI '97 Workshop on Navigation in Virtual Worlds.

Roberts, J.C. (2000) Display models - ways to classify visual representations. *International Journal of Computer Integrated Design and Construction*, 2(4):241–250, December 2000.

Roberts, J.C. (2004) Exploratory visualization using bracketing. In M. F. Constabile, editor, *Advanced Visual Interfaces (AVI 2004)*, pages 188–192, Gallipoli, Italy, May 2004. ACM Press.

Roberts, J.C. (2007) State of the Art: Coordinated & Multiple Views in Exploratory Visualization. In *Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization (CMV 2007)* (pp. 61-71). IEEE. Doi: 10.1109/CMV.2007.20

Robertson, G., Cameron, K., Czerwinski, M., & Daniel, R. (2002) Animated Visualization of Multiple Intersecting Hierarchies, in *Journal of Information Visualization*, vol. 1, no. 1, pp. 50-65, March 2002

Ross, G., & Chalmers, M. (2003) A visual workspace for constructing hybrid multidimensional scaling and coordinating multiple views. *Information Visualization*, 2(4):247–257, December 2003

Ross, G., Morrison, A., & Chalmers, M. (2004) Coordinating views for data visualisation and algorithmic profiling. In *CMV '04: Proceedings of the Second*

International Conference on Coordinated & Multiple Views in Exploratory Visualization (CMV'04), pages 3–14, Washington, DC, USA, 2004. IEEE Computer Society.

Robertson, G.G., Card, S.K., & MacKinlay, J.D. (1993) Information visualization using 3D interactive animation. *Communications of the ACM*, 36(4):56–71, 1993.

Roth, M.R., Dennis, A., & Wixom, B. A. (2012) *Systems Analysis and Design*, 5th Edition, International Student Version ISBN : 978-1-118-37939-4592 pages June 2012, ©2012. Wiley.

Sacolick, I. (2009) [blogs.staricio.com/2009/04/business-responsibilities-working-with.html](http://blogs.staricio.com/2009/04/business-responsibilities-working-with.html)

Senko, M.E., Altman, E.B., Astrahan, M.M. & Fehder, P.L. (1973). *Data structures and accessing in data-base systems: II Information organization*.

Sheldon, R. (2005) *Beginning MySQL*. John Wiley & Sons. p. 156. ISBN 0-7645-7950-9.

Schulmeister, R. ( n.d.) *Components of navigation*, Technical Report  
*Navigation in Information Space*.

Shimabukuro, M.H., Flores, E.F., de Oliveira, M.C.F., & Levkowitz, H. (2004) Coordinated views to assist exploration of spatio-temporal data: A case study. In CMV '04: Proceedings of the Second International Conference on Coordinated & Multiple Views in Exploratory Visualization (CMV'04), pages 107–117, Washington, DC, USA. IEEE Computer Society.

Shiri, A. (2007) The use of metadata in visual Interfaces to digital libraries. School of Library & Information Studies, University of Alberta

Shneiderman, B. (1987) Designing the User Interface: Strategies for Effective Human-Computer Interaction Addison-Wesley Publ. Co., Reading, MA (1987),

Shneiderman, B. (1990), "Future directions for Human Computer Interaction", International Journal of Human-Computer Interaction, Vol 2(1), pp. 73-90.

Shneiderman, B. (1996) The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations Proc. 1996 IEEE Conference on Visual Languages (Boulder, CO, Sept.3-6,1996) 336-343.

Shneiderman, B. (1997) Designing Information-abundant websites: issues and recommendations. International Journal of Human-Computer studies. 47(1), 5-29.

Shneiderman, B. (2002) Leonardo's Laptop; Human needs and the new computing technologies. Cambridge MA: MIT Press Shneiderman, B. (2002). Creativity support tools. Communications of the ACM 45 (10), 116-120

Shneiderman, B., & Aris, A. (2006). Network visualization by semantic substrates. *IEEE Transactions on Visualization and computer graphics*, 12(5), 733-40. Doi: 10.1109/TVCG.2006.166

Shneiderman, B., Plaisant, C., Cohen, M., & Jacobs, S. (2009) *Designing the user interface: strategies for effective human-computer interaction*, 5th edition. Boston, Massachusetts: Addison-Wesley.

