


Spring 12-2011

A Method for Representing Contextualized Information (MeRCI) to Improve Situational Awareness Among Electronic Message Brokering System Dashboard Users

Arunkumar Srinivasan

University of Texas Health Science Center, Houston-Texas, USA.

Follow this and additional works at: http://digitalcommons.library.tmc.edu/uthshis_dissertations

 Part of the [Bioinformatics Commons](#), and the [Medicine and Health Sciences Commons](#)

Recommended Citation

Srinivasan, Arunkumar, "A Method for Representing Contextualized Information (MeRCI) to Improve Situational Awareness Among Electronic Message Brokering System Dashboard Users" (2011). *UT SBMI Dissertations (Open Access)*. Paper 29.

This is brought to you for free and open access by the School of Biomedical Informatics at DigitalCommons@The Texas Medical Center. It has been accepted for inclusion in UT SBMI Dissertations (Open Access) by an authorized administrator of DigitalCommons@The Texas Medical Center. For more information, please contact laurel.sanders@library.tmc.edu.

Dissertation

A METHOD FOR REPRESENTING CONTEXTUALIZED INFORMATION (MeRCI)
TO IMPROVE SITUATIONAL AWARENESS AMONG ELECTRONIC MESSAGE
BROKERING SYSTEM DASHBOARD USERS
by

Arunkumar Srinivasan, MS

December, 2011

APPROVED:

Jiajie Zhang, PhD

Sriram Iyengar, PhD

Kim. Dunn, MD PhD

Jack W. Smith, MD, PhD

John Patrick Abellera, MPH

A METHOD FOR REPRESENTING CONTEXTUALIZED INFORMATION (MeRCI)
TO IMPROVE SITUATIONAL AWARENESS AMONG ELECTRONIC MESSAGE
BROKERING SYSTEM DASHBOARD USERS

A
DISSERTATION
Presented to the Faculty of
The University of Texas
School of Biomedical Informatics
at Houston
in Partial Fulfillment
of the Requirements

for the Degree of

Doctor of Philosophy

By

Arunkumar Srinivasan, MS

Committee Members:

Jiajie Zhang, PhD¹
Madurai Sriram Iyengar, PhD¹
Kim Dunn, MD PhD¹
Jack W. Smith, MD, PhD¹
John P. Abellera, MPH²

¹University of Texas Health Science Center, School of Biomedical Informatics, Houston, TX

²Centers for Disease Control and Prevention, Atlanta, GA

Copyright © 2011

by

Arunkumar Srinivasan

DEDICATION

This work is dedicated to my Parents, who believed in and unconditionally supported my endeavors in life and to my wife who stood by and encouraged me in this endeavor.

I feel very fortunate to have been a student in UT-HSC Health Informatics program. The excellent faculty, students, and staff have created an environment that promotes absorbingly interesting discussions, exciting research projects, and supportive friendships. I am also very thankful to the school for allowing me to pursue the fellowship program at CDC that has helped me obtain hands-on training and research insight into the field of public health

I thank Dr. Zhang for being a supportive research and academic advisor. His enthusiasm, insightful comments, and sympathetic ear helped me through the rough spots in my graduate training. I thank Dr. Sriram, Dr. Dunn and Dr. Smith for their meticulous scrutiny of my dissertation and their insightful comments. I am grateful to Dr. Mirhaji for providing me with guidance and research advice during the important stages of my grad school work.

I am also very grateful to the Staff and Colleagues at the Division of Notifiable Disease Surveillance at The Centers for Disease Control and Prevention. I would particularly like to thank Mr. Scott Danos, Mr. John Abellera, Dr. Scott McNabb, Ms. Jennifer Ward, Mr. Aaron Aranas, Mr. Enrique Nieves and Dr. Leslie Lenert for supporting this academic endeavor. Thanks to my colleagues at State, Local and County Departments of Public Health, public health laboratories for participating in my research and also motivating me to finish my PhD quickly.

A METHOD FOR REPRESENTING CONTEXTUALIZED INFORMATION (MeRCI)
TO IMPROVE SITUATIONAL AWARENESS AMONG ELECTRONIC MESSAGE
BROKERING SYSTEM DASHBOARD USERS

Arunkumar Srinivasan, MS

The University of Texas School of Biomedical Informatics, 2011

Dissertation Advisors: Jiajie Zhang, Sriram Iyengar, Kim Dunn, Jack W. Smith, John P. Abellera

Electronic health information brokering systems are of interest to public health informatics because they emphasize how data can be effectively shared and utilized across healthcare institutions and among providers so as to improve the quality of care, increase efficiency, and reduce costs (Lumpkin, 2002). In the domain of public health (PH) specifically, where complete and timely reporting of data is critical for all epidemiological and disease surveillance activities (Langmuir, 1976), it is imperative to ensure proper functioning of the electronic information exchange infrastructure. Receiving multiple types of data, in various formats from numerous sources, and triaging them to the appropriate surveillance system is no easy task for a department of health, whether at state, local or federal level (Magnuson, 2005).

The administrators of the electronic message brokering system, and the coordinators of surveillance systems in each public health jurisdiction, are responsible for ensuring that the data is received, archived, validated and triaged appropriately in a timely and complete fashion. This requires continuous monitoring of trends in messaging and system performance and active responses to aberrations. To achieve this, administrators depend heavily on dashboards to provide awareness of exchange system status and its reporting at any point of time. Unfortunately, current dashboards do not offer the context or cognitive support needed for interpreting the information presented. As research has demonstrated in other domains, in order

to make sense of the data and react, dashboard users are required to draw upon domain knowledge, higher level association between domains, operational rules, organizational missions, personal objectives, tasks at hand, priorities, past experiences, historic events, recent events, psychosocial and political constructs, and more (Resnick, 2005; Mirhaji, Srinivasan, Casscells, & Arafat, 2004). The burden of 'interpretation' always falls on the cognitive system of the human operator, which is prone to error and malfunctioning when risk and emergency overwhelm psychological factors (Parsa, Richesson, Smith, Zhang, & Srinivasan, 2004; Parsa, Zhang, Smith, Majid, Casscells, & Lillibridge, 2003). On the basis of the surveillance literature it can be seen that meaningful and holistic interpretation of data requires the generation of higher-level explanations based on knowledge and expertise from numerous principles (Parsa, Richesson, & Srinivasan, 2004; Parsa, Richesson, Smith, Zhang, & Srinivasan, 2004), while context is essential to illustrate the 'big picture' view of dynamic and complex problems (Parsa, Zhang, Smith, Majid, Casscells, & Lillibridge, 2003). These reservations imply that the process for building health information dashboards should consider not only user functions, tasks and goals but also the user's situational awareness (SA) requirements. This vision adds a new layer to information representation that needs to be accounted for when conceptualizing the implementation of health information dashboards. A review of the literature reveals a lack of methods to design for situational awareness in dashboard systems in complex domains (Resnick, 2005; Li, 2007).

This research introduces a new method to present contextualized information that can improve user SA. I present the design rationale, method, and results of an evaluation study that measures the situational awareness generated by adopting this new context-driven representation model.

TABLE OF CONTENTS

List of Tables	10
List of Figures	13
Preface.....	15
Research Contribution	17
Outline of this dissertation	17
Chapter 1: Public Health Data Exchange.....	19
Need for Public Health Data Exchange Systems	19
Complexity in Working with Public Health Dashboard	21
Defining Dashboards	28
Defining Context.....	31
Characteristics of Previous Approaches to Dashboard Design.....	32
Research Aims	33
Chapter Summary	34
Chapter 2: Prior Art	36
Dashboard Design Principles	37
Dashboard Design Process.....	39
Technology Centered Design Approach for Dashboards	42
User Centered Design (UCD) Approach for Dashboards.....	43
Ecological Interface Design (EID) Approach for Dashboards	46
Summary of Design Approaches	50
Situational Awareness and Mental Models.....	52
Advantages of Situational Awareness	54
Situational Awareness in Public Health.....	55
Challenges in Achieving Situational Awareness	57
Measuring Situational Awareness.....	60
Methods of Measuring Situational Awareness	62
Situational Awareness Global Assessment Technique (SAGAT)	62
Issues of Validity and Reliability of SAGAT	63
Chapter Summary	65
Chapter 3: Method for representing contextualized information.....	67
Conceptual Framework of Dashboard Design	67
Design Process	68
Goal Directed Task Analysis (GDTA)	69
Initial Interview.....	71
Preliminary Goal Structure	72
Final Goal Structure.....	73
GDTA Summary	75
Concept Maps for Knowledge Elicitation and Representation.....	78
Knowledge Acquisition Process	79

Knowledge Representation	81
Concept Mapping Activity Summary	84
Dashboard Design Principles	85
Design Principles – from Theory to Practice	86
Summary of the Design	94
Evaluation Study	94
Evaluation Instrument	95
Chapter Summary	97
Chapter 4- Results	98
Overview & Objectives	98
Comparison Systems	98
Pilot Study	99
Final Study	100
Method	101
Study Subjects	101
Procedure	102
Analysis of Results	104
Response Type Related Analysis	104
Impact of Age, Experience and Edu Background to Perception Questions	106
Impact of Age, Experience and Edu Background to Interpretation Questions	109
Impact of Age, Experience and Edu Background to Forecasting Questions	112
Relationship between SA level to Perception & Interpretation Questions	115
Relationship between SA level to Interpretation & Forecasting Questions	117
Response Time Related Analysis	119
Response Time Analysis by Type of Response for Level 1 Questions	121
Response Time Analysis by Types of Response for Level 2 Questions	123
Response Time Analysis by Types of Response for Level 3 Questions	124
Impact of Age, Experience and Edu on Response Time for Level 1 Questions	126
Impact of Age, Experience and Edu on Response Time for Level 2 Questions	128
Impact of Age, Experience and Edu on Response Time for Level 3 Questions	131
Confidence Analysis	137
Confidence Related Analysis for All SA Levels	137
Confidence Related Analysis by response type for SA Level 1	140
Confidence Related Analysis by response type for SA Level 2	142
Confidence Related Analysis by response type for SA Level 3	143
Impact of Age, Experience and Edu on Confidence for SA Level 1 Questions	145
Impact of Age, Experience and Edu on Confidence for SA Level 2 Questions	148
Impact of Age, Experience and Edu on Confidence for SA Level 3 Questions	150
Confidence analysis by response type between SA Levels 1 and 2	153
Confidence analysis by response type between SA Levels 2 and 3	155
Chapter Summary	156
Chapter 5: Discussion	158
Introduction	158
MeRCI – Evaluation for Situational Awareness	158
Guidelines for Building towards Situational Awareness	166
Guidelines for choosing a Rule based System	173

Chapter Summary	177
Chapter 6: Summary & Conclusion.....	178
Contextualized Information Representation for Situational Awareness.....	178
Contributions.....	179
Limitations and Future Work.....	181
Concluding Remarks.....	183
APPENDIX A – Goal Directed Task Analysis.....	185
APPENDIX B – Concept Map Formal Representation.....	201
APPENDIX C – Evaluation Study Instruments.....	219
APPENDIX D – Dashboard Prototypes	221
Bibliography	227

LIST OF TABLES

Table 1 Measurement strategies for situational awareness.....	62
Table 2 Preliminary Goal Hierarchy.....	73
Table 3 Scenarios Identified using Concept Mapping.....	83
Table 4 Drools Representation of a Message Trend Scenario.....	83
Table 5: Comparison of Information available among system components.....	88
Table 6: Scenarios for Evaluation Study.....	97
Table 7 Demographic characteristics of the sample.....	102
Table 8 Cross-tabulation of correct and incorrect responses to perception questions.....	105
Table 9 Cross-tabulation of correct and incorrect responses to interpretation questions.....	105
Table 10 Cross-tabulation of correct and incorrect responses to forecasting questions.....	106
Table 11 Cross-tabulation of levels of correct responses to perception questions for the two age groups.....	107
Table 12 Cross-tabulation of levels of correct responses to perception questions for the four occupational experience levels.....	108
Table 13 Cross-tabulation of levels of correct responses to perception questions for the two educational background groups.....	109
Table 14 Cross-tabulation of levels of correct responses to interpretation questions for the two age groups.....	110
Table 15 Cross-tabulation of levels of correct responses to interpretation questions for the four occupational experience levels.....	111
Table 16 Cross-tabulation of levels of correct responses to interpretation questions for the two educational background groups.....	112
Table 17 Cross-tabulation of levels of correct responses to forecasting questions for the two age groups.....	113
Table 18 Cross-tabulation of levels of correct responses to forecasting questions for the four occupational experience levels.....	114
Table 19 Cross-tabulation of levels of correct responses to forecasting questions for the two educational background groups.....	115
Table 20 Correlations between response correctness scores on perception and interpretation questions for the two interfaces.....	116
Table 21 Contingency Table for Both Correct vs. One or More Incorrect Responses to the Perception and Interpretation Questions by Interface Type.....	117
Table 22 Correlations between response correctness scores on perception and interpretation questions for the two interfaces.....	118
Table 23 Contingency Table for Both Correct vs. One or More Incorrect Responses to the Perception and Interpretation Questions by Interface Type.....	119
Table 24 Results of t-test of difference between the two interfaces in their mean response times to the perception questions.....	120
Table 25 Results of t-test of difference between the two interfaces in their mean response times to the interpretation questions.....	120
Table 26 Results of t-test of difference between the two interfaces in their mean response times to the forecasting questions.....	121

Table 27 Results of ANOVA of response time for perception questions by correctness of response and subject for each interface	122
Table 28 Results of ANOVA of response time for interpretation questions by correctness of response and subject for each interface	123
Table 29 Results of ANOVA of response time for forecasting questions by correctness of response and subject for each interface	125
Table 30 Results of t-test of difference between under 40 and 40 or older age groups in mean response times for perception questions	126
Table 31 Results of analysis of variance of perception question response times by occupational experience levels for the new and conventional interfaces.....	127
Table 32 Results of t-test of difference between IT/Informatics vs. other educational background groups in mean response times for perception questions.....	128
Table 33 Results of t-test of difference between under 40 and 40 or older age groups in mean response times for interpretation questions.....	129
Table 34 Results of t-test of difference between under 40 and 40 or older age groups in mean response times for interpretation questions.....	130
Table 35 Results of t-test of difference between IT/Informatics vs. other educational background groups in mean response times for interpretation questions.....	131
Table 36 Results of t-test of difference between under 40 and 40 or older age groups in mean response times for forecasting questions	132
Table 37 Results of analysis of variance of forecasting question response times by occupational experience levels for the new and conventional interfaces.....	133
Table 38 Results of t-test of difference between IT/Informatics vs. other educational background groups in mean response times for forecasting n questions.....	134
Table 39 Results of t-tests of the difference in interpretation question response times between correct and incorrect perception question responders under the two interface conditions	135
Table 40 Results of t-tests of the difference in forecasting question response times between correct and incorrect interpretation question responders under the two interface conditions	136
Table 41 Results of ANOVA of confidence ratings of responses to perception questions by interface type and user	138
Table 42 Results of ANOVA of confidence ratings of responses to interpretation questions by interface type and user	139
Table 43 Results of ANOVA of confidence ratings of responses to forecasting questions by interface type and user	140
Table 44 Results of ANOVA of confidence ratings of responses to perception questions by correctness of response and subject for each interface	141
Table 45 Results of ANOVA of confidence ratings of responses to interpretation questions by correctness of response and subject for each interface	142
Table 46 Results of ANOVA of confidence ratings of responses to forecasting questions by correctness of response and subject for each interface	144
Table 47 Results of t-test of difference between under 40 and 40 or older age groups in mean confidence ratings of responses to perception questions	145
Table 48 Results of analysis of variance of perception question confidence ratings by occupational experience levels for the new and conventional interfaces	146
Table 49 Results of t-test of difference between IT/Informatics vs. other educational background groups in mean confidence ratings of responses to perception questions	147

Table 50 Results of t-test of difference between under 40 and 40 or older age groups in mean confidence ratings of responses to interpretation questions	148
Table 51 Results of analysis of variance of interpretation question confidence ratings by occupational experience levels for the new and conventional interfaces	149
Table 52 Results of t-test of difference between IT/Informatics vs. other educational background groups in mean confidence ratings of responses to interpretation questions.....	150
Table 53 Results of t-test of difference between under 40 and 40 or older age groups in mean confidence ratings of responses to forecasting questions	151
Table 54 Results of analysis of variance of forecasting question confidence ratings by occupational experience levels for the new and conventional interfaces	152
Table 55 Results of t-test of difference between IT/Informatics vs. other educational background groups in mean confidence ratings of responses to forecasting questions.....	153
Table 56 Results of t-tests of the difference in mean confidence ratings for interpretation questions between correct and incorrect perception question responders under the two interface conditions.....	154
Table 57 Results of t-tests of the difference in forecasting question response times between correct and incorrect interpretation question responders under the two interface conditions	155
Table 58: Comparison of Rules Engine Features	176

LIST OF FIGURES

Figure 1 Data Exchange between health care partners and systems.....	21
Figure 2: NEDSS ELR Message Volume.....	22
Figure 3: ELR Trend from different sources indicating a pattern	23
Figure 4. Disciplines required for Public Health Preparedness	24
Figure 5: Patient Traffic in Houston Hospital ER	25
Figure 6: State Surveillance Dashboard with ELR component distributed	26
Figure 7 Guidelines to specify what gets into the dashboard (Few, 2006).....	38
Figure 8: The unmodified "waterfall model" as defined by Royce	40
Figure 9 Spiral Design Process- Example of a Concurrent Engineering Model	41
Figure 10 Interface design for complex systems using EID, adapted from Vicente, 1992.	49
Figure 11 Endsley's Model of Situation Awareness.....	54
Figure 12 SAGAT Queries for Public Health Surveillance.....	64
Figure 13 SA Drives Decision Making and Performance	68
Figure 14 Goal - Decision - SA Requirements Hierarchy Representation	70
Figure 15: Preliminary Goal Hierarchy	72
Figure 16 GDTA Summary Status.....	75
Figure 17: Final Goal Hierarchy	76
Figure 18: Final GDTA Hierarchy.....	77
Figure 19 Concept Map Interview Notes.....	79
Figure 20 Parking Lot of Lab Reporting Volume.....	80
Figure 21 Concept Map for Lab Reporting Volume Trend	81
Figure 22 Parking Lot of Message Trend Concepts	82
Figure 23: System and Source Status in the new Interface.....	90
Figure 24: Breadcrumbs to allow awareness	91
Figure 25: Windowed environment	91
Figure 26: Annotated Graphs providing historical information	92
Figure 27: Annotation Identified using Drool Rule.....	92
Figure 28: Web-Based Survey Instrument.....	96
Figure 29 User SA levels measured by response type when using both the interfaces	159
Figure 30 User Response Time in Seconds for all SA Levels.....	161
Figure 31 Comparison of Response Time for Correct and Incorrect Responses for Perception questions	161
Figure 32 Comparison of Response Time for Correct and Incorrect Responses for Interpretation questions	162
Figure 33 Comparison of Response Time for Correct and Incorrect Responses for forecasting questions	163
Figure 34 User Confidence Level when responding to different SA levels	163
Figure 35 Confidence Rating for Correct and Incorrect responses to Perception Questions	164
Figure 36 Confidence Rating for Correct and Incorrect responses to Interpretation Questions.	165
Figure 37 Confidence Rating for Correct and Incorrect Forecasting Questions	165
Figure 38: GDTA - Overall Goal.....	185
Figure 39: Monitoring System Operation.....	186
Figure 40: Assess System Status.....	187
Figure 41: Assess Comm pt & Route Status.....	188

Figure 42: Assess System Performance	189
Figure 43: Determine Cause of System Failures	190
Figure 44: Evaluate and Execute Response Plan	191
Figure 45: Assess Impact of Failures	192
Figure 46: Identify Response Options	193
Figure 47: Execute Response Options	194
Figure 48: Monitor Reporting Volume and Trends	195
Figure 49: Assess Reporting Trends by Source	196
Figure 50: Assess Reporting Trends by Source (continued)	197
Figure 51: Assess Reporting Trends by Diseases	198
Figure 52: Assess Reporting Trends by Diseases (continued).....	199
Figure 53: Monitor Data Quality	200
Figure 54: Traditional (old) MSS Dashboard integrated with BIRT Reporting.....	221
Figure 55: Orion Rhapsody System Dashboard.....	222
Figure 56: Comm Point Status	223
Figure 57: Rhapsody Route Monitor	224
Figure 58 New Prototype Main Page with Labcorp Trend.....	225
Figure 59: MSS System Window	225
Figure 60: Labcorp Reporting Window	226
Figure 61: Labcorp TB Reporting Window	226

PREFACE

Access to electronic data from traditional sources (e.g. Hospitals, Laboratories) and non-traditional sources (e.g. school absenteeism records, pharmacy sales, news feeds etc.) is critical for effective surveillance practice (Parsa, Richesson, & Srinivasan, 2004). Proper functioning of the message brokering system is critical for collecting, processing and triaging data for all surveillance and epidemiological activities. Dashboards are widely used by system administrators to monitor the activities and performance of the brokering systems (Srinivasan A, 2009). Current dashboards designed on the basis of traditional approaches do not provide a context to interpret reporting trends and events; instead they rely on the limited cognitive resources of expert users to characterize these trends and signals (Resnick, 2005). Understanding aberrations in reporting trends and following them up with effective response action depends on situational awareness (Kunapareddy, Mirhaji, Michea, Casscells, & Zhang, 2005). Studies in similar complex environments have shown that, in order to achieve awareness while using information systems, users should alternate between goal-directed and data-driven processing of information (Endsley, Bolte, & Jones, 2003). While goal-directed processing helps users to determine which elements in the environment they should pay attention to, data-driven processing presents information based on its perpetual characteristics (e.g. severity).