Simison, G.C. (2009) *Data and Reality*. Technic Publications. 3rd edition.

Sjolinder, M. (1998) *Spatial Cognition and Environmental Descriptions*. In *Exploring Navigation: Towards a Framework for Design and Evaluation of Navigation in Electronic Spaces*. Ed. Dahlbäck, N. SICS Technical Report T98:01, March 1998, ISSN: 1100-3154, ISRN: SICS-T-98/01-SE.

Smith, P. A. (1996) *Towards a practical measure of hypertext usability*, *Interacting with Computers* 8(4), 365–381.

Smith, P. R., & Sarfsty, R. (1993). *Creating a strategic plan for configuration management using Computer Aided Software Engineering (CASE) tools*. Paper for 1999 DOE/contractors and facilities CAD/CAE users group.

Solodovnik, I. (2011) "Metadata issues in Digital Libraries: key concepts and perspectives". JLIS. (University of Florence) 2 (2). doi:10.4403/jlis.it-4663. Retrieved 29/06/2013.

Solso, R. L. (1998) Cognitive Psychology (5th edn). Boston: Allyn and Bacon.

Soukup, T. (2002) Visual data mining : techniques and tools for data visualization

Spence, R. (1996) Visualisation really has nothing to do with computers, Proceedings Euro graphics UK, Pages:1-8

Spence, R. (1997) Navigation in Real and Virtual Worlds, Workshop on Personalised and Social Navigation in Information Space

Spence, R. (1997) The Use of Interactive Histogram Visualization in the interpretation of complex geochemical data sets, Proceedings Migration '97

Spence, R. (1999) A framework for navigation. International Journal of Human-Computer Studies, 51

Spence, R. (2001) Information visualization, Harlow, Addison-Wesley, ISBN:9780201596267

Spence, R. (2006) Information visualization: design for interaction, ISBN:9780132065504



Spence, R. (2007) *Information Visualization: Design for Interaction* (2nd edition), Prentice Hall, 2007, ISBN 0-13-206550-9 and mining. New York: Chichester

Staples, M. L., & Bieman, J. M. (1999) *3-D Visualization of Software Structure*. In *Advances in Computers* (Vol. 49, pp. 96–143): Academic Press, London.

Stasko, J. T., Brown, M. H., & Price, B. A. (1997) *Software Visualization*: MIT Press

Theus, M. (2002) *Interactive data visualization using mondrian*.  
*Journal of Statistical Software*, 7(11), 2002.

Thuring, M., Hannemann, J., & Haake, J. (1995) *Hypermedia and cognition: Designing for comprehension*, *Comm. Of the ACM* 38(1), 57–66.

Theng, Y. (1999) "Lostness" and Digital Libraries, *Proc. Digital Libraries*, Berkeley, CA.

Thomas, J. J., & Cook, K. (2005) *Illuminating the path: the R&D agenda for visual analytics*. IEEE.

Trepess, D. (2013) *Information Foraging*. Interaction Design Foundation, Denmark.  
Retrieved May, 16th 2014 [www.interaction-design.org](http://www.interaction-design.org).

Tufte, E. R. (1983) *The Visual Display of Quantitative Information*, Graphics Press, Cheshire, CT.

Tufte, E. R. (1991) *Envisioning Information* Graphics Press, Cheshire, CT.

Tweedie, L. (1997) Characterizing interactive externalisations. Proceedings of CHI'97, ACM Conference of Human Factors in Computing Systems, Atlanta, 375-382.

Upton, C., Faulhaber, T., Kamins, D., Schlegel, D., Laidlaw, D., Vroom, F., Gurwitz, R., & VanDam A. (1989) The application visualization system: A computational environment for scientific visualization. *IEEE Computer Graphics and Applications*, 9(4):30–42, 1989.

Van Dam, S., Foley, J., Feiner, S., & Hughes, J. (1995) *Computer Graphics, Principles and Practice*. Reading MA USA: Addison-Wesley p1174.

Van Dyke Parunak, H. (1989) Hypermedia Topologies and User Navigation, Proceedings of the Second Annual ACM Conference on Hypertext, 43-50, 183.

Vellucci, Sherry L. (1998) Metadata. *Annual Review of Information Science and Technology (ARIST)*, v33 p187-222 1998.

Walker, J. (1990) *Through the Looking Glass* In Laurel, B. (ed.) *The Art of Human-Computer Interface Design* Addison-Wesley, Cambridge, Ma

Warwick, C. (1997). *The British National Corpus. Digital Resources for the Humanities*, Conference St Anne's College, Oxford, 14 th - 17th September 1997

Ware, C. (2004). Information Visualization: perception for Design (2nd Edition). Morgan Kaufman.

Weaver, C. (2004) Multidimensional Visual Analysis Using Cross-Filtered Views. The GeoVista Centre and Department of geography. The pennsylvania State University

Weinschenk, S., & Barker, D.T. (2000) Designing Effective Speech Interfaces. Wiley.

Ware, C. (2000) Information Visualization: Perception for Design. San Diego, CA, USA: Academic Press, 2000.

Ware, C. (2008) Visual thinking for design. Morgan Kaufman.

West, M., & Fowler, J. (1999) Developing High Quality Data Models. The European Process Industries STEP Technical Liasion Executive (EPISTLE).

Wodke, C. (2003) Information Architecture: Blueprints for the web. Test (p.312). doi: 10.1093/intimm/dxs023

Wong, P., & Thomas, J. (2004) Visual Analytics. IEEE Computer Graphics and Applications, 24(5), 20–21.

WWW. [lidi.napier.ac.uk/c/groups/ Publications site/iidi](http://lidi.napier.ac.uk/c/groups/Publications%20site/iidi)

Yin, R. K.(2003) Case study research, design and methods, 3rd ed. Newbury Park:  
Sage. ISBN 0-7619-2553-8

# Appendices

## Appendix One:

### Task Sheets

#### Empirical work one task sheet: Interface A and B

#### Data Visualization and Navigation

##### Task sheet 1

Thank you for agreeing to participate in this study of data visualisation and navigation.

Please undertake as many of the following tasks as you can using the data visualisation of the data about the Institute of Informatics and Digital innovation (IIDi).

Identify and list two different things on the visualisation and write down the information they provide.	
How many publications has member X made?	
Locate two people in the centre for interaction design (CID) that have made joint publications with people from other centres.	
Which centre does member V belong?	
What subject is the centre for emergent computing (CEC) interested in?	

How many grants have been awarded to the Centre for Social Informatics (CSI)?	
Which individual has made the highest number of publications? Which centre does the person belong to?	
How many mutual publications has member W made in the last three years and what topic did they address?	
From the overview, list the five centres in the institute and identify the centre with the least number of members.	
Which centre has the highest number of publications?	

### **Empirical work two task sheet: Interface A**

Thank you for agreeing to participate in this study of data visualization and navigation.

Please undertake as many of the following tasks as you can using the visualization of the data about the Institute of Informatics and Digital Innovation (IIDI). However, you may choose to discontinue your participation in this study at any time.

#### **Interface A**

**URL:** <http://www.soc.napier.ac.uk/~cs22/test2/iidivislog.html>

<p><b><i>((Please enter task number 1 before you start))</i></b></p> <p>1. Identify two objects on the given interface and write down the information they provide.</p>	
<p><b><i>((Please enter task number 2 before you start))</i></b></p> <p>2. How many publications were made by CISS in 2010? What was the subject of the publication?</p>	
<p><b><i>((Please enter task number 3 before you start))</i></b></p> <p>3. Locate and list two centres that have made joint publications. Write the names of the people involved in this publication.</p>	
<p><b><i>((Please enter task number 4 before you start))</i></b></p> <p>4. Which centre has been awarded the highest number of grants? Name the member who received it?</p>	

<p><b><i>(Please enter task number 5 before you start )</i></b></p> <p>5. Identify the centre which focuses on social informatics.</p> <p>In which year did it make its highest number of publications?</p>	
<p><b><i>(Please enter task number 6 before you start)</i></b></p> <p>6. Locate person X. Which centre does the member belong to?</p> <p>How many publications did the member make in 2009?</p> <p>How many grants has the member received?</p>	

On completion of the first six tasks, you may attempt the tasks below.