In a complex environment, where information resides within multiple systems and relevant information is not readily available for triggering the appropriate goal and tasks, dashboard designers face immense challenges in system implementation. Previous studies have shown that providing context can help to identify data relevant to the user's goal and tasks. In the domain of public health, the concept of contextualized information representation has not yet been applied and evaluated to show whether it can improve user's understanding and situation

awareness (3). This research work proposes to *develop a method* to create contextualized information representations for improving the user's situation awareness during the signal characterization task, and to *evaluate its effect* on the user's ability to perceive, interpret and project the data generated by public health systems.

Specifically, in this research work, I applied two Cognitive Task Analysis (CTA) approaches. A Goal Directed Task Analysis is first conducted so as to understand the various user goals, tasks and needs for SA information. A Context Map is then created to understand the operational and domain knowledge of the work domain. I merged the goals and domain knowledge in real time to create a hybrid representation that will provide more contextual data relevant to the user's goal and to enhance awareness. I relied upon information representation and cognitive theories to construct the tailored information representation. In order to evaluate the impact objectively, I evaluated the situational awareness of the system's user by employing the Situational Awareness Global Assessment Technique (SAGAT). The SAGAT instrument is customized following task analysis that measures the users' understanding and interpretation of the representation by probing cognitive constructs, such as perception of information, task-related interpretation, forecasting or explaining the near future or immediate past.

Better understanding of the contextualized information representation in public health systems will enable the construction of a model for design and evaluation of information representation in health information systems. The study has also resulted in identifying some guidelines for developing future systems for SA.

Research Contribution

- Identification of a method for building user-centered health information dashboards in complex real time systems for Situational Awareness
- Document design principles for user-centered health information dashboards for better situational awareness
- Introduction to the public health field of a validated method to investigate the impact of HI dashboards by objective measurement of awareness from system users.

Outline of this dissertation

Chapter 1 overviews the background, and outlines the domain problems, the environment, and the context of this dissertation. In this section basic principles of dashboard design and its significance are discussed in the light of current problems in the public health domain. Existing frameworks for dashboard design are introduced and major challenges of design, conceptualization, and implementation of robust human-centered dashboards are discussed. I highlight some of the core criteria that are required for measuring the impact of the system interface.

Chapter 2 reviews the prior art and describes the design and conceptualization of information dashboards. A comparative discussion of the pros and cons and design implications of each system is provided. This chapter concludes with a gap analysis that set the stage for further research and development in this area and rationalizes and motivates this work.

Chapter 3 formulates the problem from the author's perspective, provides the motivation, rationale and criteria that informed the conceptualization of the MeRCI system and the methods used to implement it. This chapter continues with an in-depth discussion of the system design,

and its components. At the end, there is a brief review of the challenges facing the evaluation of the health information system, followed by a detailed explanation of the evaluation methods used to assess its validity and reliability.

Chapter 4 presents the results of a comprehensive and methodological evaluation described in Chapter 3.

Chapter 5 is devoted to the in-depth analysis of the MeRCI design and its conceptualization. The discussions are focused on the design rationale and outcomes of the evaluation in light of the desiderata put forward in Chapter 1 for the next generation information representation, the gap analysis provided in Chapter 2, and the motivations introduced in Chapter 3. I have also documented the key design principles that were identified during the research study.

Chapter 6 concludes the dissertation, recapitulates its main points, and highlights the contributions and the significance of the MeRCI design to the field of health information sciences. Plans for the improvement of the system to address its known shortcomings are discussed, and future directions for research and development in the field are highlighted.

Each chapter ends with a summary of its content recapitulating the main points and concepts introduced.

CHAPTER 1: PUBLIC HEALTH DATA EXCHANGE

In this chapter, I will discuss the public health need for data exchange and the current challenges in monitoring and responding to issues in data exchange, visit the concept of the dashboard and dashboard design principles, and introduce some current data exchange system dashboards and the problems that users face in using these interfaces. I will also present the need for user-centered design (UCD) and discuss some of the common mistakes that people make when designing for UCD. Finally I will summarize the UCD approach and the methods and evaluation criteria adopted in this research.

Need for Public Health Data Exchange Systems

Public health practice is built on a distinctive science basis of epidemiology and biostatistics that facilitates the analysis of large sets of data for describing, understanding and reacting to health problems (Lumpkin, 2002). The advent of the computer and the development of information systems have increased the effectiveness of public health practice by delivering data in a timely and complete fashion for analysis (Ball, 2002). For public health agencies developing integrated health information systems, new risks, as well as benefits, are emerging rapidly on the horizon (Arzt, 2007). The ways in which public health information is increasingly exchanged among health-care providers, hospitals, government, insurers and families demand a closer look at the networked information environment (Srinivasan A. C., 2008). Information is one commodity that gains value the more it is used, and public health stands to benefit from a landscape of increasing opportunities for exchanging information among more sources and users.

The National Health Info Network initiative by the Office of National Coordinator (ONC) for Health is geared to establishing a standardized exchange of health information across

the whole health care entity, overlapping with the Public Health focus of exchange between PH entities, a commonly-known reference in PHIN. For the past few years public health has channelized its effort through national and regional efforts so as to effectively procure data. The 2010 report from Trust for America's Health indicate that, in FY 2010, \$13 billion was available via Cooperative Grants (Epi Lab Capacity (ELC), Affordability Care Act (ACA), Bioterrorism Funds (BT grants), and ARRA HITECH grants to improve electronic exchanges between public health stakeholders (TFAH, 2010).

With the current national push toward electronic medical records (EMR), clinical systems will increasingly need to comply with the Health Level 7 Electronic Health Record standard (HL7 EHR), and, to stay viable in the marketplace, will need to comply with minimum functional standards and be independently certified as compliant. Many different solutions—large and small—are available to provider practices today, and these products are likely to be consolidated as standards compliance becomes more important (HHS, 2010).

Two philosophically distinct approaches are adopted by data exchange partners to share data (**Error! Not a valid bookmark self-reference.**Figure 1). An interface engine is a software program designed to simplify the creation and management of interfaces between separate applications and systems within and between organizations (Mclead, 2004). Interface engines undertake messaging between systems, and normally manage any mapping, translation and data modification necessary to ensure the effective exchange of data among and within the organization. Without a healthcare integration engine, hospital, lab and public health administrators will have to operate in cumbersome, manual and time-intensive IT environments to move large volumes of health data (Srinivasan, Danos, McNabb, & Rhodes, 2008). Non-standardized message content, disparate message structure, reporting process irregularities,

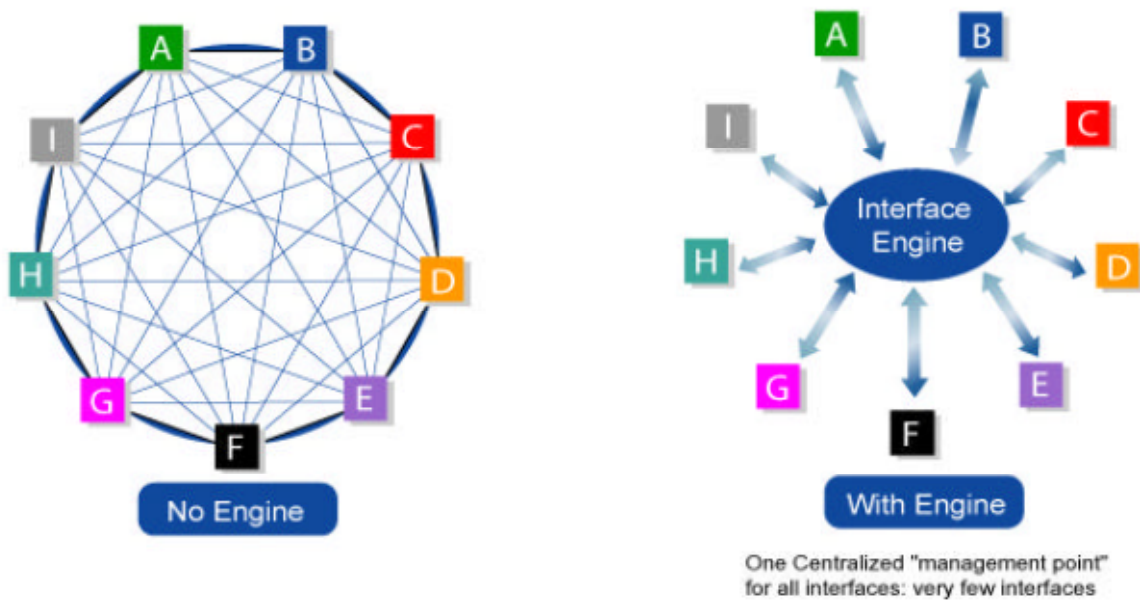


Figure 1 Data Exchange between health care partners and systems

changing reporting requirements, and disparate reporting protocols, all play a critical role in the health care data exchange space (Srinivasan & Abellera, 2010), particularly in public health where delivering a complete, quality message in a timely fashion will require continuous monitoring. The responsibility for data quality lies with the public health informaticists who rely on system dashboards to provide awareness. However current dashboards do not offer them enough depth in representation to understand the trends but rely on their knowledge for integrating data and providing interpretations (Srinivasan, Abellera, Danos, & McNabb, 2007).

Complexity in Working with Public Health Dashboard

The PH Informaticist operates in a world where disciplines ranging from social sciences (e.g., organizational theory, management and political science) to engineering (e.g., information science, computer science) and health sciences (e.g., public health epidemiology, infectious disease, behavioral science) are applied. Figure 4. Disciplines required for Public Health Preparedness represents the core competencies and knowledge domains involved in interpretation of data and information relevant to public health situation awareness

(Kunapareddy, Mirhaji, Michea, Casscells, & Zhang, 2005). The knowledge to interpret the PH data is distributed between many disparate domains, and the data to be interpreted is also distributed over many domains. Hence our understanding and interpretation of the data will differ as the context changes within which inferences are made (Parsa, Richesson, & Srinivasan, 2004).

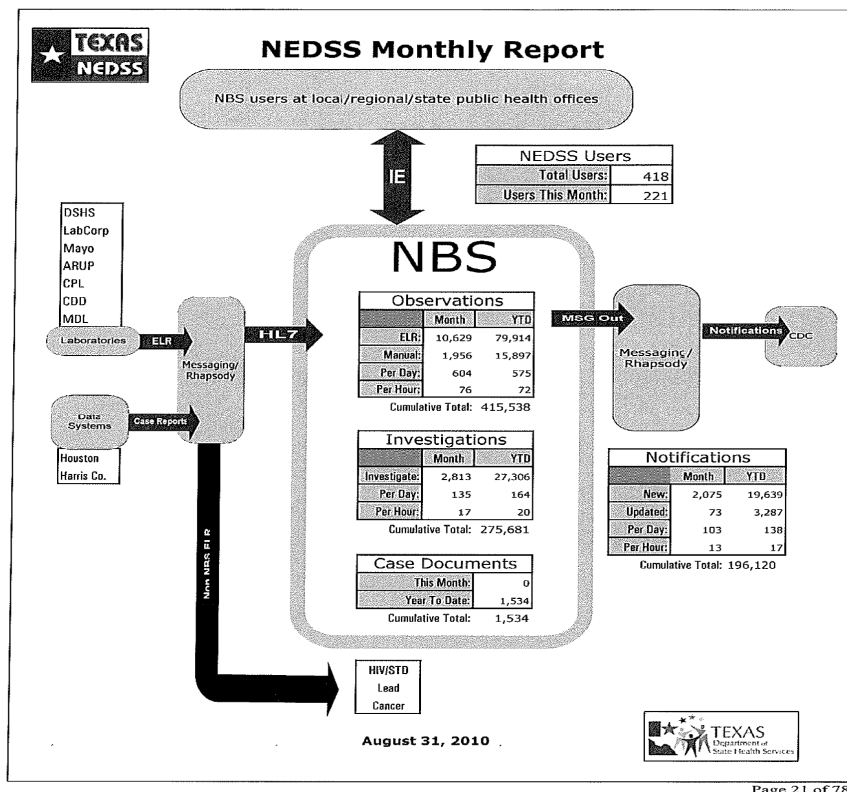


Figure 2: NEDSS ELR Message Volume

For example, in a state like Texas, the NEDSS ELR coordinator deals with an average of 470 ELR messages a day (see Figure 2). This represents only a part of the provision of laboratory result information as it covers only communicable disease surveillance. This number will be in many thousands when Syndromic indicator data is added to the pipeline. It is absolutely critical to utilize a messaging dashboard to monitor the message exchange trend. For this, the informaticist has to consider the following factors when characterizing a message reporting trend

from a state laboratory: due to seasonal variations the volume of messages from the state public health lab (SPHL) is more likely to have a higher volume trend in winter, due to influenza test results, whereas local hospitals and commercial labs are more likely to have seasonal summer bumps due to people being outdoors and in swimming pools, and prone to more infections like Giardia, Cryptosporidium, etc. But this interpretation can also be skewed when outbreaks such as West Nile virus happen during the winter season and then the State Lab trends get both a summer and winter seasonal bump.

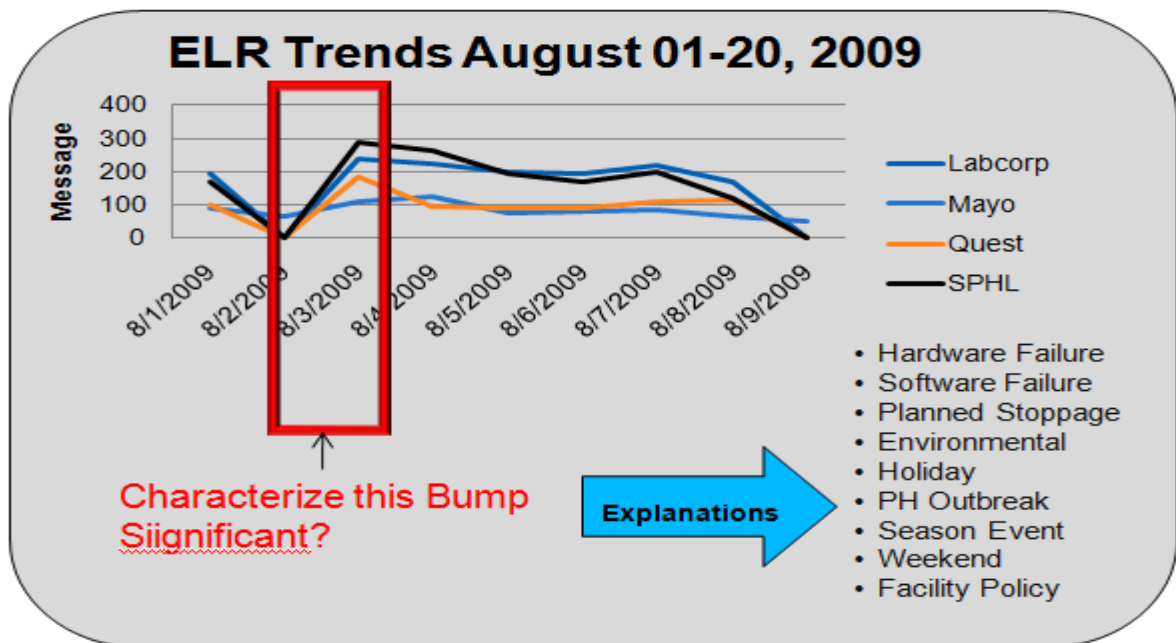


Figure 3: ELR Trend from different sources indicating a pattern

As another example, while characterizing a respiratory syndrome signal, public health practitioners will synthesize and integrate information regarding the pattern of outpatient visits and patient complaints to identify the most probable explanations in a complex list of relevant options. But to understand the significance (rise or fall in the respiratory syndrome) they look for knowledge of a recent intervention (e.g., flu vaccination) or event (e.g., a Rodeo), environmental factors (water quality, air quality) and other findings. Hence to attribute the rise in

the respiratory syndrome to any of these reasons they need to bring together and integrate all this information.

Although all the data are available in the system they are distributed and their representations do not provide any cues to help users integrate them (Parsa, Richesson, Smith, Zhang, & Srinivasan, 2004; Kunapareddy, Mirhaji, Michea, Casscells, & Zhang, 2005). This gives rise to the notion of scoping, defined as: “*given the data, what and how much do we need to describe the data*” (Kunapareddy, Mirhaji, Michea, Casscells, & Zhang, 2005; Parsa, Zhang, Smith, Majid, Casscells, & Lillibridge, 2003).

In current systems, a predefined set of variables deemed important by the system designer is presented to the user. The system places the burden on the user for querying their relevance, integrating them and finally narrowing them down to a few that can be considered to be attributes of the problem.

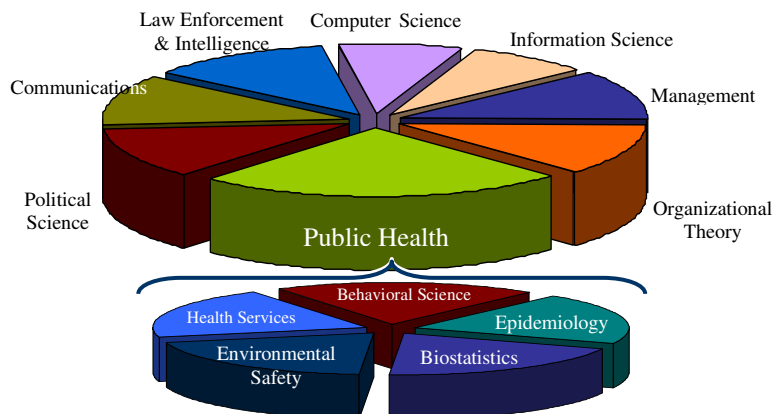


Figure 4. Disciplines required for Public Health

For instance, in Figure 5: Patient Traffic in Houston Hospital ER, the epidemiologist has to understand why the patient traffic fell drastically on a particular day. He can attribute this problem to many reasons, such as closing down of the hospital, or failure in the transfer of data, or maybe it was a day when no one got sick or all the people left the city. To scope the problem

he needs relevant information about the events before or on that day, environmental conditions before that day, and so on, to come to a decision that “yes – 3/4th of the population left the city due to a event, which is enough to explain the drastic fall in patient traffic”. Now the user can scope or narrow things down from a set of five different possible problems to a single problem that can explain the situation. Current representations do not account for external environments, user experience, goal-driven behavior, the information environment or resource constraints and provide no context for interpretation, even if a particular variable could be a possible reason for a given situation. In the next section I will define Situational Awareness (SA) and review situations where it is crucial.

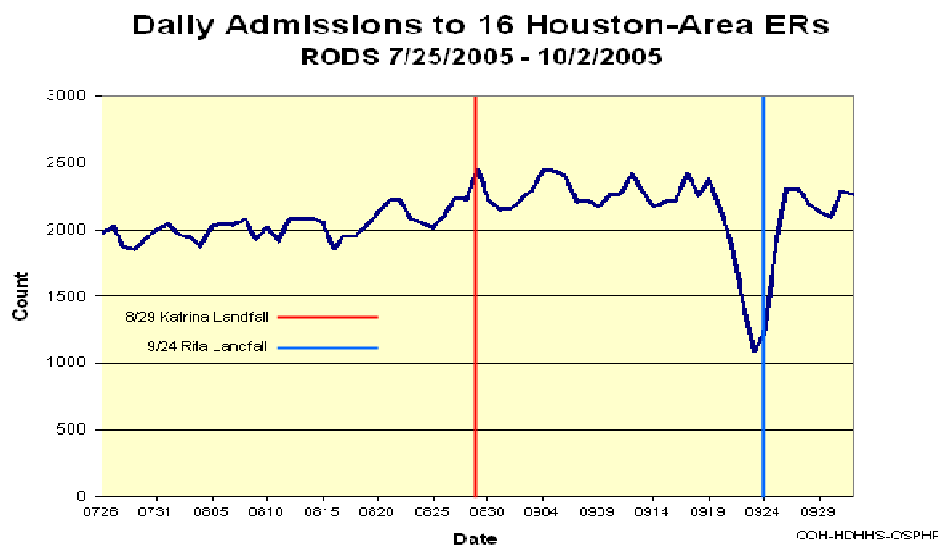


Figure 5: Patient Traffic in Houston Hospital ER

Figure 6 show some screenshots of the most commonly used surveillance system dashboards. The common theme among all these is that they are designed for a specific purpose; what they all lack is providing a global picture. According to the surveillance literature, meaningful and holistic interpretation of data requires generation of higher-level explanations

based on knowledge and expertise using multiple principles. Context is essential to provide a big-picture view of any dynamic and complex problem.

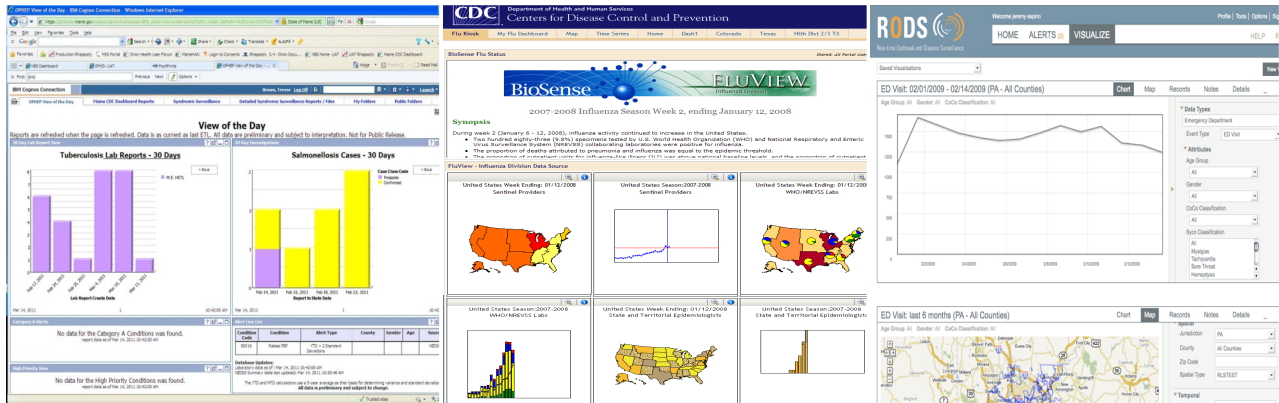


Figure 6: State Surveillance Dashboard with ELR component distributed

This implies that the process for building health information dashboards should consider elements beyond user functions, tasks and goals, by including the user’s situational awareness (SA) requirements.

Defining Situational Awareness

Situational Awareness (SA) is defined by various researchers working on the operational domain (Beringer & Hancock, 1989). According to Endsley, SA as a mental construct is defined as *“the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future”* (Endsley, 1988). Toner further simplified Endsley’s definition of SA as “Understanding what is going on around you. But there is more to this statement than first meets the eye. Understanding is more than information gathering. It implies gathering the right information (all that is needed, but not too much), being able to analyze it, and making projections based on the analysis” (Toner, 2009). This fits a complex domain like public health surveillance where there are new events and new knowledge that evolves continuously, the users must be able to learn from the system to improve their performance (Parsa, Richesson, & Srinivasan, 2004). Current systems

fail to adapt to changes and present the same information regardless of the scenario (Kunapareddy, Mirhaji, Michea, Casscells, & Zhang, 2005). They do not provide information to users that is based on their situations, making it hard for them to understand the scenarios and theorize from the events. For example, a sudden rise in child respiratory symptoms reports in an area every year during the month of August can be attributed to the reopening of the schools. We know that children spend the vacation in different areas and pick up infections. Later, when the schools reopen they pass on the infections upon contact. So we have learnt the lesson that school reopening presents a possible respiratory outbreak in a particular zip code. But this cannot be learnt by using the current systems, as the user is not provided with the relevant data points (school reopening date, school zip codes etc.) on the same occasion so as to present the holistic view. PH surveillance takes place in a dynamic environment and the background knowledge of this environment plays a vital role in decision-making (Parsa, Richesson, & Srinivasan, 2004). For instance, public health experts utilize their previous experiences, for instance, that the average number of respiratory syndromes in some particular zip code is always higher than other syndromes because of its proximity to an industrial area, or an economically impacted zone, or an area with unhealthy living conditions and so forth. Such background information and knowledge are not currently given to users to take account of during decision-making tasks (Parsa, Zhang, Smith, Majid, Casscells, & Lillibridge, 2003). This explains the need for a better information representation system for improving awareness. In the next section, I shall define dashboards and discuss the design principles that govern the current dashboard development process and proceed to explore their gaps.

Defining Dashboards

The dashboards considered here are a refined second generation of the Executive Information Systems (EIS) of the 1970s (Watson, Houdeshel, & Rainer, 1996). The intent of the EIS was to provide executives and managers with an integrated view of the information needed to manage the business (Few, 2009). Despite a great deal of interest in the concept, it was just too hard in those early days of computing to build effective solutions without the advances in processing power, database technology, and data warehousing methodology that have arrived in later years (Few, 2009; Morrissey, 2007). Today, the technology is ripe enough to meet the needs in representation, data processing and management; however, the concept of dashboard itself is still muddled within the industry. Every system user wants one, but not always for the right reasons, and often with little clue as to what is needed for. Like all new technologies, dashboards are surrounded by hype and confusion (Few, 2008). This is because very few dashboard designers (or even dashboard vendors) have fully understood, appreciated, and responded to the unique challenges and opportunities they present (Few, 2009). Caught up in the race to out-play one another, few designers and developers have taken the time to acquire more than a superficial understanding of effective dashboard design.

A search to identify an appropriate definition of a dashboard returned interesting results. A simple search in the Internet for examples of dashboards results in a mix of discoveries that is too eclectic to fit a single definition. Searching through the Business Intelligence (BI) literature for a definition, we find that while much is said about dashboards, few articles try to define them (Eckerson, 2006; Few, 2009), while the definitions that are found are stated, generally, in ways that conveniently fit the software they're promoting. For the purpose of this work, the following

definition by Stephen Few (Few, 2009), which is relatively unbiased, practical, and rooted in real-world experience has been adopted:

“A dashboard is a visual display of the most important information needed to achieve one or more objectives, consolidated and arranged on a single screen so the information can be monitored at a glance.” – Stephen Few (2009)

The process of implementing a dashboard is complex. Designers need to understand the specific user objectives and identify the most important information one must know to achieve them. Information requirements are often not interrelated and come from diverse sources. In a complex environment, the amount of data relevant to a given objective is more than can be presented in a single screen. Unfortunately, current dashboards dump a lot of information with little context or relevance to the objectives of the users. Another major drawback in the design process usually adopted is that few designers spend time planning how to present the information so that human eyes can quickly take it in, and human brains can easily extract the correct and most important meanings from it. To design dashboards effectively, one must understand the objective, the context and some aspects of visual perception—what works, what doesn't, and why. The following section will examine some of the dashboard design techniques currently in use in the industry.

Dashboards that communicate clearly, accurately, and efficiently are the product of careful and informed visual design (Tufte, 1983, 2005). Designing a dashboard starts with choosing the right information to include. CTA techniques, hierarchical task analysis in particular, have been found the most efficient ways of identifying the data needs of the end user (Few, 2006). The next step in the design process is deciding how to display all the required information on a single screen, clearly and without distraction, in a manner that can be

assimilated quickly (Few, 2006). Literature is sparse on the use of information-seeking patterns, or the meaning of data itself, to drive the representation.

According to Stephen Few (Few, 2009), the characteristics required for dashboards are:

- Exceptional organization
- Concise, clear, and often small display widgets
- Emphasis on summaries and exceptions
- Information that is finely customized for the task.

However, today's dashboards tend to be highly visual, lacking the proper emphasis and context, having improper measures and inaccuracies, cluttered and extending beyond a single screen. Current dashboards, built without proper design principles, present information that is not comprehensible, and tend to focus on cute or entertaining elements (Few, 2006). A dashboard representation should support the following processes of visual monitoring, by helping the user to:

- See the big picture.
- Focus in on the specific items of information that need attention.
- Quickly drill into additional information that is needed to take action.

Current tools in the market are driving designers to use technology-centered development rather than User Centered Design. This perspective of development does not factor in the human limitations around information processing (Sexton, AG, 1988); dynamic, colorful widgets lead to increased overload while the operator is handling changing tasks and situations. The operator can only pay attention to a certain amount of information, and if it is scattered, cluttered or not ideally represented then it increases overload and leads to operator error (Nagel, DC, 1988). This research work aims to adopt technologies fitted to the capability of the users instead of forcing

them to adapt to technology. In this research the human factors of data perception, processing, memory and capacity are taken into account while designing the dashboard and the technology adapted to fit to the user's expectations.

Defining Context

While most people subliminally understand what context is, they find it hard to elucidate. Previous definitions of context are done by enumeration of examples or by choosing synonyms for context (Dey, 2001). Dey defines Context as ***“any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves”*** (Dey & Abowd, 1999). Though it is widely accepted that context is needed in improving the human ability to input and interact with computers in both traditional and dynamic settings, there is only a vague understanding of how to apply context to systems (Dey, 2000).

Contextualization is defined as ***“Process of adding context to data”*** and ***“information is the output of contextualization”*** (Edmondson & Meech, 1994). Contextualization involves the immediate data, its history, and the knowledge already possessed by the recipient (Edmondson & Meech, 1994).

A review of the human computer interaction literature identified three different styles of interaction used in contextualization of data as *Communication*: where contextualization happens dynamically between two human users based on their responses to each other; *Tool Usage*: Here contextualization exists as a process only within the human, with no active contextualization by the tool; and *Agency Mediation*: Here one agent adopts a subordinate role, usually that of the tool (Suchman, 1987). Edmondson identified the major factors to be considered during agent

mediation system design as *Experience of the user*: the nature and extent of context needed changes drastically as a function of experience; *Focus of attention*: the level of granularity of the context to be applied (local vs. global); *Filtering of information*: tailoring the amount of information flow to reduce information overload; *Representation*: effectively presenting the information using techniques like spatial organization, layering and navigational paths (Maskery, H & J, 1992). In public health practice today, that context is provided exclusively by human experts (Mirhaji, Srinivasan, Casscells, & Arafat, 2004). A systematic literature review in the field of public health informatics found that there has been no work done in utilizing contextualized information, whilst in other domains, like information and data mining, learning technologies and EMR it is well studied. The next section presents some relevant studies and discusses how the presentation of contextualized information in a dashboard can improve situational awareness.

Characteristics of Previous Approaches to Dashboard Design

Today's dashboards are designed with a philosophy of taking advantage of the latest visualization technology and complex data analysis techniques, however they are not designed to match the mental schemata of the user nor to support situational awareness requirements (Resnick, 2005). Previous studies with dashboards for executives discussed the challenges in current dashboards being designed to deliver data to address a particular problem space but not well adapted to drive decision making when there are changes to the environment (Drews & Westenskow, 2006). Further, situational awareness in today's dashboards is limited to visual cues, such as using colors or indicators, like "a red for critical outcomes and a green for positive outcomes". Information System Dashboards are not designed to match the mental schemata of

the user nor to support the following situational awareness requirements (Few 2006; Resnick, 2005):

- Meaningful and holistic interpretation of data requires generation of higher level explanations based on knowledge and expertise from multiple principles (Parsa, 2004)
- Information systems should consider elements beyond user functions, tasks and goals but also the user's situational awareness (SA) requirements (Endsley, 2001)
- Lack of methods to design for situational awareness in dashboard systems in complex domains (Resnick, 2005; Gledhill, 2002; Huang 2003)

Research Aims

This research work proposes a method of representing contextualized information to improve a user's public health situation awareness, i.e. to perceive, interpret and forecast when utilizing and performing a signal characterization task with public health information system dashboards.

Research Aim 1

- Develop a method to build dashboard systems that will meet a user's SA requirements

Research Aim 2

- Implement a prototype health information dashboard using the new method and empirically evaluate for the Situational Awareness delivered by the system.

Specifically, for aim 1, I undertook a cognitive task analysis (CTA), commonly known as goal directed task analysis (GDTA), to identify the user goals and information needs of a user, followed by another cognitive task analysis process called concept mapping to elaborate the domain and the operational knowledge into a formal, machine processable representation. The

outcomes of the 2 CTA processes are used in developing a representation that will provide higher levels of situational awareness for users by addressing the information needs for their goals and the domain knowledge for delivering context to interpret this information. For aim 2, the proposal is to empirically evaluate the method by implementing a prototype and comparing the SA levels of the users when performing a signal characterization task while using the prototype, when compared to the traditional interface. Finally, based on the lessons learned while developing the method and results from experimental data, preliminary guidelines are presented to assist developers to design systems to produce better SA.

Chapter Summary

In this section, I have presented an introduction to the problem domain. It is evident that in the public health domain, meaningful and holistic interpretation of data requires the generation of higher-level explanations based on knowledge and expertise from multiple disciplines (Parsa, Zhang, Smith, Majid, Casscells, & Lillibridge, 2003). The distribution and the multiplicity of domains, along with the unprecedented and complex nature of events, require access to information sources from a variety of domains (infectious disease, epidemiology, bio-statistics, information sciences, policy making, law enforcement, intelligence, clinical science, pharmacology, environment and others) to enable interpretation of the data (Kunapareddy, Mirhaji, Michea, Casscells, & Zhang, 2005). However, such coordination is unfeasible in real world situations unless the information system can meaningfully integrate and apply context and present appropriate information to the user. These problems highlight the importance of contextualized representation of the information for decision-making in a complex environment. Context (comprising domain knowledge, higher level association between domains, organizational missions, goals, tasks at hand, priorities, past experiences, historic events, recent

events, psychosocial and political constructs, among others) is essential to provide a big picture view of this dynamic and complex problem (Parsa, Richesson, & Srinivasan, 2004). Based on the context awareness literature review it is evident that contextualizing and representing the data appropriately will elucidate non-obvious relationships and new patterns in the data. Presenting context reduces cognitive overload on the user, refines/defines search domains and reveals structures.

To summarize, public health practitioners need context to interpret the data presented to them in the dashboards. Contextualized information can explicate the non-obvious relationships between data and can be used to improve the SA coverage of the public health user by allowing them to tie together the relationships between data. Although the literature stresses the importance of using context, to date there has been very little work advocating how to represent context to improve awareness. It is also unclear whether the principles of using context will be appropriate if they are translated to the public health domain. The next chapter looks into the prior arts used in contextual and human centered approaches to improve situational awareness. I will discuss the gaps in these approaches and lay the foundations for the method adopted in this work.

CHAPTER 2: PRIOR ART

Designing dashboards is a creative process. Recent trends in information delivery have inspired much enthusiasm for delivering dashboards to key decision makers (Few, 2009; Marcus, 2006). When they work properly, they provide a powerful means to digest loads of data; however, most dashboards live up to only a fraction of their potential (Few, 2008). The fundamental challenge of dashboard design is to display all the required information on a single screen, clearly and without distraction, in a manner that can be assimilated quickly (Few, 2006; Farcot, 2010). Most dashboards are used once a day to monitor information, because more frequent use is unwarranted, given the rate at which the information changes and speed at which responses must be made (Few, 2008). Some jobs, however, require constant monitoring in real time, or close to it, because the activities being tracked are happening right at the moment and delays in responding cannot be tolerated.

There are perhaps no better examples of this type of dashboard than those that monitor the message exchange trends at a public health (PH) institution such as the department of health. Much like air traffic control systems or cockpits in airplanes, PH dashboards must be designed to support real-time situation awareness, as defined by Endsley (Endsley, Bolte, & Jones, 2003). They must grab user attention when it's needed, they must make it easy to spot what is most important in a screen full of data, and they must give users the means to understand what's happening and respond without delay (Few, 2008). To do this, they require expert visual design and must display measures of performance clearly, accurately, directly, and without distraction. Traditional dashboard design focuses almost exclusively on defining the right success metrics, and then piecing together a bunch of charts and gauges on a single page. These techniques result in solutions with a hodgepodge appearance that presents confusing information.

Dashboard Design Principles

Existing guidelines for dashboard design are focused primarily on perceptual and physical design of the systems. For example, Figure 4 provides a template to specify what information is going to be presented and the purpose of this representation. A review of the literature lists the following principles when it comes to designing a dashboard (Few, 2009; Inc., 2009; Hansoti, 2010; Clark, Lyons, & Hoover, 2004).

- A dashboard should be guided by important and actionable information and not novel and whimsical desires.
- A dashboard should have a core theme based on the essence of the problem.
- Do not treat all information as equally important
- Suppress ancillary information
- Choose the right metric based on goal
- Deliver only credible data
- Delivery should be transparent and simple; avoid overcrowding of information
- Design based on common interpretation
- Use simple visual indicators.

Other core design principles laid out by researchers include:

- **Compactness / Modularity:** Present information in a way that matches the underlying conceptual model and fits within the user's capacity to receive and process.
- **Gradual Reveal:** Reveal information as the user expresses interest.
- **Guide Attention:** Use visual cues and functionality to draw the user to the things that matter most.

- **Support Causal Use:** Minimize the cognitive barriers by avoiding feature overload, minimizing clicks for each task, and providing clear, concise descriptions of what things mean.

Scope	<input type="checkbox"/> Broad: Displaying information about the entire organization	<input type="checkbox"/> Specific: Focusing on a specific function, process, product, etc.		
Business role	<input type="checkbox"/> Strategic: Provides a high-level, broad, and long-term view of performance	<input type="checkbox"/> Operational: Provides a focused, near-term, and tactical view of performance		
Time horizon	<input type="checkbox"/> Historical: Looking backwards to track trends	<input type="checkbox"/> Snapshot: Showing performance at a single point in time	<input type="checkbox"/> Real-time: Monitoring activity as it happens	<input type="checkbox"/> Predictive: Using past performance to predict future performance
Customization	<input type="checkbox"/> One-size-fits-all: Presented as a single view for all users	<input type="checkbox"/> Customizable: Functionality to let users create a view that reflects their needs		
Level of detail	<input type="checkbox"/> High: Presenting only the most critical top-level numbers	<input type="checkbox"/> Drill-able: Providing the ability to drill down to detailed numbers to gain more context		
Point of view	<input type="checkbox"/> Prescriptive: The dashboard explicitly tells the user what the data means and what to do about it	<input type="checkbox"/> Exploratory: User has latitude to interpret the results as they see fit		

Figure 7 Guidelines to specify what gets into the dashboard (Few, 2006)

- **Lead to Action:** Empower users to finish their task quickly and/or understand the action that should be taken on the basis of the results.