<p><b><i>(Please enter task number7 before you start)</i></b></p> <p>7. From the overview, list the five centres in the institute.</p>	
--	--



<p><b><i>(Please enter task number 8 before you start)</i></b></p> <p>8. Who is the head of the CDCS?</p>	
<p><b><i>(Please enter task number 9 before you start)</i></b></p> <p>9. List four types of publications made by the institute.</p>	
<p><b><i>(Please enter task number 10 before you star)</i></b></p> <p>10. Name the centre with the least number of grants.</p>	

Date:

Time:

Group:

### **Empirical study two task Sheet: Interface B**

Thank you for agreeing to participate in this study of data visualization and navigation. Please undertake as many of the following tasks as you can using the visualization of the data about the Institute of Informatics and Digital Innovation (IIDi). However, you may choose to discontinue your participation in this study at anytime.

### **Interface B**

URL: <http://www.soc.napier.ac.uk/~cs22/test3/iidi/iidivislog3.html>

<p><b><i>(Please enter task number 1 before you start )</i></b></p> <p>1. Identify two objects on the given interface and write down the information they provide.</p>	
<p><b><i>(Please enter task number 2 before you start)</i></b></p> <p>2. How many publications were made by CID in 2010? What was the subject of the publication?</p>	
<p><b><i>(Please enter task number 3 before you start)</i></b></p> <p>3. Locate and list two centres that have made joint publications. Write the names of the people involved in this publication.</p>	
<p><b><i>(Please enter task number 4 before you start)</i></b></p>	

<p>4. Which centre has been awarded the highest number of grants?</p> <p>Name the member who received it?</p>	
<p><b><i>(Please enter task number 5 before you start)</i></b></p> <p>5. Identify the centre which focuses on networking.</p> <p>In which year did it make its highest number of publications?</p>	
<p><b><i>(Please enter task number 6 before you start)</i></b></p> <p>6. Locate member Y. Which centre does the member belong to?</p> <p>How many publications did the member make in 2009?</p> <p>How many grants has the member received?</p>	

On completion of the first six tasks, you may attempt the tasks below.

<p><b><i>(Please enter task number 7 before you start)</i></b></p>	
--	--

<p>7. From the overview, list the five centres in the institute.</p>	
<p><b><i>(Please enter task number 8 before you start)</i></b></p> <p>8. Who is the head of the CEC?</p>	
<p><b><i>(Please enter task number 9 before you start)</i></b></p> <p>9. List four types of publications made by the institute.</p>	
<p><b><i>(Please enter task number 10 before you start)</i></b></p> <p>10. Name the centre with the least number of grants.</p>	

Date:

Time:

Group:

**Empirical study three task sheet**

Thank you for agreeing to participate in this study of data visualization and navigation. Please undertake as many of the following tasks as you can using the Institute of Informatics and Digital Innovation (IIDi) data visualization.

Tasks	Response
1. Identify and list two objects on the IIDi visualization and write down the information they provide.	
2. Locate two people in the centre for interaction design (CID) that have made joint publications with people from other centres.	
3. Write down the name of the person who has received the highest number of grants in each centre.	
4. Find the member with the highest number of publications? How many grants has the person received?	
5. From the overview, list the five centres in the institute. Which centre has made the highest number of publications? How many grants has the centre been awarded?	

6. How many conference papers have been published by CDCS.	
7. a). Think of a task you would like to carry out and write it down. b). Perform the task.  c). Please state if you were able to perform the task. d). If your answer is No. Please state why?	

**Informed consent form for research participation**

**Edinburgh Napier University**

Faculty of Engineering, Computing and Creative Industries

**Informed Consent Form to Participate in Research**

**Title:** Data Visualization and Navigation: Navigating Coordinate Multiple Views of Visualization

**Student Researcher:** Igoniderigha Nseabasi E.

**Director of Studies:** Benyon David

**Second Supervisor:** Tom Mc Ewan

**Research:**

The research, 'Data Visualization and Navigation' focuses mainly on identifying novel methods to make navigation more effective while visualising large sets of data on information spaces.