The dashboard design guidelines discussed above focus on the perceptual and physical design of the system and not on the human cognitive attributes. A number of studies based on human factors include detailed data on the presentation of various types of visual, auditory and tactile information and their impact on human perception of information (Boff, Kaufman, & Thomas, 1986; Sanders & McCormick, 1992; Salvendy, 1987). These studies help in the design

of various information representation widgets, such as buttons and graphs, to minimize human error in the perception and interpretation of information. These studies help in designing systems that help to control the influences of environment, noise and external events and serve as a foundation to build situational awareness. These guidelines serve as building blocks for achieving SA, however they do not remain constructive of SA when the user operates in a demanding setting requiring interaction between different components that tax the human memory with data and cognitive overload. While guidelines for integration still continue to be studied (including standardization across the system, compatibility between control and display and methods for arrangement of controls), there is a need for guidelines that address SA. Research in behavioral psychology has led to the in-depth study of human cognition, specifically work in the fields of cognitive engineering and naturalistic decision-making has opened the door towards designing systems for SA by targeting human cognition. Before we expand dashboard design principles to accommodate SA, we need to understand the design processes in place. The following sections discuss current design approaches and how their principles can be implemented for improved SA.

Dashboard Design Process

Implementing a dashboard is not much different from typical system design processes used across many different industries. This section discusses briefly current processes of system implementation to better understand where and how considerations of SA can be introduced. The most typical systems implementation approach used in small and medium sized projects is the waterfall design model (Hoffer, 2002; Royce, 1987), which is the simplest and easiest to implement due to its linear nature. The original model published by Royce in 1987 had seven stages (see Figure 5). However, within the software industry various modifications were brought

into the model to accommodate project specific requirements. The main advantage of the waterfall model is the minimal resource requirements needed for its implementation. Furthermore, the documentation produced at every development stage of the waterfall model is extensive and complete due to its defined functions. In spite of these advantages, the waterfall model does have significant limitations; ironically its greatest advantage is also its biggest drawback. The linear approach allows very little space for correction, for example, if the design phase has gone wrong, things can get very complicated in the implementation phase (MacLean, 1989; Parnas, 1986). The model doesn't offer sufficient room to allow for backtracking and correction at a later stage (Parnas, 1986). Furthermore, the model doesn't scale well to support projects where there are unclear or evolving requirements.

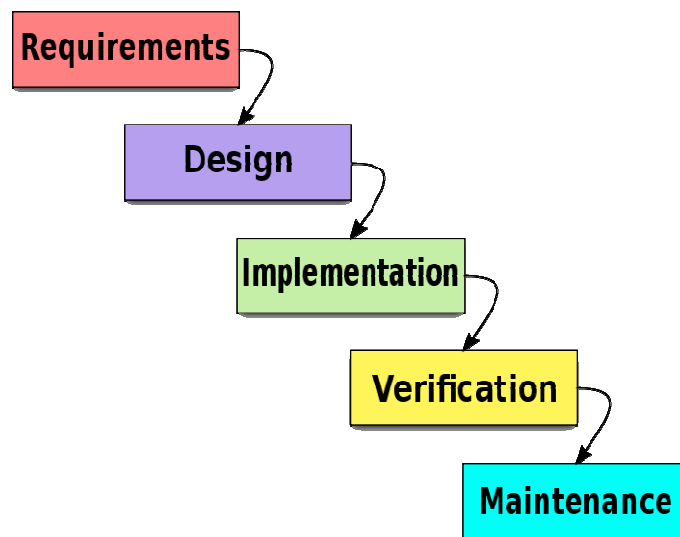


Figure 8: The unmodified "waterfall model" as defined by Royce

In order to address the gaps in the waterfall model, the industry started to implement a concurrent engineering model. The philosophy of the concurrent approach is to parallelize selected activities in the waterfall model, or the entire process itself. Various types of implementation of concurrent models are widely used in the industry. The spiral design model is

one of the most frequently used variations (see Figure 6). This approach accounts for rapid development of software solutions with information from each phase feeding to the next phase to accommodate changes (Boehm, 1986). The spiral approach has the advantage of quicker releases with a subset of capabilities, while the system evolves as requirements of users become clearer and more extensive based on feedback from previous stages.

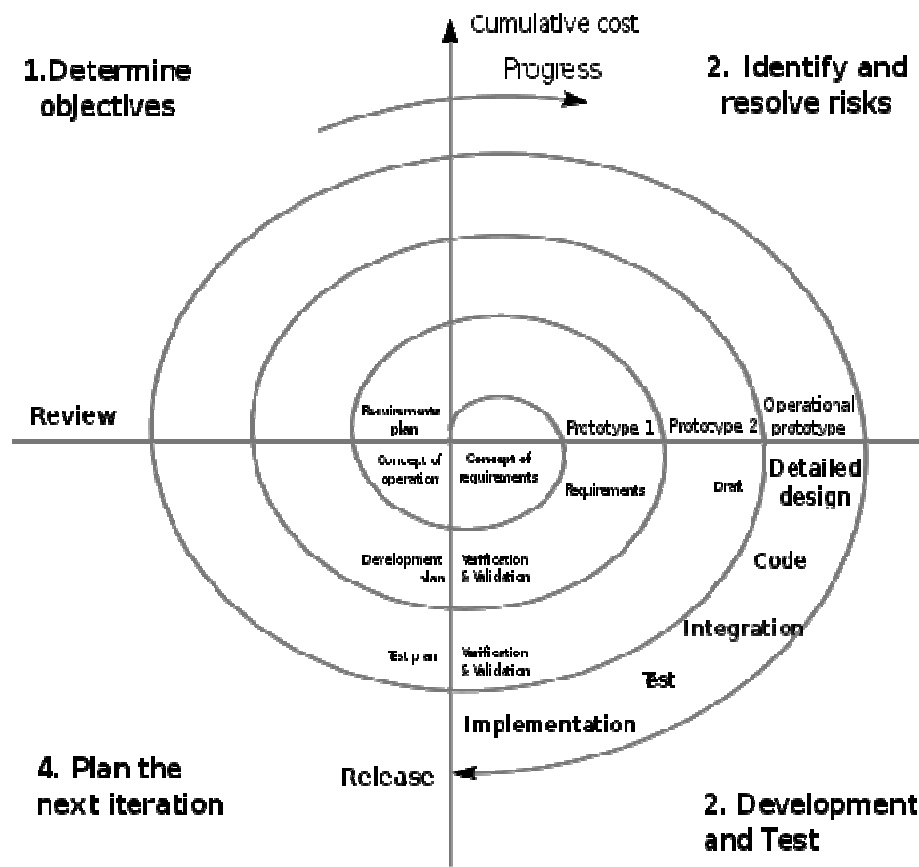


Figure 9 Spiral Design Process- Example of a Concurrent Engineering Model

The concurrent process does have major challenges in implementation; however when managed wisely it can implement systems that capable of delivering SA. In the user interface design process, the requirements phase is the most critical. Requirements analysis focuses on identifying or translating the broad goals and objectives of the user into system capabilities; it is a challenging activity that requires an understanding of the vision and concepts of the current

systems that will have to be mapped to the user's mission, which then needs to be translated to capabilities delivered by the system. There are three schools of thought in how to design user interfaces: technology-centered, user-centered and ecological. The following section will focus on the conceptual frameworks that are currently promulgated in designing user interfaces, and will discuss the advantages and current challenges in delivering situational awareness.

Technology Centered Design Approach for Dashboards

This approach is based on displaying data pertinent to a user function from all available sensors, e.g. speedometer, engine temperature and gas gauge for the driver of a vehicle. This model would work only if the information to process is limited; the huge volume of information available for users to process and the reality of changing tasks and situations facing the operators present a challenge (Sexton, 1988). The users are required to find, sort, integrate and process the information available, inevitably leading to gaps in processing. Human operators have some inherent challenges in processing information; various human cognitive factors come into play because of the need to pay attention to information that is often scattered and sometimes not even granular enough to process. Technology-centered design is often blamed as a causal factor for 60% to 85% of all accidents (Nagel, 1988). The Union Carbide accident in Bhopal, India is a perfect example of an error induced by technology-centered design. Casey in his study found that the system's interface design did not support the operator in detecting significant cues to the problem but was more designed to present information on status at a point in time. The study further showed that information needs by users were different from the actual information presented (Casey, 1993). This led to an approach placing importance on user mission and goals and on designing an interface to support these. The following sections discuss in detail the user centered design approach towards implementing dashboards.

User Centered Design (UCD) Approach for Dashboards

Wikipedia defines UCD as a design philosophy and a process in which the needs, wants, and limitations of end users of a product are given comprehensive attention at each stage of the design process (Wikipedia, 2010). UCD is one of domains intensively studied in the cognitive science world. It is a human-centered approach, in contrast to the technology-oriented approach. UCD prioritizes user mission, goals, task and needs when dealing with information. The philosophy was born not from a humanistic desire but to obtain optimal functioning of the human machine system (Endsley, Bolte, & Jones, 2003). To better understand UCD it is critical to know what it is and what it not? UCD is not about asking users what they want and giving it to them (Endsley, Bolte, & Jones, 2003), the primary reason being that users have incomplete ideas about what is a better way to access information than they are currently used to. This will produce complications when an integrated view of data is needed to achieve the goals (Endsley, Bolte, & Jones, 2003). Furthermore, a single interface might need to support multiple users and so the requirements may quickly become overwhelming. Hence designing a system interface based on user input although considering user's vision, working environment, external factors, and information needs, can neglect the importance of the dynamics surrounding human understanding.

The scientific literature also recognizes that UCD is not presenting only the information needed by users at a given moment. Although it sounds logically sound that information should be presented to address user tasks, it is critical to understand that this leads to significant problems, the most significant among them being the inability to track the goals of the users (Graeber, 1996). In real-world scenarios user goals tend to change based on the information available. This dynamic scenario could lead either to presenting or distracting the user from

critical information needed for specific tasks (Hancock, 1988; Jones, 2000). Even if the system is constructed to retrieve user goals and tasks via probing techniques, it may lead to a situation where the user constantly needs to perceive and interpret changing information. New information may not always fit within the same representation scheme and this could lead to mental overload for the user. This information filtering concept has been well studied in published literature, finding that it tends to make the user reactive rather than proactive as he/she waits for cues from constantly changing information (Jones, 2000; Moray, 2000). Further scientific papers also indicate that UCD is not making the systems take decisions for the users. There is a potential risk that utilizing decision support or expert systems to aid in user interface design could lead to failures, especially if there are ambiguous responses. Studies show negative impact and bias if the expert systems' recommendations are incorrect (Selcon, 1990). Further studies have also shown that the overall decision-making and performance was slower with expert systems, compared to traditional representation (Endsley & Kiris, 1994). Other studies have also shown that if the goal of the user is not transparent enough, decision making systems could provide advice that is not useful but forces the user to modify activities, resulting in poorer performance. Also UCD is not producing benefits for the user. The scientific literature clearly indicates that shifting roles to a system has a negative impact on user performance (Endsley, Bolte, & Jones, 2003). The user is left out of the loop in the task and it imposes a burden on the users in catching up with the situation. The process of automation is a well-studied area, where careful consideration needs to be given to solving the problem space.

So what are the principles governing UCD? Endsley discusses three core principles that form the foundation blocks for this research. The next section discusses why each principle is critical and how they help deliver the building blocks for SA. The first principle is to organize

technology around user goals, tasks and abilities. This stems from a shift of focus away from developing an interface to address human tasks toward designing interfaces that conform to human abilities (Sanders & McCormick, 1992). Traditionally, Human Factor (HF) and ergonomics studies have sought to design systems that will not require users to perform tasks that exceed their mental or physical abilities (Endsley & Kiris, 1994). In recent years more focus has been directed on the mental ability of the users. Task analysis in the design process has become a *de facto* method to determine what information is needed to support user tasks. The HF approach is most suitable for linear and repetitive tasks, whereas the scientific community has invested in User Centered Design process (UCD) that is focused more on addressing complex scenarios in which the user doesn't follow a linear set of activities and the goals change over time.

The UCD approach is still a goal-oriented approach and the interfaces are designed to aid the goal-oriented information processing of the users. For example, in the battlefield a commander should be able to switch from one goal to another, e.g. offensive to defensive (Selcon, 1990). But to do this the commander should have access to all pertinent information from various sources to make this decision and achieve the goal. The focus of the UCD approach will be to identify the information needs for various goals and to deliver the information specific to those. The challenge lies in dealing with changing environments and uncertainty (Vicente K. , 2002). UCD approaches this by aiding the decision-making process of the user by keeping the representation closer to the user's mental model of the situation. Studies done across multiple fields show that experts operate by performing pattern matching to search long term memory to better understand a situation (Mintzberg, 1973; Kuhn, 1970). Situational awareness (SA) is a key mental construct that needs to be achieved through the system interface to perform decision (Endsley, 1988). The

UCD process needs to support the cognitive processes of the operator by delivering better SA. In this research the UCD core principles of the goal-oriented approach will be used in developing a method for designing complex dashboards. Chapter 3 will expand on the method and its implementation. In the following section, another philosophically different approach to interface design is discussed and the challenges and advantages are discussed.

Ecological Interface Design (EID) Approach for Dashboards

As an evolution from the principles of ecological psychology and direct perception, a concept of interface design focusing on presenting objects to operators to make them active problem solvers as opposed to passive monitors was developed (Thorvald, 2009; Vicente K. R., 1992). The goals of EID approach are twofold: first, not to force processing to a higher level than the demands of the task require, and second, to support each of the three levels of cognitive control (Rasmussen, 1983). In the past decade, researchers have reported progress in applying EID as a framework to design interfaces for a variety of work domains of increasing complexity (see Vicente, 2002). Some of the complex areas where EID have been demonstrated are in Anesthesia (Drews & Westenskow, 2006), Transport Safety (Lee, 2006), Nuclear power plant (Itoh, 1995) and Aviation (Dinadis, 1995).

The focus of EID is to design displays by presenting system users with the constraints and opportunities for action in the environment. Users in an environment are often presented with two types of situation, anticipated and unanticipated. EID's goal is to deliver a representation that will aid the process of responding to both situations (Jamieson, 2003). EID strategy for the anticipated case is by identifying the best path of action based on what is known about the work domain. Once the path is determined, the process and steps are organized efficiently to reach the user goal. This strategy can be effective when the process states are

relatively static and predictable, and have a predetermined starting point. However in the alternative, unanticipated, case, the strategy is to show the constraints of the environment, and to depend on the operator to choose a feasible path to get to the goal. This can be effective when the constraints are dynamic, changing in predictable and unpredictable ways, and when the starting point is not predetermined (Flach, 1995). So EID is based on 3 basic premises, the first one being the challenges around users dealing with unanticipated (or abnormal events beyond the normal) and anticipated events (Jamieson, 2003). The second premise is that people have different ways of carrying out their tasks. While some activities are so routine (linear) that operators don't even have to think about it; they simply see an indication that the task should be done, and they do it automatically. In many cases in complex domains tasks operate under a rule-based model e.g. if the temperature of the boiler is high then confirm that the pressure in the pipes are within limit. In this case, the person consciously reviews the situation and interprets why an undesirable process state has occurred, and plans the appropriate sequence of actions to bring the process to a more desirable state (Christofferson, 1998). The third premise of EID is that an effective visual display can present information in such a way that people directly perceive process relations and states (Gibson, 1988).

The EID design approach is based on two theoretical foundations (Flach, 1995; Vicente, 1992; Rasmussen, 1983) a. Abstraction hierarchy, b. Skills, rules & knowledge taxonomy (SRK). The abstraction hierarchy is a framework that is used to develop models of work domains. It is used to represent the constraints in the work domain in a way that will allow the user to handle events. Accordingly, EID begins with work domain analysis (Vicente, 1992). The abstraction hierarchy is a framework that can be used to develop models of particular work domains. It contains function information that describes the state of the objects of interest to a particular goal

(Vicente, 2002). Work domain analysis is different from task analysis because the latter is conducted only for anticipated tasks, while the former is focused on extrapolating the system function, independent of any particular user, automation, event, task, goal, or interface.

So once the information is identified the next question is how can the user process it? The SRK framework is used to describe the mechanism by which the users process information. It is a widely accepted framework within which three mutually exclusive ways of interpretation of information determine the level of cognitive control that is activated for processing the information, namely skill based behavior (SBB), Rule based behavior (RBB), or Knowledge Based Behavior (KBB). The *skills, rules, knowledge* (SRK) taxonomy describes three qualitatively different ways in which people can interact with their environment (Rasmussen, 1983). Skill-based behavior involves parallel, automated, direct behavioral interaction with the world. Rule-based behavior involves associating a familiar perceptual cue in the world with an action or intent, without any intervening cognitive processing. Knowledge-based behavior involves serial, analytical problem solving based on a symbolic mental model. To achieve these aims, the framework comprises three design principles, each directed at supporting one level of the SRK taxonomy:

- *Skill-based behavior (SBB)*: Workers should be able to act directly on the interface.
- *Rule-based behavior (RBB)*: There should be a consistent one-to-one mapping between the work domain constraints and the perceptual information in the interface.
- *Knowledge-based behavior (KBB)*: The interface should represent the work domain in the form of an abstraction hierarchy to serve as an externalized mental model for problem solving. The design goal is to adopt these two theoretical constructs and use them to build the

interface. In the following section I will look at how the two frameworks are applied in design process of a user interface.

The problem of interface design for complex systems is summarized in Figure 7. There are three general parts to the representation: complex work domain, interface, and operator/user (Jamieson, 2001). There are three general steps for designing graphical user interfaces based on EID principles.

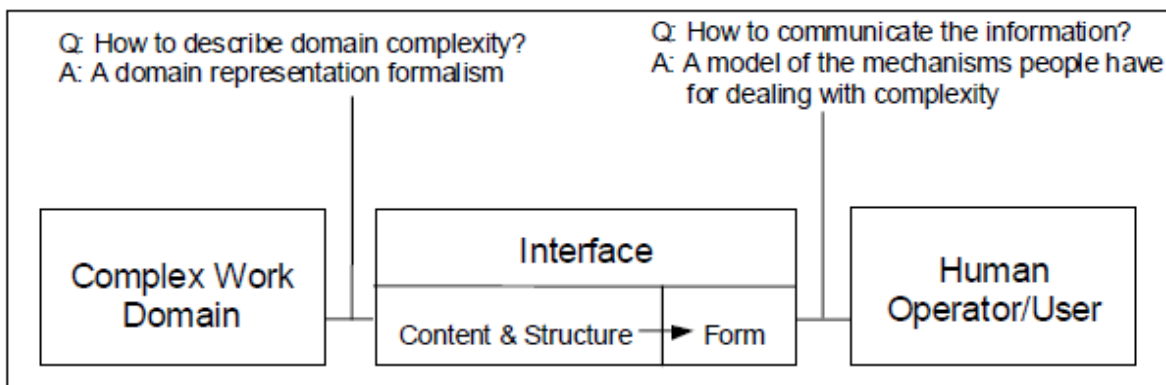


Figure 10 Interface design for complex systems using EID, adapted from Vicente, 1992.

- The first step is to conduct a work domain analysis, which specifies the functional relations that a user should be aware of. These functional relations become information requirements for the user interface.
- The second step is to conduct task analyses, which specify the context-specific decision and execution requirements for assorted tasks that a user is expected to do. These decision and execution requirements also become information requirements for the user interface.
- The third step is to integrate the information requirements from the first two steps into a meaningful graphical representation.

A system implemented using the EID principles will deliver direct manipulation of the data and ensure that all domain constraints are available readily via a graphical interface.

Furthermore, all information identified using the abstraction hierarchy framework (work domain analysis) should be available to the user. This suits EID based interface systems for handling both anticipated and unanticipated situations. Studies have shown that systems developed using EID process have produced improved performance, for example, (Reising & Sanderson 1998, 2000a, 2000b) in a milk pasteurization unit, (Sharp & Helmicki, 1998) tested the value added by EID in the context of neonatal intensive care unit.

Summary of Design Approaches

In this section I discussed two philosophically different approaches towards implementing user interfaces. The UCD approach is more geared towards linear and anticipated tasks and the other is focused on unanticipated tasks. In a complex and dynamic environment where there is a constantly evolving and changing state, it is critical that a goal-oriented interface is delivered to fulfill each user objective. The UCD recommended approach of Goal Directed Task Analysis methodology will help determine the data needs related to the user goal. These form the building block for SA. There are well-published guidelines and principles for designing systems using the UCD principles; common steps include a user analysis, an environmental analysis, a task analysis, a functional analysis, and a representational analysis. Each of these analyses provides different, but necessary, components in order to build a comprehensive system. One of the areas where UCD has gaps in the design process is that, when it comes to delivering support for unanticipated events or external impacted events within the same infrastructure, there are challenges in the design process and in deciding what information to present.

The alternative approach developed by the scientific community to address the gap in UCD approach is the ecological interface design process. This process works on the two fundamental premises; a. Abstraction Hierarchy and Skill-rule-knowledge based (SRK). The abstraction hierarchy offers a framework for understanding the work domain, while the SRK taxonomy provides a way of classifying a user's cognitive task demand. Although the theoretical model sounds solid, the literature does list the practical challenges in adopting EID. Further, the scientific literature available in this area lacks depth in two critical areas. First, it offers few applications of the framework to real work domains. Second, it tends to focus on the design *product* rather than on the design *process* (Reising & Sanderson, 2002). Both EID and SA contribute to the development of information displays that improve operator insight into decision-making spaces. But a review of the scientific literature available indicates that the EID researchers are not accounting for SA in their work when they employ EID in their system design (Burns, 2007).

In complex and dynamic environments, decision-making is highly dependent on situational awareness. In reviewing the two distinct approaches for interface design, it is evident that SA is essential for decision making that in turn leads to performance. Despite the convergence in these objectives of UCD and EID with SA, the concepts have independently evolved and needs to overlap. High SA will depend on delivering goal-driven representation of information that is relevant to the work domain. In this research the design concepts from both UCD and EID will be adopted to develop a hybrid method of leveraging the work domain concepts, user skill, rule & knowledge based behavior to address particular user goals. Chapter 3 describes the method in detail followed by evaluation of the SA in a system developed using this hybrid model.

Situational Awareness and Mental Models

Understanding the interactions of cognitive constructs, such as attention and mental workload, with higher order psychosocial concepts, such as mission, goal, task and function, can be useful in the study of human performance (Wright, 2004). Situation awareness is a higher-level cognitive construct that can be conceptualized as a cognitive state that corresponds to the mental and perceptual state of operational insight of mission and task in a situation, its progression and its relationship with the environment (Wright, 2004). Situation awareness can be explained as the internal mental model of an individual that represents the current state of a dynamic environment.

Even though the conceptual framework of SA can be applied to almost any domain, it has been especially evaluated for air traffic control, aircraft piloting, combat command and control, tele-operations, and some medical procedures (anesthesiology). What is common to all these domains is that:

- a) Multiple competing goals are active at any given time. Operators need to prioritize and time-share between competing goals and tasks;
- b) Multiple and diverse sources of information need to be objectively and constantly inspected for cues. This may overload the limited cognitive resources available to a human operator and increase the probability of error;
- c) Limited time resources are available for interpretation of information and making high impact decisions (Endsley, 1999).

Four generic patterns of functions are identifiable in the domains where SA framework has significant relevance:

- a) **Monitoring:** active and systematic collection of information to analyze and understand the status of the environment;
- b) **Generativity:** formulating opinions about the significance of the events in the environment, projecting the future status of the system and developing strategies to achieve goals;
- c) **Selection:** realizing the relevant courses of action available at any time and selecting a particular option or strategy;
- d) **Execution:** carrying out the selected option successfully (performance).

Endsley has formally defined SA as *'the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future'* (Endsley, 1988). According to this model, SA is a construct meaningful within the context of the mission and tasks, and it interacts directly with the decision-making, performance and other cognitive processes (Endsley, 2000) (Figure 8). Endsley has recognized three levels of SA: Perception, Comprehension, and Projection ((Endsley M. , 2000):

Level 1: Perception of the important informational cues from environment

Level 2: Understanding of the meaning and significance of the information in the context of the task at hand,

Level 3: Forecasting future status of the environment (events or actions) accordingly

Figure 11 schematically demonstrates Endsley's model of SA, defined in this model as pertaining to the individual's knowledge about the state of the dynamic environment and does not include background knowledge, experiences, and established rules that are static knowledge sources (Wright, 2004). This is not to dispute the importance of such static components, as they might influence and support the SA. For instance, experiences from the past can guide or bias the

individual's attention, affecting the formation and quality of his SA. As obvious from the this model, SA is a dynamic construct and it is continuously changing as the environment changes, either due to the new information from environment, results of decisions and actions of the individual or due to other outside influences (distraction, workload, limitations of human cognition, etc.).

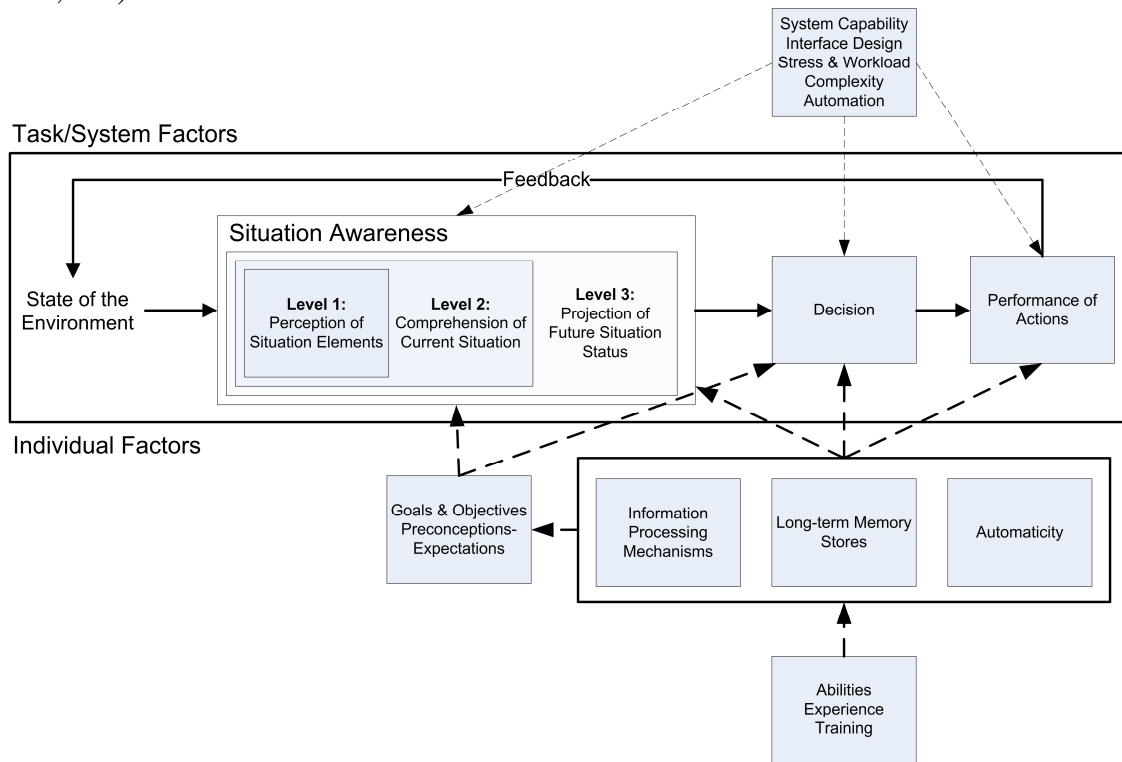


Figure 11 Endsley's Model of Situation Awareness.

Advantages of Situational Awareness

There are four reasons why we believe that understanding and applying the SA construct is important in the design of strategically important information systems (Klein, 2000):

- a) *SA can be linked to performance.* This claim has obvious face validity as it is expected that the more up-to-date cues from environment and the better the understanding of a dynamic situation, the more adaptive the responses can be. Measures of SA have been correlated with performance in aviation research (Wright, 2004).

- b) *Inadequate SA may be associated with errors.* If relevant and required information is not available or is not correctly interpreted and understood, (due to failures of memory or attention or due to system failures, etc.), the probability of errors is increased. Bell and Lyon found that fighter pilots with lower ratings of SA during a combat scenario had significantly greater number of decision errors than pilots rated highly for their SA (Endsley & Kaber, 1999).
- c) *SA may be related to expertise.* Experienced physicists have been shown to classify physics problems differently from novices (Goldberg, 1970). This might imply that the mental map of the environment and its dynamism is modeled and constructed differently in an expert brain. Such models, if captured and formally represented, can be a basis for evaluation of SA in others and for design of systems that support formation of SA comparable to an expert.
- d) *SA is the basis for decision making in most cases* (Endsley, 2000). Endsley's model of SA and Klein's recognition primed decision model have been proposed to explain this phenomenon. It is important to note that Endsley believes in precedence and separation of SA from the process of decision-making and performance. It is possible to have the perfect SA and make less than optimal decisions. For example, two practitioners may have the same SA, but choose different courses of action based on their prior experience, training, goal or personal preferences.

Situational Awareness in Public Health

Although applied in other areas, in the domain of public health situational awareness is an explored concept. In the following examples I try to illustrate the relevance of Endsley's three-layered SA model in a typical public health preparedness setting:

Level 1 SA: Perception of the cues and important elements of the environment. The first step in achieving SA is to perceive the status, attributes, and dynamics of relevant elements in the environment. For a public health practitioner, this may include awareness of numbers of patients visiting health care facilities throughout a geographical distribution, types of health problems and their relative frequency and spatial distribution (vital signs, chief patient complaints, laboratory test results, etc.), availability and relative distribution of health services resources (practitioners, medications, vaccines, beds, etc.). They should also be consistently aware of those community events with potential relevance to the psychosocial and behavioral aspects of public health. For example events such as sports (Super bowl, Rodeo, etc.), promotions and advertisement campaigns (a new OTC medication advertisement can locally and abruptly increase its sales and consumption), holidays and other special occasions can affect public health behavior and the expectations of public health practitioners. Awareness of the actions of other collaborators (such as vaccination campaigns, results of investigation under way for certain incidents) and awareness of availability of resources such as (available emergency beds, antibiotics stockpiles, ventilators, etc.) are also critical elements of level 1 SA for the public health practitioner in a preparedness setting.

Level 2 SA: Understanding the meaning and significance of events and observations in the context of the current situation and missions and tasks. Situation awareness level 2 involves comprehension of the current situation based on a synthesis of the separate level 1 elements, and combining this data to form a holistic picture of the environment in the light of one's goals. For example, public health practitioners will synthesize and integrate information regarding patterns of outpatient visits and patient complaints to identify the most probable explanations in a complex list of relevant options. They will understand the significance of a sudden rise or drop in

certain data elements (number of respiratory distress cases) based on knowledge of a recent intervention (e.g., flu vaccination) or event (e.g., a rodeo started last week and will go on for two weeks) and other findings. They will interpret findings to understand whether they represent an expected and temporary event or are a serious problem. In this example a public health practitioner may consider a sudden surge in number of respiratory distress cases in certain areas of the city, despite undergoing flu vaccination, as an expected and normal finding, considering the fact that the Texas Rodeo has started since last week (overcrowded environment, increased population mix between local and non-local population, etc.) and areas with most cases are areas closest to that event.

Level 3 SA: Projection of the future status of the information by integrating the composition and the dynamics of the environment (level 1 and level 2 SA) in a temporal perspective. SA level 3 is the highest level of situation awareness and may include as well the backward projection of the events (predicting the past). In our scenario, public health practitioners with a high degree of level 3 SA will be able to project the rise of the respiratory distress syndromes in the few coming days and anticipate the increase in OTC medication sales as reported by pharmacies and grocery stores. This type of projection is very important in enabling a proactive surveillance by foreseeing future needs and planning ahead.

Challenges in Achieving Situational Awareness

Achieving and continuing to maintain SA is a difficult process, especially in a domain where there are numerous information sources. Endsley argues that users spend a majority of the time ensuring that their mental picture (snapshot) of the situation is correct and is updated to reflect the current state (Endsley, Bolte, & Jones, 2003). The issue in achieving and managing SA stems from both the human processing mechanism and the complex domain system that the

user interacts with. Endsley coins the term “SA Demons” to describe both these human and domain challenges. This section briefly discusses some of the key SA demons and why they should be considered in the design process:

- a. Attention Tunneling: When users are processing various pieces of information, this is highly likely in situations where the user is fixated on one set of information and excludes the rest (Baddeley, 1972). This could lead result in critical loss of SA. A real world example of this issue was the Eastern Airlines crash in Florida where the pilots ignored the flight path but were fixated on an indicator (National Transportation Safety Board, 1973). The system design should account for the effects of tunneling and provide mechanisms to counteract them (Endsley, Bolte, & Jones, 2003).
- b. Requisite Memory Trap: This is created due to the over reliance on human memory to retrieve information that was available to the user earlier to make decisions. Miller’s contention of human ability to hold seven plus or minus two chunks of related information is a core principle that needs to be considered while designing an interface (Miller, 1956). When users need to collect situation information, they need to hold it in memory and relate back to access it. This could be a problem if the volume of information is large, because then we begin to see lack of memory space to hold new information or a decay of existing information. An example of this demon is the LA air traffic controller failing to retrieve the situational information of having a flight on the runway ready to take off and at the same time allowing a flight to land on the same runway (National Transportation Safety Board, 1991). System designers should consider not relying on users to recollect information for interpretation but allowing direct perception as much as possible

- c. Workload Anxiety, Fatigue and other Stressors (WAFOS): SA can be severely dented if the user is stressed by workload or fatigue (Endsley, Bolte, & Jones, 2003). The stressors can also be caused by external environment the user is exposed to e.g. lighting in the room, noise etc. When users are stressed, studies have shown that people pay less attention and gather less information compared to normal situations. Often people tend to make decisions before perceiving all the information available to them (Klein, 2000). Systems should be designed to counteract this critical shortcoming.
- d. Data Overload: A complex environment can have a number of variables and measures to track. If the user were to face all this information and also keep in step with changes, it can quickly create a mental overload and also overburdens the person's sensory capability (Endsley & Kaber, 1999). Further, if data representation is disorganized and the user has to search for different pieces of information, SA will fail. In system design the concept of goals and information pertinent to the task should be considered while choosing the representation model.
- e. Misplaced Saliency: The human perceptual mechanism is tuned to react to certain triggers. E.g. a flashing red light or a colorful billboard or a stop sign. This is caused by a concept called saliency. Saliency is the compellingness of certain forms of information determined by physical characteristics (Few, 2006). So saliency can be used to improve SA or hinder it. In the design of the dashboard, proper attention is to be provided when certain elements or effects are placed in the interface. Unfortunately, tools and gadgets are overused in many places in the real world. Less important information can be made to appear important by providing alarms, buzzing or flashing elements. This dilutes the importance contrast between less important and critical information.

- f. **Complexity Creep:** External systems often deliver a complex set of information because they are designed to support a myriad of features. E.g. Television Remote Control. Studies have shown that even after years of experience, users still pose significant problems in understanding the operations of certain features. This refers to the overload of information added with complexity of tasks supported by the systems. It is critical to be transparent about the operations but at the same time we need to ensure that the ability to interpret. Training is often offered as the solution to this problem.
- g. **Errant Mental Models:** Studies shows that human operators utilize mental models and schemata to relate and act on situations. However, there is always the risk of mapping to incorrect mental models. This can be caused by external system display added with human factors. In one study, 66% of the participants failed to recognize the mistaken cues and began associating them with wrong mental model. So it is critical that system designers avoid leading users to errant models. Standardized usage of display elements will reduce the occurrence of such errors.
- h. **Out of Loop Syndrome:** This is an error caused by excessive automation in the system, in which the user is kept out of the loop. This will lead the user to believe that the system is in one state when actually it is not. When automation is on course, being out of loop may not be a problem, but when it fails, systems should have the capacity to notify the user efficiently to bring the user into the loop.

Measuring Situational Awareness

Multidimensional measures of SA have been shown to be sensitive to differences in information seeking (level 1), information interpretation (level 2), and projection of future courses of events (level 3), that are not reflected in traditional performance measures (Endsley,

1990). For example, Endsley conducted a study comparing the SA of pilots using a new avionics system with that of others using the old system. Although mission performance measures showed no differences between the two, a direct multidimensional SA measurement technique, Situation Awareness Global Assessment Technique: SAGAT (Endsley, 2000) showed that the new system provided pilots with significantly better SA regarding knowledge of enemy aircraft location and other critical factors compared to the old system (Endsley, Mogford, & Allendoerfer, 1997). The multidimensional evaluation of the SA of pilots made it evident that pilots using the old system were aware of significantly fewer enemy aircraft (level 1 SA), that they had a significantly reduced understanding of what was happening in the overall situation (level 2 SA), and that pilots had reduced knowledge of where aircraft in the field were going (level 3 SA). These studies suggest that measures of SA can have diagnostic and explanatory powers beyond traditional measures of performance. It also suggests that SA measures may be predictive of performance problems or errors that are not seen within the limited sensitivity, scope, or time involved in laboratory studies or in a real world situation.

The measurement of SA could help in the identification of performance problems and error mechanisms (possibly induced by information systems with poor user interface, poor information representations, poor information seeking strategies, or poor communication and teamwork among the collaborators in a distributed environment) (Wright, 2004).

It can be also used as a method of evaluating training needs by identifying areas of deficiency (that is, areas where individuals fail to attain the needed levels of SA). The results can be used to improve training and education of public health practitioners. It is also possible to use measures of SA to evaluate the efficacy of training programs or new procedures, tools and systems in improving the performance and addressing needs (Endsley, 1988).

Methods of Measuring Situational Awareness

In the situational awareness literature there are three major types of measurement strategies that have been employed. As shown in Table 1, the explicit and implicit measures are objective measures, where the latter assumes that a subject's performance correlates with SA and improved SA will lead to improved performance while the former does not.

Table 1 Measurement strategies for situational awareness

Categories	Subcategories
Explicit Measures	<ul style="list-style-type: none">• Retrospective Measures• Concurrent Measures• Utilizing freeze Technique
Implicit Measures	<ul style="list-style-type: none">• Global Measures• External Task Measures• Embedded Task Measures
Subjective Measures	<ul style="list-style-type: none">• Direct Self-Rating• Comparative Self Rating• Observer Rating

Situational Awareness Global Assessment Technique (SAGAT)

SAGAT provides an objective measure of SA based on queries during freezes in simulations (Endsley, 2000). Using SAGAT, a simulation of a scenario is frozen at randomly selected times and the operators are queried as to their perception of the situation at that time (Endsley, 1988). SAGAT is a global measure and queries all SA requirements (perception, interpretation and forecast). SAGAT queries allow for detailed information about subject SA to be collected, on an element-by-element basis that can be evaluated against reality, thus providing an objective assessment of the operator SA. This type of assessment is called direct measure, because it does not rely on judging situation knowledge on the basis of incomplete or subjective assessment (Endsley, 2000). SAGAT requires a comprehensive goal-directed task analysis to be

performed to identify the sequence of tasks, and the list of required variables for them, specifying what questions to be asked. The SAGAT queries are categorized into three groups (perception, interpretation and projection) and will be randomized when evaluating the subject. Figure 3 provides a sample set of queries in the domain of public health surveillance. Some of the major issues that were discussed in the literature are the impact of limitations on working memory (Fracker, 2001), questions acting as cues (Sarter & Woods, 1995), predictive validity of the freeze technique (Pritchett, Hansman, & Johnson, 1995). However, Endsley showed that the accuracy of answers is not affected by time elapsed (working memory) and that task performance was not affected by the duration or frequency of freezes (interruptions) (Endsley, 1995).