**Purpose:**

I am a computing science research student at Edinburgh Napier University. I am carrying out investigations so as to unearth novel methods that will improve navigation in information spaces to allow easier and more effective visualization of large data sets.

**Procedure:**

During this study, you will be asked to carry out some tasks and write down the steps you took while performing the task. In addition, you will also be asked to write down any difficulties you encountered during the exercise and suggestions on how to make the task easier. A form has been provided for you to enter this information. You may ask any questions about the tasks. The total testing time should be about 30 minutes. Thank you for your time.

**Risks:**

There are no risks directly related to participating in this research.

**Participant's Agreement:**

I am aware that my participation in this research tasks is voluntary. If for any reason, at any time, I wish to stop the tasks, I may do so without having to give an explanation.

I understand the purpose of this research.

I agree to participate in the tasks for this research.

**Signature:** \_\_\_\_\_

**Date:** \_\_\_\_\_

**Sample of log file analysis (Task 1)**

[TS] – 9sec, [SM] – 10 sec, [SLM] – 62sec, [FPB] – 10sec, [MN] – 9 sec

[SLM] – 59sec

Time taken: 159 sec

Table 5: Navigation Pattern illustration for Interface A, Tutorial 1: n = 18

[TS] – 26sec, [FPB] – 28 secs, [MN] – 16 secs, [SN] – 3sec, [MM] – 6 secs,

[SLM] – 11 sec, [TS] – 8 sec, [MM] – 19 sec, [RS] – 35sec and [TS] – 1 sec.

Time taken: 162 sec

Table 6: Navigation Pattern illustration for Interface A, Tutorial 2: n = 37

[RS] – 4 sec, [TS] – 9 sec, [MN] – 1 sec, [SN] – 5 sec, [MN] – 7 sec, [SN] – 2sec, [MN] – 7 sec, [SN] – 2 sec, [MN] – 12 sec, [SN] – 3 sec, [MN] – 1 sec, [SN] – 11 sec, [MN] – 2 sec, [SN] – 8 sec, [MN] – 7 sec, [MM] – 2 sec, [SLM] – 17 sec, [MN] – 1 sec, [MM] – 3 sec, [SLM] – 3 sec, [MM] – 45 sec

Time taken: 72 sec

Table 7: Navigation Pattern illustration for Interface B, Tutorial 3: n = 17

[TS] – 12 sec, [MN] 1 sec, [SN] – 7 sec, [MN] – 13 sec, [SN] – 1 sec, [MN] – 9 sec, [MN] – 5 sec, [SN] – 3 sec, [MN] – 51 sec, [MM] – 5sec, [MN] – 7 sec, [MM] – 1 sec, [SLM] – 3 sec, [SM] – 5 sec.

Time taken: 123sec



## Appendix Two:

### Redesign of SIMD CMV by Alistair Rae

In 2012, a lecturer, Alasdair Rae, found the SIMD map (figure A) of most deprived areas in Scotland 'frustrating to use'. He created an alternate map (figure B) to illustrate the same data in figure A.



Figure A: Map of most deprived areas in Scotland in 2012

Source: Adapted from [www.holyrood.com](http://www.holyrood.com)