Issues of Validity and Reliability of SAGAT

A literature review of the SAGAT technique shows a high degree of validity, sensitivity and reliability in measuring SA (Endsley, 2000). Vidulich found a good level of sensitivity for the SAGAT across a range of studies when a broad range of queries was used (Vidulich, 1992). In a study, fighter pilots who reported the existence of an enemy aircraft using SAGAT were three times more likely to kill the target later in the simulation (Endsley, 1990), showing the criterion validity of the SAGAT task. In another study of air traffic controllers' awareness it was found that the SAGAT measure is sensitive to changes in the task load and to factors that affect operator attention (Endsley, 1995).

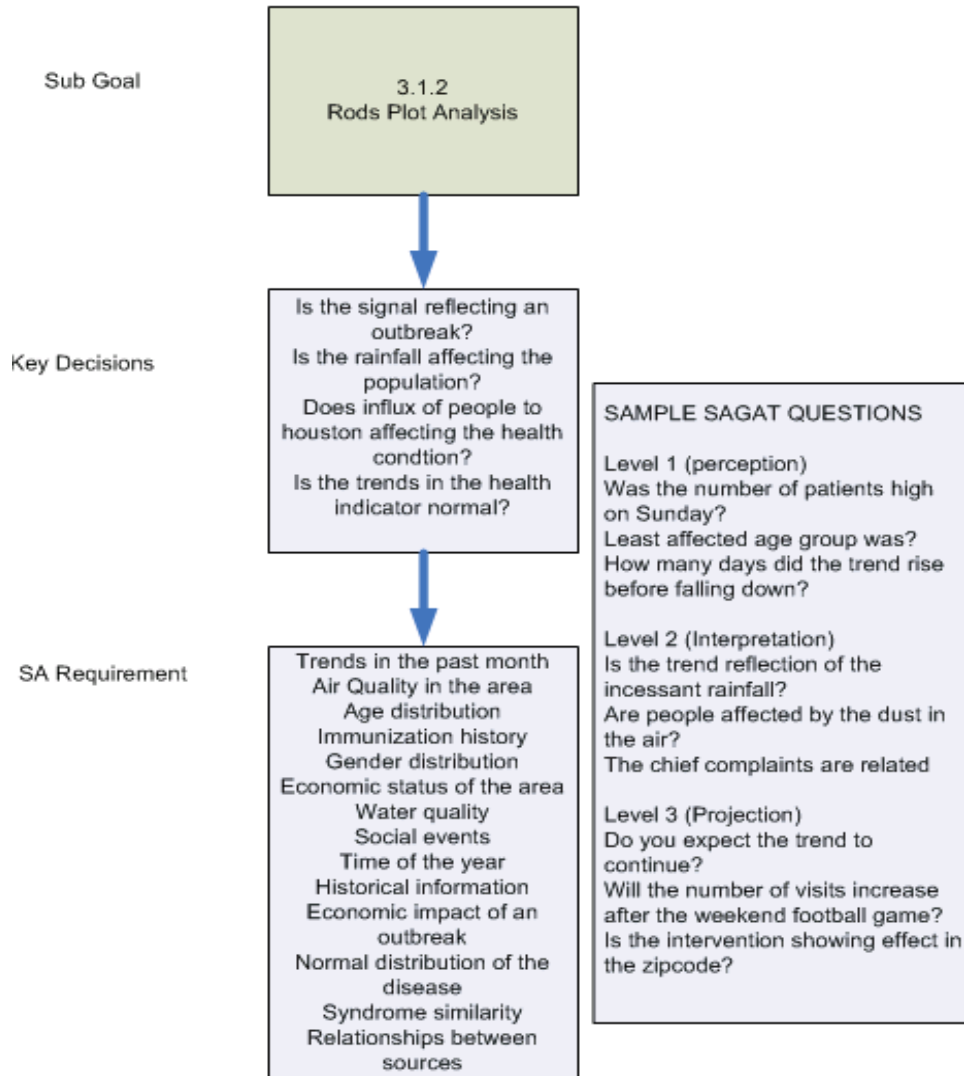


Figure 12 SAGAT Queries for Public Health Surveillance

Measurements in the studies using SAGAT showed a high reliability. Studies in pilot awareness and nuclear plant controllers' awareness support the reliability of the measure (Klein, 2000). Results of studies also show that constructive validity issues, like intrusiveness when performing SAGAT freezes, do not impact performance (Endsley, 1988, 1995, 1990). SAGAT also showed that long-term memory along with working memory was used by expert users.

Chapter Summary

This chapter started by discussing the current status of the dashboard design process. Literature review shows that the current adopted design process is heavily technology-centered and not human-centered. In those studies where human-centered design was adopted, the concept of situational awareness was adopted shallowly. Both UCD and EID contribute to the development of information displays that improve operator insight into decision-making spaces. They share a mutual objective of designing for good decision-making and good human performance in complex environments. Despite the convergence in these objectives, SA has evolved independently in each of these. This is rather surprising, considering that they must, at practical levels, overlap. A review of the empirical EID and SA literature reveals next to no co-occurrences of the terms, in over 112 cited articles, but the term SA is not properly defined nor SA used with reference to the adoption of perceptual elements in the representation. Links between the theoretical foundation of EID and the three levels of SA have not been well understood. In the UCD world, however, the links are somewhat understood. There are some gaps in adopting either approach as a whole towards building a dashboard for complex environment where user needs are constantly changing and there is a need to address anticipated and unanticipated situations.

The key product of the two methods is the design process. In the first one (UCD) this is more goal-oriented whereas in the second (EID) it is more domain-oriented. In order to address challenges with continuously evolving situations, it is critical to map the goals and the domain working knowledge so that a context-sensitive interface is available for users to operate. This study will adopt the approaches from both frameworks to create system interfaces.

The second part of the section introduced the concept of situational awareness, and the challenges in establishing and managing SA over a period of time. It also looked at some of the SA demons and why they should be borne in mind while designing a system. Finally it looked at various subjective and objective methods to assess SA from the users' viewpoint, identifying SAGAT as a potential method for evaluating SA. The following chapter will discuss the new hybrid method that is used to design systems for delivering high SA and a study to test SA using the SAGAT technique.

CHAPTER 3: METHOD FOR REPRESENTING CONTEXTUALIZED INFORMATION

The summary of Chapter 1 specified four desirable characteristics for a public health information system dashboard:

- Meaningful and holistic interpretation of data requires generation of higher level explanations based on knowledge and expertise from multiple principles
- Information systems should consider not only the elements of user functions, tasks and goals but also the user's situational awareness (SA) requirements.
- Information systems should present representations that match the mental schemata of the user
- Systems should have the mechanisms to organize and optimize the interaction and performance based on new situations, user interactions and knowledge of the domain.

This chapter specifies the design process necessary for achieving these characteristics in a dashboard system, employing the following design activities: Goal Directed Task Analysis, Concept Mapping, Classifier, and information presentation interface. Following the design process, it discusses the design for an evaluation study to measure SA levels, the user's response time and the user's confidence level.

Conceptual Framework of Dashboard Design

Chapter 2 discussed the various prior approaches in designing an interface for information systems and also presented the challenges and gaps in these approaches in supporting situational awareness using the technology driven design process. The alternative approach is to use User Centered Design (UCD). As discussed in Chapter 2, UCD focuses on organizing information presented around user's goals and tasks. The UCD process has been applied for years now in various fields in the implementation of interfaces in domains of various

complexities. This section discusses a methodology that adopts multiple UCD methods and applies them to a complex domain problem in public health, so that awareness is provided to the end user on the state of the system. The rationale towards relying on awareness is that, as shown in Figure 10, SA is the key mental construct that drives decision making and performance in complex dynamic systems (Endsley, Bolte, & Jones, 2003). Keeping users on top of the situation and in control is fundamental to good situational awareness. This requires the interface to directly support the cognitive processes of the user. The following section presents the methodology for designing and implementing a dashboard that delivers the information needed by users to achieve all 3 levels of SA.



Figure 13 SA Drives Decision Making and Performance

Design Process

In order to design a system that delivers situational awareness, it is critical to evaluate the information needs of the users in performing their goals. The information needs often found are dynamic, due to the continuously changing goals of the user as various tasks are performed and the information is unveiled. This stresses the importance of determining the information-seeking process and the role of the information needed in achieving each SA level. To delineate the SA requirements and the goals, I have performed a Goal Directed Task Analysis (GDTA) where the users' basic goals, decisions and information needed to support the design process for each SA

level is identified and organized. This step focuses only on the dynamic information requirements relevant in a particular goal but does not explicate the operational and system knowledge (rules and procedures) that the user applies to achieve a goal. To extract this information a second CTA approach, commonly known as context mapping (CM) is applied. Concept Mapping is a CTA toolkit that adopts a knowledge elicitation process, involving a systematic empirical procedure that results in detailed and sometimes formal or even computable representations of knowledge (Crandell, Klein, & Hoffman, 2006). Using CM methods, the domain rules, and operational logics are explicated and represented in a machine-readable format. In the third step of the design process the interface is built based on these goals, and when the data is presented the domain rules and operational rules are applied on the representation flags to offer further guidance for meaningful interpretation and analysis of data. In the following section, I present the GDTA and CM activities for designing a dashboard for a health information exchange system. At the end of this section, I describe the evaluation method and the study instruments developed in this study.

Goal Directed Task Analysis (GDTA)

The GDTA seeks to determine what the information system users would like to know in an ideal situation to meet each of their goals. This information could be made currently available in the same infrastructure or in an external one. One of the common limitations in traditional system implementation is tailoring the design to what information is accessible. The GDTA approach tends to avoid this artificial ceiling effect by explicating and listing all the information needed for the end user to successfully complete the task (Endsley & Kiris, 1994). The GDTA process involves a number of interactions with end users, both formal and semiformal. The end product of the GDTA process is a list of the goals, the decisions made in those goals and the

information requirements and tasks to achieve them. As shown in Figure 14, the GDTA outcomes are organized into charts depicting the hierarchy of goals, sub-goals, decisions relevant to sub-goals and the information required for making each decision.

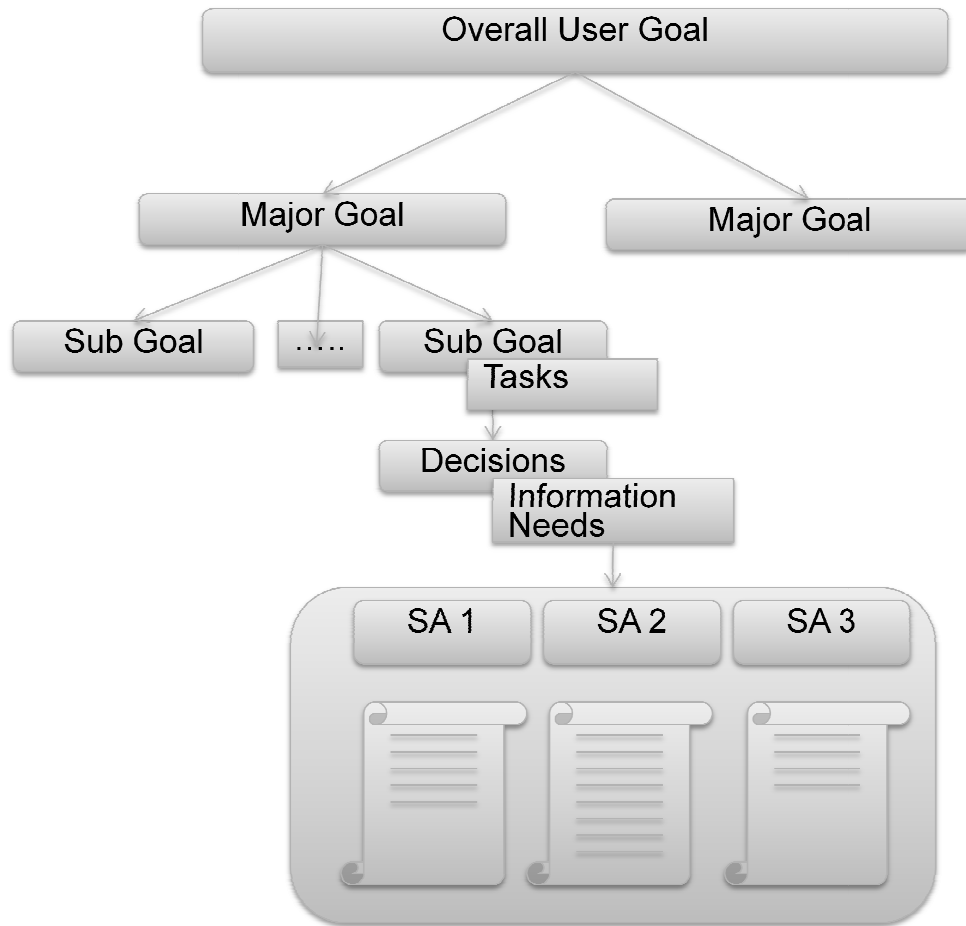


Figure 14 Goal - Decision - SA Requirements Hierarchy Representation

The key towards successfully performing a GDTA is to clearly capture the goals and information needs. The goal hierarchy is truly the foundation of the GDTA and hence, to be clear, this study adopted the following principles in identifying and differentiating goals, tasks, decisions and information needs.

- Goals: Goals are higher order objectives essential to successful job performance (e.g. Communicate updates to partners effectively).

- Task: These are activities that operators must physically accomplish (e.g. pick up phone and call a reporter)
- Cognitive Demands: Items requiring expenditure of higher order cognitive resources. (E.g. determine the effects of a system failure at the reporting facility).
- Decisions: Questions posed by the user to effectively meet each goal in a goal hierarchy.

The next step performed in the GDTA is to become familiar with the project scope and domain. For this, I spent time reviewing all the business requirements documents, software requirements specification and software design documents, so as to understand the nature of the job of the end user and also to help me guide the interviewing process. The next step performed in the GDTA included conducting interviews with subject matter experts to obtain the list of goals and the information needed by the SME in performing the goals. This is discussed in the following section.

Initial Interview

I recruited 7 subject matter experts (SME) who met the inclusion criteria established for the study. The SMEs were invited for a kick-off meeting over the phone. In the meeting, the purpose of the research work and the expectations were established. As a follow-up to the initial call, individual meetings were held in a semi-formal setting. Typical interviews included a review of the research scope, purpose, and intent of data collection and a verification of each interviewee's professional experience; following this, the interviewee's goals in the project were discussed. Due to the number of interacting domains there were many major goals identified during the process of inquiry. The interview process allowed the identification of a preliminary goal structure in which the overall goal was broken down into sub goals with a hierarchy spanning multiple levels. This required the use of a page as a parking lot for reminders of the

need to analyze the topics after completion of the first. The preliminary goal structure is presented in the next section.

Preliminary Goal Structure

The usual outcome of the interview was notes with highlighted goals and potential decisions that were part of the goals. Figure 15 shows how the goals were categorized into specific areas based on reorganizing the notes from interviews. At this time the goals were not sequenced, but just grouped into specific domain areas e.g. monitoring, communicating etc. The preliminary goal structure helped to channel future interviews. The key challenge in this GDTA process is perfecting the goal structure. There was an inordinate amount of time spent in organizing the information from 7 SME sources, for whom terminology differences were significant.

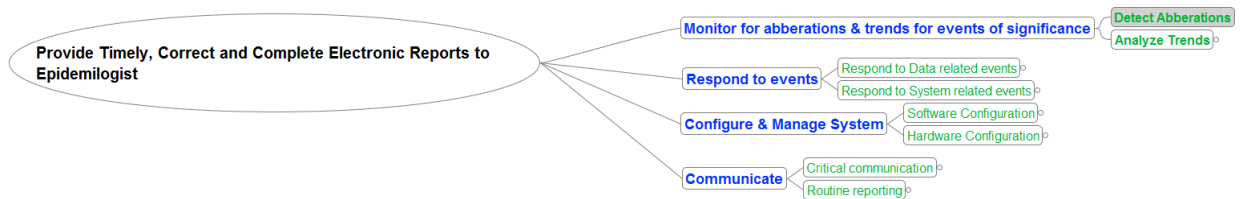


Figure 15: Preliminary Goal Hierarchy

The preliminary goal hierarchy was shared again with the 7 SMEs via email, and input was obtained on their interpretation to assure correctness. A follow-up individual call was scheduled with each of the SMEs, based on the comments received. The process involved clarification of comments and also validating changes made to the goal structure. Table 2 provides an expanded view of some of the goals under the ‘monitor’ goal. The next stage included multiple rounds of refinement of the goal hierarchy. The following section will discuss the final goal structure that was agreed and some of the key considerations made in this study while defining the goal structure

Preliminary Goals	
Overall Goal: Provide Timely, Correct and Complete Electronic Reports to Epidemiologists	
1.	Monitor for aberrations & trends for events of significance
1.1.	Detect Aberrations
1.1.1.	Detect Aberrations in System Performance
1.1.1.1.	Monitor the HW operations of the system
1.1.1.1.1.	Monitor HW Status
1.1.1.1.2.	Monitor memory usage
1.1.1.1.3.	Monitor disk space
1.1.1.1.4.	Monitor Network Connection
1.1.1.2.	Monitor the SW operations of the system
1.1.1.2.1.	Monitor Orion Rhapsody status
1.1.1.2.2.	Monitor NEDSS Webservices status
1.1.1.3.	Monitor the communications with external systems
1.1.1.3.1.	Monitor the email server connection
1.1.1.3.2.	Monitor the FTP server status
1.1.1.3.3.	Monitor the HTTPS server status
1.1.1.3.4.	Monitor the PHIN MS server status
1.1.2.	Detect Aberrations in reporting of messages
1.1.2.1.	Monitor the volume of messages received
1.1.2.2.	Monitor the volume of messages distributed
1.1.3.	Detect Aberrations in message transformation process
1.1.3.1.	Monitor the subscription web service health status
1.1.3.2.	Monitor message content validation checks

Table 2 Preliminary Goal Hierarchy

Final Goal Structure

After the preliminary rounds of interviews the SME's were queried only on the higher-level requirements. Once the goals were realigned, as shown in Figure 13, the next step involved narrowing down to the information requirements level. Furthermore, to achieve specific goals certain decisions would be needed. To make these, the user would need to seek and utilize certain information bits. This step in GDTA aims to explicate the decisions made by the users

and the information queried by the user to achieve a goal. In the third round of individual interviews with the SMEs, each sub-goal was used as the starting point of discussion and the decisions needed to effectively meet the goal identified. Some goals have more than a few decisions. For the first pass, the decisions were listed individually to see whether they had similar or different requirements. For each decision, the subsequent SA requirements were collected and presented as a list of variables to be made available. At this point, no consideration is provided as to whether the variable is available in the current system or even exists in a different system. This list considers an ideal situation and hence usually lot of thought needs to be gone into, carefully stating whether there is anything missing that might be needed to validate a theory or an assumption. Further clarity needs to be provided in this list. Figure 14 provides a view of the decision and information requirements in one of the sub-goals. Typical list of requirements are organized at level 1 at the lowest, with level 2 and level 3 staked below. Level 1 variables were mainly focused on delivering information availability; level 2 focused on interpreting the current situation and used level 1 to perceive the information on the screen. Level 3 information requirements are staked below Level 2 and Level 1 variables and are generally used to project the impact of the situation into the future.

It was typically found that there were many situations where a sub-goal was part of more than one goal; in such situations, the sub-goals were called out from all future locations. This cross-referencing drastically reduced redundancies in the goal hierarchy. The final version of the goals hierarchy was shared repeatedly with the SMEs and input was constantly received and incorporated into the goals lists. Commonalities became evident and the redundancies of tasks were reduced many times. The complete list of the goals hierarchy for the monitor tasks that was used in this study is listed in Appendix A. The other major goals including “respond to events”

and “communication” were not assessed in depth as the study goal was limited to the signal characterization task and did not lead up to performance activities and decision-making.

GDTA Summary

The process of GDTA is a complex and taxing activity, but the outcome of the work is extremely critical and lays out the core data elements needed for successfully achieving the SA levels needed for the goals. The GDTA statistics for this study are shown in figure 15. The GDTA process in this study lasted over 8 months due to the varying availability of the SMEs and the complexity of the domain. The GDTA hierarchy constructed had repeated refinements centered over the definition of the goal and whether a certain activity was a goal or a task. During the process, confusion persisted over whether specific decision points were applicable to the situation. Differences among the SMEs on the role of specific variables and their validity to be considered for interpretation was a complex problem, for example, socio-economic distribution and its role in reporting. Some SMEs referred to the fact that certain regions due to socio economic strata do not get a lot of medical attention until there is an outbreak, but some SMEs did not want to even consider this as an variable as it was misleading and not scientifically proven. Other situations for which the GDTA approach was complex were the roles of some of the SA variables and their classification as a Level 2 or Level 3. This caused some confusion when the sub-goals were used in the call outs.

- | | |
|--------------------------|---------------------------------|
| ▪ Sub- Goals Level 1: 3 | ▪ Call Outs : 49 |
| ▪ Sub-Goals Level 2 : 6 | ▪ Unique Decisions : 67 |
| ▪ Sub- Goals Level 3: 23 | ▪ Unique Information needs: 169 |
| ▪ Sub- Goals Level 4: 79 | |

Figure 16 GDTA Summary Status

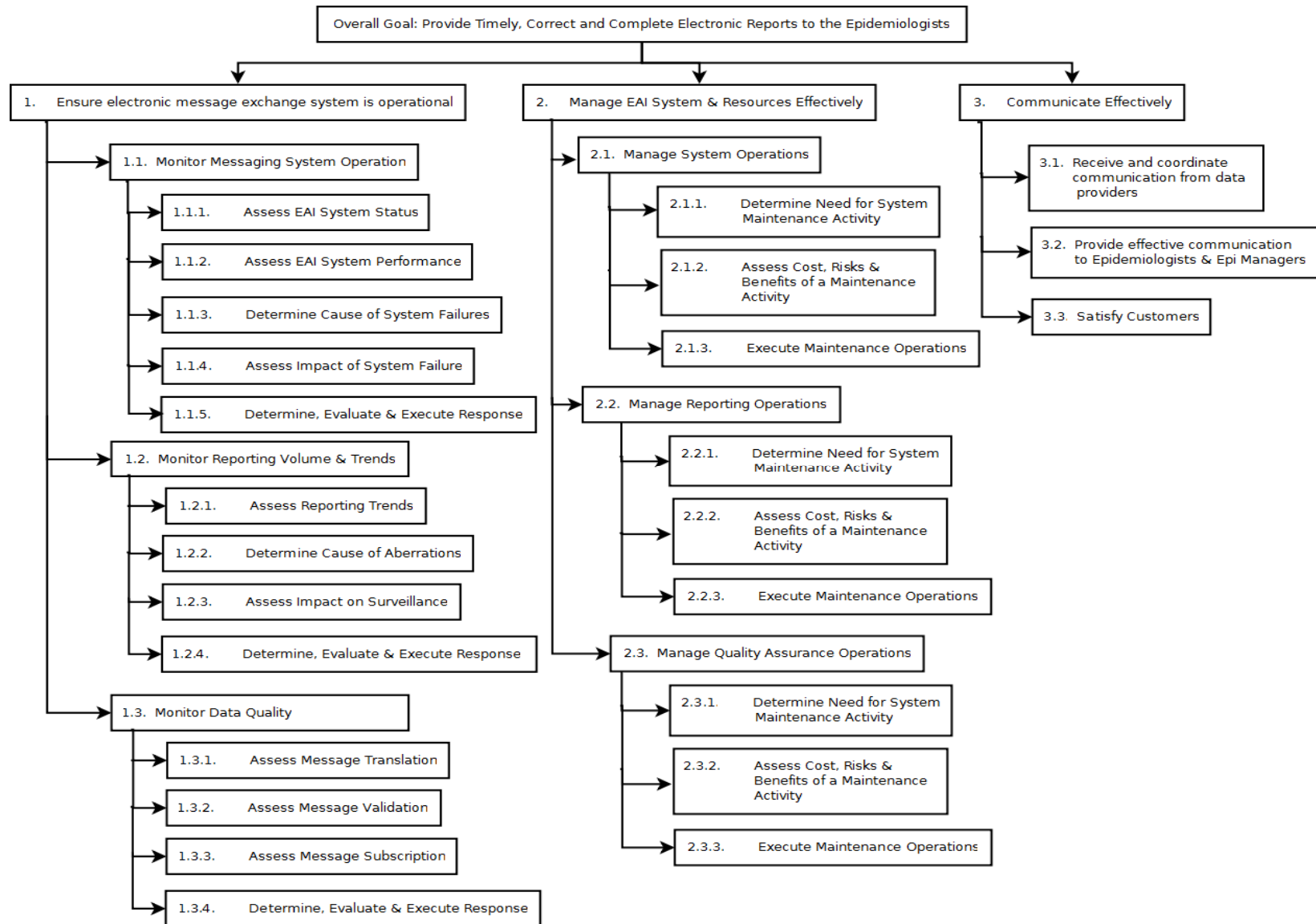


Figure 17: Final Goal Hierarchy

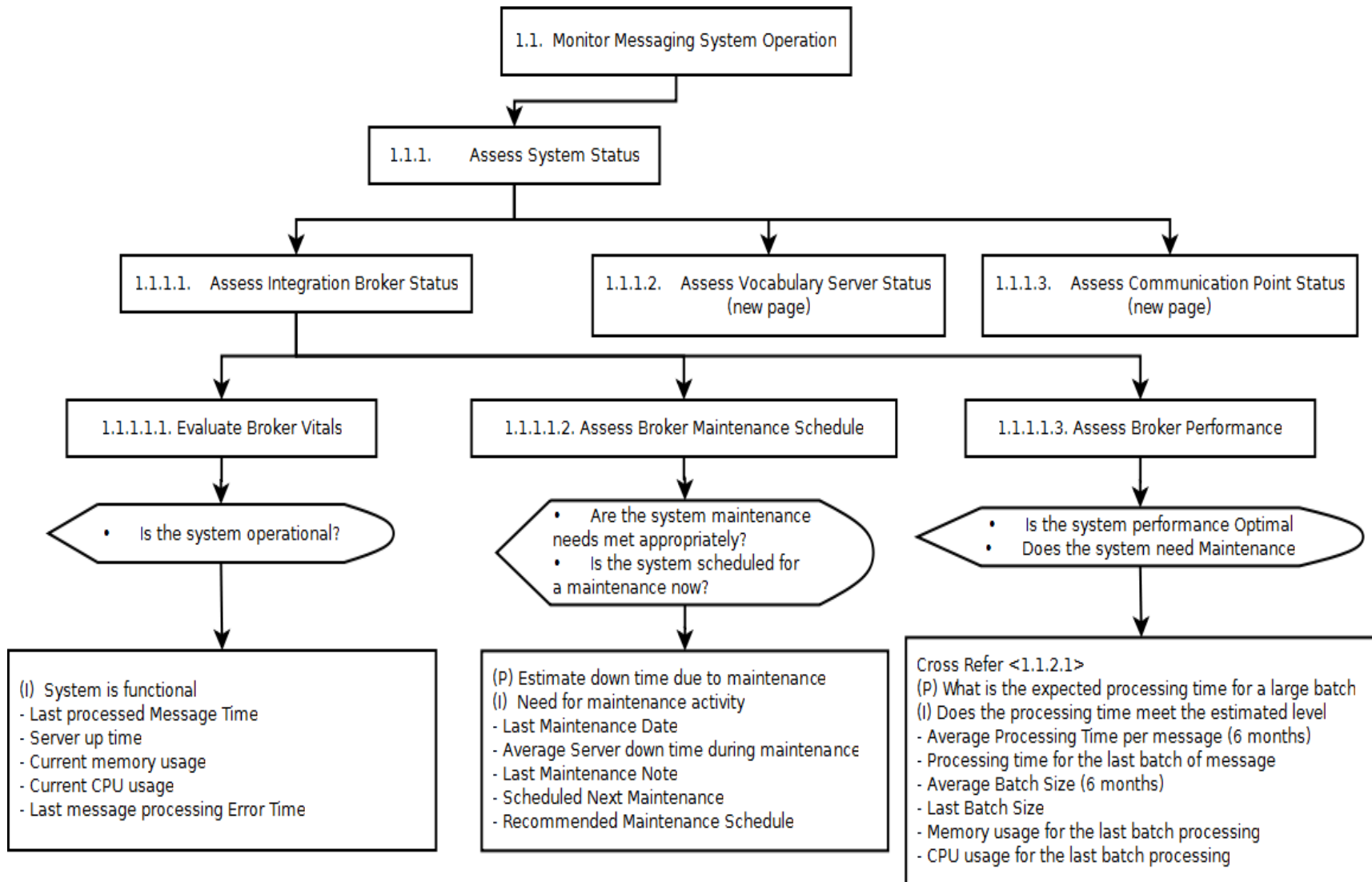


Figure 18: Final GDTA Hierarchy

To validate the GDTA, 2 SMEs who did not participate in the development of the hierarchy were asked to review a hardcopy of the GDTA, for missing information or errors in the representation. No critical errors were identified. Some suggestions on organization were provided and two changes were incorporated into the final goals hierarchy.

Concept Maps for Knowledge Elicitation and Representation

With the completion of the GDTA process, it was clearly evident that the needs of the end user in performing the tasks are often quite long and most of the interpretation and linkages are left to the end user to do. Simply incorporating GDTA into the design leaves the system as the syntactic representation platform with no semantics applied to understand the role, value and context of the data. The Concept Mapping (CM) process is added to this design methodology to allow designers to extract the semantics of the data for use during system design to help users interpret information meaningfully.

The concept map as a knowledge elicitation tool has been applied in many fields to explicate the user's understanding of the conceptual, methodological and multi-domain linkages that exists in a domain (Hoffman, 1998). This process involves a series of one-on-one interviews with end users. In order to perform CM, it is critical to be familiar with the domain. I started off reviewing the annual ELR and MMWR reports to understand the trends and specific reports in which matters of environmental, social and organization had been cited. These preparatory activities aided in triggering the integration system managers to recall past cases in which one or more of the concepts were salient. In the following section, I discuss the methods adopted and how the information was collected and documented in ways that could be used meaningfully to improve the end user's awareness.

Knowledge Acquisition Process

The concept mapping process involves a series of interviews with subject matter experts. Concept map interviewing requires two researchers to participate, so that they can document the discussion in a human readable format. The main researcher began to facilitate the discussion with the SME in the form of probe questions, while the other one captured the discussion and documented the key concepts as a graph. The critical step in a successful concept mapping process is the selection of the domain subset and focus that are directly relevant to the user goals. The GDTA sub goals were used as focus questions in this work. E.g. “why do message brokering systems fail?” These questions are usually the broadest, most overarching and general but still relevant to the user goals. They usually triggered a train of thought that was followed up by specific queries like “Have you had to deal with such failures? And what caused them?” The outcome of this interview led to additions to the parking lots. Below is an example of a discussion about the reporting trends among labs in the region.

What are some reason for varying trends among labs reporting in your jurisdiction?

“The focus would specifically be on volume reported by State Labs and commercial labs in the region. Due to seasonal variations this can become a complex interpretation, for example the State Lab is more likely to have a seasonal bump in winter due to Influenza whereas local hospitals and commercial labs are more likely to have seasonal summer bumps due to people being outdoors and in swimming pools more so things like Giardia, Cryptosporidium, etc. come into play. Of course this can be skewed when outbreaks like West Nile come around and then the State Lab gets both a summer and winter seasonal bump.”

Figure 19 Concept Map Interview Notes

The above elicitation by the SME was captured into a parking lot (Figure 20) where the focus questions and all concepts related to the questions were grouped. The concepts were moved around so that relevant concepts were grouped. In a similar study in some domains this parking lot grouping stage was skipped and the concepts were directly linked, as shown in Figure

21. The linking process involved identifying linking words that are meaningful and represent the relationship between the two concepts. Due to the nature of the domain, linking words fell under the categories of: classification (is an example of), nominal (is known as), property (consist of), explanatory (is the cause of), methods (is done by), dependency (requires) and probabilistic (is less likely). Figure 21 shows a linked concept map. Similarly, based on SME interviews, over 70 trend-specific concepts were captured and organized. Figure 22 lists the grouping of all concepts gathered. In the next section I will describe the steps involved in representing the information in a machine processable format.

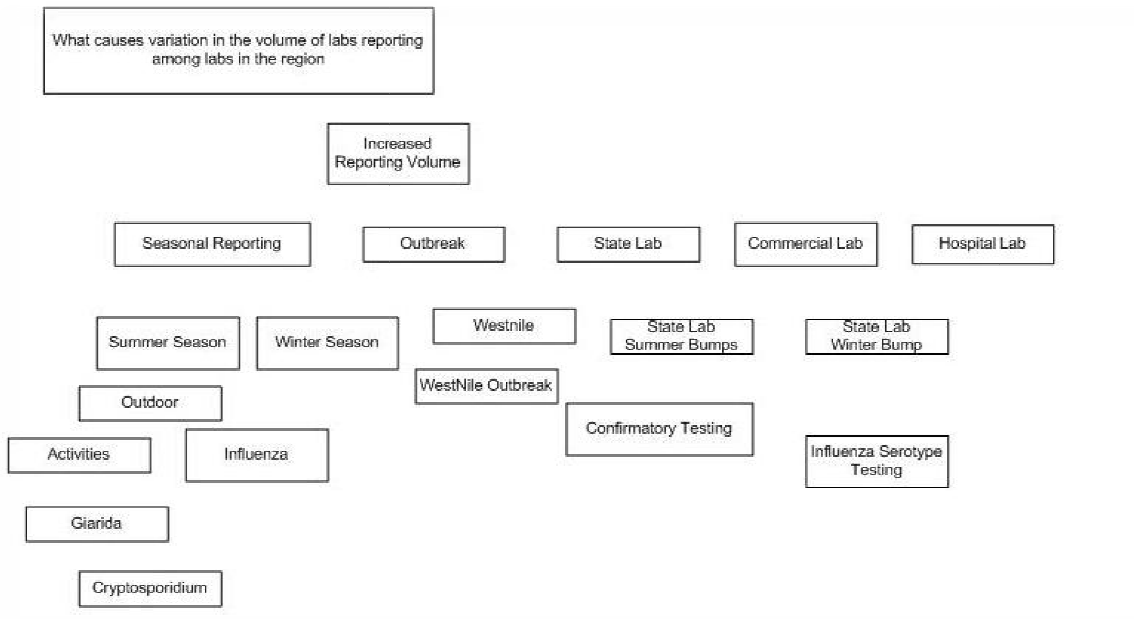


Figure 20 Parking Lot of Lab Reporting Volume

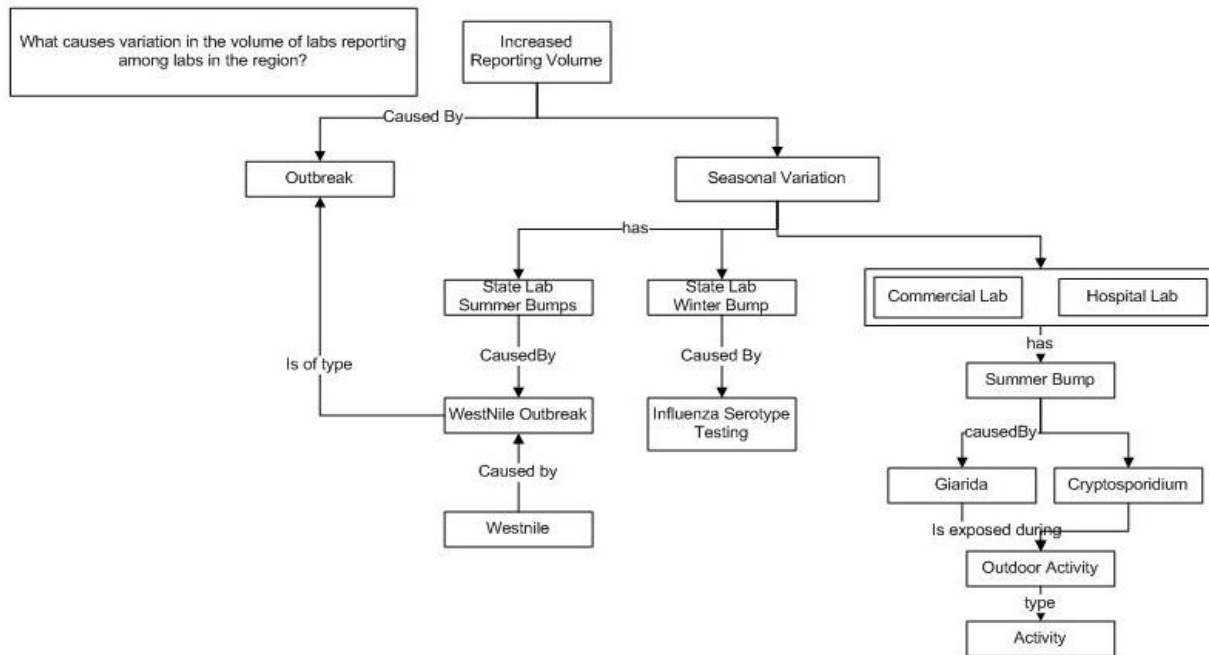


Figure 21 Concept Map for Lab Reporting Volume Trend

Knowledge Representation

The concept maps that were created based on the SME input were validated through multiple rounds of peer review. Once the concept map was finalized, the next step was to represent the map in machine-interpretable format that could be used to aid in the signal characterization task while monitoring the message trends. Choosing a machine interpretable representation is a highly complex task. Current representation standards include the W3C approved formats like OWL, RDF, and XML. Other representations include adoption of rule languages like Drools. For this study due to nature of the domain and the study scope, Drools rules were adopted. An effort to represent the domain using OWL ontology was also attempted, see appendix C. In the following section, I will describe how the Drools rule language was used in representing some of the domain knowledge that was identified during concept mapping. Table 3 presents three scenarios that need to be represented, so that, when a certain situation