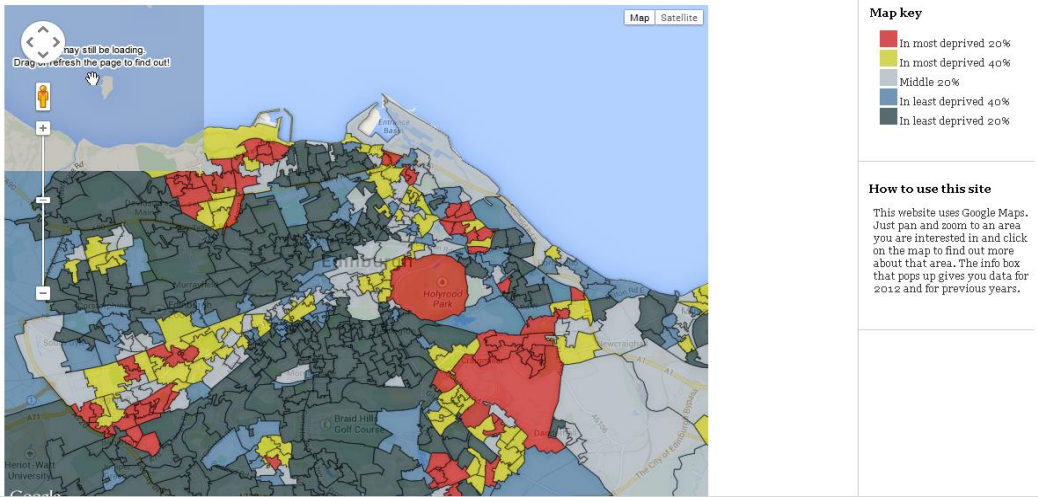


Figure B: Map of most deprived areas in Scotland 2012

Source: Adapted from Alasdair Rae (2012) [www.holyrood.com](http://www.holyrood.com)

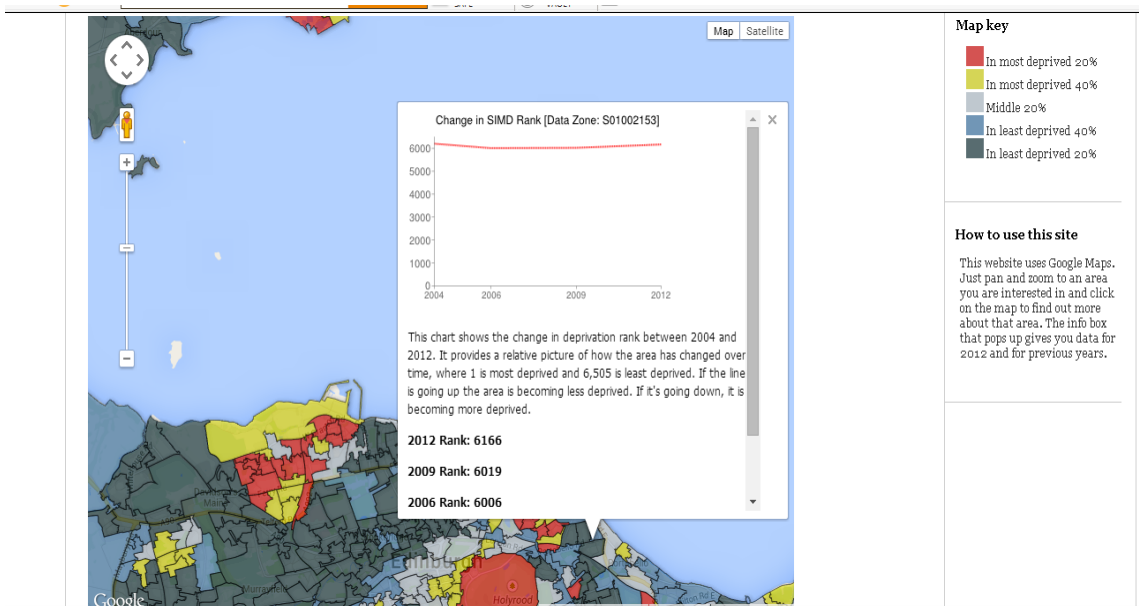


Figure C: Map of most deprived areas in Scotland showing metadata in 2012

Source: Adapted from Alasdair Rae in 2012 [www.holyrood.com](http://www.holyrood.com)

Alasdair Rae developed an alternate representation of most deprived areas in Scotland 2012 by superimposing figure A onto Google maps to enable user's mouseover the various datazones to obtain information. According to him, his aim was

to develop an interactive mapping website to allow people interact with the data easily. It allows users view the data from three perspectives; satellite, map or terrain. Users can zoom in and out of regions and when they click on a datazone, a graph pops up giving details about how the area has performed as regards its deprivation ranking since 2004. He claims that the SIMD interactive mapping has good features but is 'burdensome', for instance users spend a long time aggregating data.