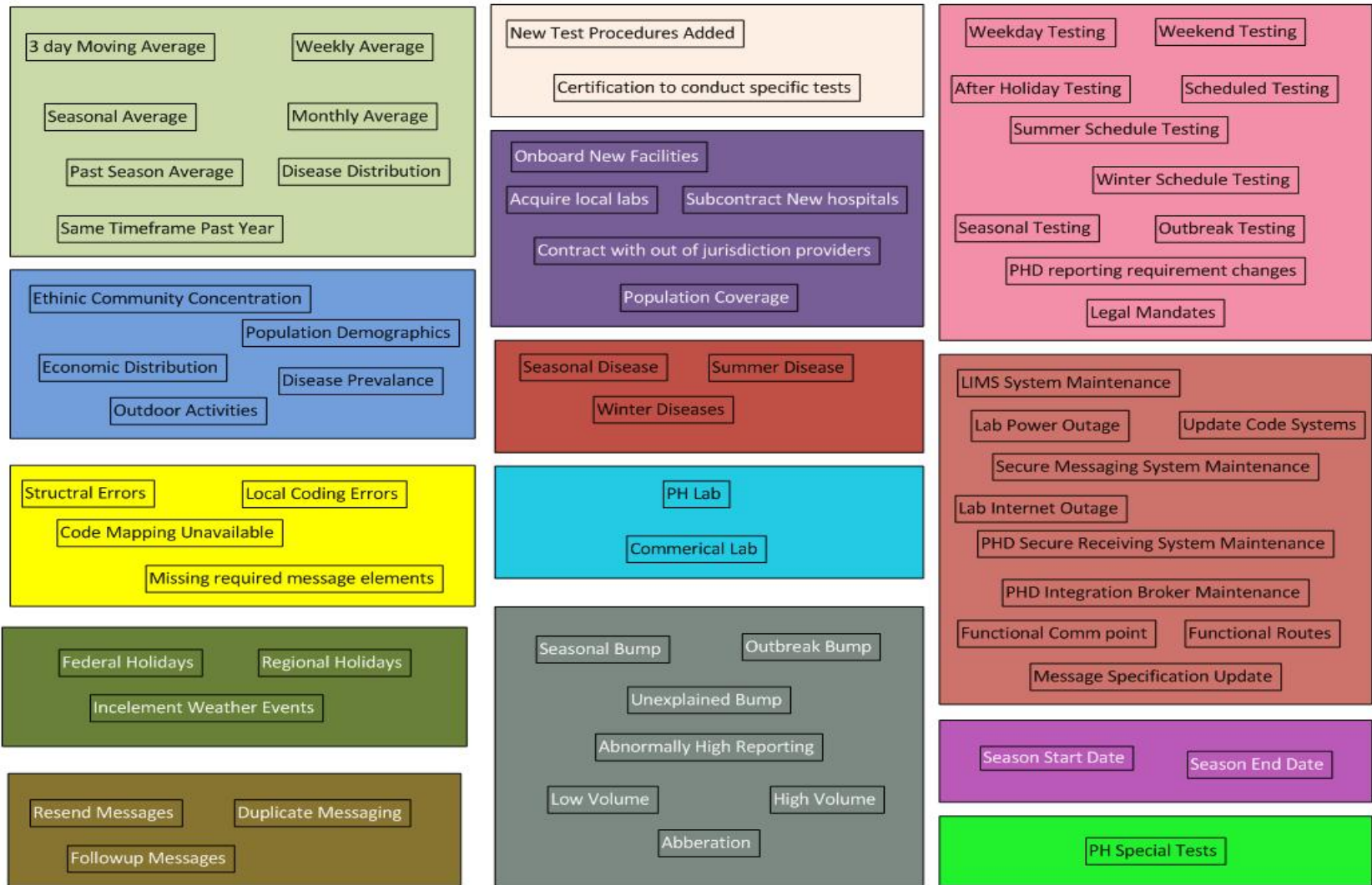


Figure 22 Parking Lot of Message Trend Concepts

<p>Scenario 1: Alert -Holiday Effect</p> <ul style="list-style-type: none"> • If volume > 2011, 3 day moving Avg • If prev day == Sunday or Prev day == Fed Holiday • Alert: “Holiday Effect” <p>Scenario 2: Alert system failure</p> <ul style="list-style-type: none"> • Lab primary FTP system • FTP Comm PT current Status != operational • FTP Comm PT 24 Hr Status != operational • If volume received == 0 or volume < Current Season week Avg • Raise Alarm: “Comm pt failure Effect” <p>Scenario 3: High Volume Monday</p> <ul style="list-style-type: none"> • If volume > 2011 weekly avg • If day == Wed • If season == summer • Raise Alarm : “ Lab PCR test day”

Table 3 Scenarios Identified using Concept Mapping

arises, they will help the user understand the situation better by providing contextual information. In scenario 1, the volume can be represented using the Drools rule language as shown in Table 4.

```

rule "Rule_HolidayEffect"
dialect "mvel"
    when
        m: LabMessage ( res : result, vol : volume, date : currDate, comm : labCommSystem, season : season)
        eval ( vol > MessageTrendConstants.WEEKLY_AVERAGE_2011)
        eval ((DroolsMessageTrendHelper.getPrevDay(date) == Calendar.SUNDAY)|| (
DroolsMessageTrendHelper.isPrevDayFedHoliday (date)))
    then
        m.setResult(MessageTrendConstants.FALSE);
        System.out.println("Holiday Effect");
    End

```

Table 4 Drools Representation of a Message Trend Scenario

Concept Mapping Activity Summary

The concept mapping activity led to the extraction and representation of the domain and operational knowledge in a clear and concise manner. The concept mapping process was comprehensive as it related to each of the focus questions. In this study, during that process some important concepts had been overlooked due to views on their relevance to the scope of the study. The final outcome of the concept map, however, allowed the researcher to get crystal clarity on the operational knowledge required to assess the events in the domain. The concept mapping process itself was not a simple task, requiring numerous discussions with the subject matter experts. Especially in this domain, where theories and logical understanding were widely different from jurisdiction to jurisdiction and from person to person, it was a very complex activity. Some of the challenges faced during this interview process were

- a. SME's significantly different language in referring to the same events (e.g. seasonal bump vs. seasonal aberration)
- b. SME's perception about the impact of certain elements during the monitoring task (e.g. use of weather data)
- c. SME's retraction and negation of certain elements in different scenarios (e.g. use of air quality data while assessing alerts).

Fourteen knowledge elicitation sessions among three subject matter experts yielded seven concept maps and about seventy digital resources. The concept maps were then coded as Drools rules as explained in the knowledge representation section, and the rules used to identify what potential contextual information could be added to the user interface to guide the dashboard system user while performing the signal characterization activity. The model can be expanded or removed with no impact to the actual system, as each rule is fired independently. In the

following section I will present the principles adopted in taking the outcomes of the GDTA and concept mapping and translating them into representation.

Dashboard Design Principles

The goal of the manager of an electronic message broker engine using the dashboard is to acquire sufficient information about the status of the system, so that it can be used to determine to a certain degree the potential reasons for the current state of the system and its reporting. The way information is presented to the user greatly influences the SA gained by the user. Hence the information presentation should aim at allowing the user to gain information needed as quickly as possible without undue cognitive effort. This section presents some of the key design principles adopted to enhance operators' SA while using the system, in this case to implement a dashboard for electronic disease surveillance lab reporting. The design is also compared with a current dashboard built using the traditional waterfall approach. The results of the study are discussed in detail in Chapters 4 and 5.

In this study the following design principles are adopted in building a dashboard to deliver SA to end-users. Each of the principles is discussed in detail in Chapter 5.

- Design around user goals
- Organize information available to support Levels 2 & 3
- Make information explicit
- Reduce data overload
- Integrate information
- Acknowledge missing information
- Include domain knowledge in the representation
- Present critical information needed to trigger mental schemata

- Consider semantics of data to support SA
- Provide SA support and avoid making decisions
- Support Global SA at all times
- System functionality should address all user goals
- Adopt consistent representation across all goals
- Allow systematic querying of information
- Minimize the levels in logical branches
- Implement an interface requiring less cognitive effort
- Support uncertainty and higher levels of complexity
- Accommodate to changes in the domain knowledge
- Use the data salience property with caution
- Deliver consistent representation across the system.

Design Principles – from Theory to Practice

To more clearly articulate the design implementation phase, I compared the goals identified during the GDTA with the current dashboard’s capacity to meet the goal requirements in terms of SA levels. Based on gap analysis, information availability is assessed and then a new interface was designed by applying the above principles. The NEDSS Message Subscription Service System Dashboard in combination with the Orion Rhapsody Admin Dashboard is widely used in this domain by administrators. A sample of the comparison of the MSS dashboard with the GDTA goal and information requirements is shown in Table 5. Cells filled in green indicate that the information was available in the system, yellow indicates that partial information was available, whereas blue indicates that the information was inferred based on other information in

the system. Red indicates that the information was not available from any of the systems available to the user.

Closely reviewing the dashboard, it was evident that a fair amount of information queried by the integration broker manager was present in one form or other. Very little information was unavailable. However, the information was distributed across various sections or even systems. Table 5 shows that all the information needed by the user was distributed across multiple systems.

The first step in the new design is to integrate information from various systems into a single frame organized by user goal. In cases where information was inferred, it was critical to identify the specific information needed for interpretation. In some cases there was a need to review historical and comparative data, requiring multiple steps to search the specific data range. It was essential to ensure that the number of tasks to search information is limited. Predefined concepts of past season, last week, same week last year were included and made available for the user to select, reducing memory overload in trying to determine or remember information, e.g. season start date. Some information needed to be more explicit to support SA levels 2 and 3, e.g. percentage of use of a comm pt by a certain source. This is critical to assess the impact of a situation on overall goals of the user. Another critical gap identified was the lack of information regarding global SA. At any given point of time during information retrieval, the global SA elements, system status and reporting status are not always all available. Furthermore, the system does not consider dynamic events to help the user to alter goals. Some past or current events are buried under a lot of information that is critical for achieving the user goals.



Goals-Tasks-Decisions-Information	MSS	Rhap	Email	Feeds	Other	Not Available
1. Monitor for aberrations & trends for events of significance						
1.1. Detect Aberrations						
1.1.1 Detect Aberrations in System						
Performance						
1.1.1.1. Monitor the HW operations of the system						
Status						
1.1.1.1.1. Monitor HW						
<i>Is the Server on?</i>						
Server Name	Green	Green				
Operating System					Yellow	
IP Address	Green	Green				
Server status	Yellow	Yellow				
Last transaction time	Yellow	Green				
Server uptime		Green				
1.1.2. Detect Aberrations in reporting of messages						
1.1.2.1. Assess Reporting Trend by Source						
<i>Is there an abnormal increase in message volume by source?</i>						
<i>Is there an abnormality?</i>						
<i>What is the impact of this abnormality?</i>						
<i>What caused the abnormality?</i>						
Message count by source	Green	Blue				
Average count of messages by source over the month	Green	Blue				
Trend compared to last season						Red
Current status of the communication path used by the source		Green				
Interfacing systems status	Blue	Green				
Message processing system	Green					
status						
Trends of diseases reported by this source	Green					
(outbreaks)						
Impacting PH events			Yellow		Yellow	
Impacting system events at source & destination				Yellow		
Population demographics of the region					Green	
Source Market Share					Blue	
Last processed message time		Green				
Last batch receive time		Green				
Last batch volume		Green				
Weekend average		Blue				

Table 5: Comparison of Information available among system components

Unless the information is available in a timely fashion, the user is more than likely to miss the information and thus the SA. Lack of proper interface components and linkages, e.g. breadcrumbs; and overcrowding of graphical elements, e.g. dynamic charts, frames, windows; overuse or incorrect use of color coding, e.g. alert vs. warning proved too taxing on the end user.

A prototype of a new interface was developed to address these shortcomings. In the new interface, information was oriented around goals. The system supported global SA by delivering appropriate event-based information. All information was based on the outcome of GDTA. Perception was improved by delivering mission-critical elements with appropriate salience (e.g. color codes and trend markers). Interpretation was improved by bringing relevant goal-specific information closer, for comparison and interpretation. Projection was supported by providing information on current and past trends of the domain. The new interface had the following specific design that promoted SA:

1. The dashboard has a status section that delivers the global system status at all time. The status map was broken into sub-systems, namely hardware systems, sources, diseases and error status. The status section delivered a simple graphical cue which indicated whether the system was operational, under scheduled maintenance, or not operational. The source status was indicated using a graphical cue that showed whether the source reports are received as expected, or not operational. Similarly the disease trend was indicated using a trend marker that showed whether there was an increased or decreased volume in reporting. If all sources were operational and if the disease count was zero then the disease list showed green, otherwise a red. The error status indicated if there were any errors in the message processing.

System	Height
NEDSS MSS	✓
Rhapsody	
PHIN VADS	
Webserver	✓
PHIN MS	✓

Source	Status
LabCorp	▲ 69%
Quest	▲ 2%
Mayo	0%
SPHL	▼ -87%
University	▲ 22%

(a) System Status

(b) Source Status

Figure 23: System and Source Status in the new Interface

2. A breadcrumb approach was chosen to allow dissection of the information presented to the end user, also allowing the system to have an understanding of what goal the user is pursuing. E.g. if the user chose ALL LABS>> Tuberculosis then the goal of the user is to understand the trend of TB reporting in the jurisdictions. If the user chose LabCorp>> All Diseases then the goal is to see the trends among LabCorp messages, helping to choose the right goal and the information needed for that goal.
3. The status cues allow user to drill for more specific information. The windowed environment allowed display to be organized around goal attainment. This also allowed switching between goals. The cross reference between goals identified in the GDTA allowed the implementation of links between windows relevant to a specific goal.

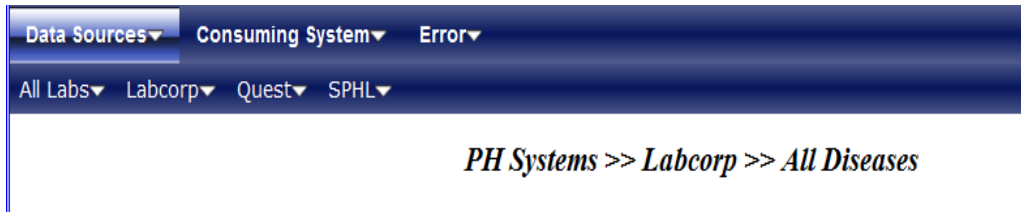


Figure 24: Breadcrumbs to allow awareness

4. Level 3 SA is supported by the presentation of historic trends and prior behavior of the system on specific events. These parameters help in the comprehension of domain specific trends that are critical in estimating the future and describe the past. The trends are shown in Figures 25 and 26.

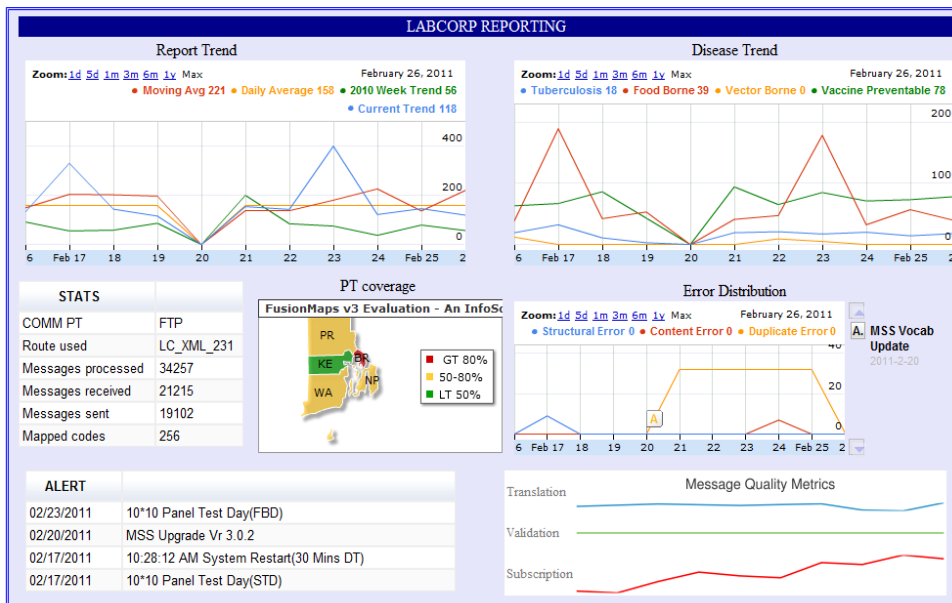


Figure 25: Windowed environment

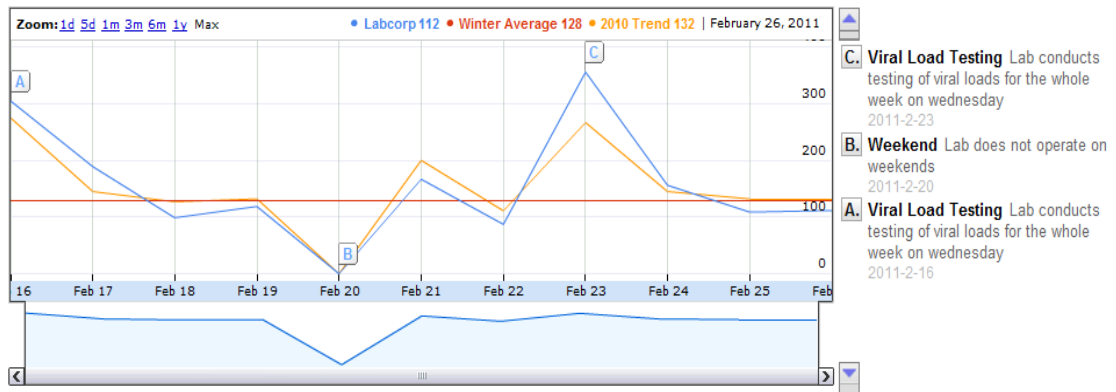


Figure 26: Annotated Graphs providing historical information

- The use of concept map information, which triggers rules based on the data, is used to annotate the trends and provide specific alerts that are caused by domain specific events. These annotations are again goal driven which means only those annotations that are specific to the particular situation are presented, e.g. an FTP failure alert is applied to the trend only if the user is reviewing trends of reporting from the source that uses the FTP comm pt. The annotation is not available while reviewing a source that does not use the FTP comm pt.

```

rule "Rule_SystemFailure"
dialect "mvel"

    when

        m: LabMessage ( res : result, vol : volume, date : currDate, comm : labCommSystem, season : season)

        eval ( comm == MessageTrendConstants.LAB_COMMUNICATION_SYSTEM_FTP &&
!DroolsMessageTrendHelper.checkFtpComm())

        eval ( vol ==0 || vol < MessageTrendConstants.CURRENT_SEASON_WEEKLY_AVERAGE)

    then

        m.setResult(MessageTrendConstants.FALSE);

        System.out.println("Comm pt failure Effect");

End

```

Figure 27: Annotation Identified using Drool Rule

6. The alerts are just suggestions for SA and are not decision points. If alerts are applicable to a specific domain but inapplicable to the domain trend, an alert box is available to display alerts that are relevant to the goal.
7. The interface design allowed flexibility in conducting repetitive tasks by reducing the number of steps. The system also allowed the users to expand beyond the goals to retrieve information not fitting into the domain knowledge that has been explained using CTA process.
8. The system required minimal training and low memory load as relevant data are grouped.

By adopting the above approach the system design addresses some of the core SA demons discussed in Chapter 2. By providing indicators with salience for system, reporting, disease and quality status, the user is always aware of the overall system, avoiding the risk of the user getting into attention narrowing. By providing global SA the system is able to trigger appropriate schemata that increase the probability that system users will redirect goals and priorities relevant to the situation. By presenting information appropriate for the goals, the risk of relying on human memory to perform correctly is reduced because the SA requirements directly address the user requirement. By providing level 2 support directly, the design directly supports comprehension, relieving users' mental activities like remembering. This in turn reduces the impact of stressors in the external environment. Stressors are factors that cannot be directly controlled by the interface, but by reducing the cognitive and physical requirements, the system provides an ambience that minimizes their impact. Having adopting the GDTA, the system clearly reduces the risk of data overload. The domain and operational knowledge further helps by placing salience appropriately. One of the major risks in this approach is triggering the incorrect

mental model. The system design focusing on mapping user goals to system functions helps to reduce the risk of users applying incorrect mental models. By ensuring that the presentation of information stays consistent across the system, the user will run less risk of failing to identify the critical cues essential to map to a mental schema. These design principles directly address the key demons identified in the SA literature. The empirical study will demonstrate the outcomes of the design and will be discussed in the next chapter.

Summary of the Design

The new interface illustrates several of the design principles presented in this chapter. All information is organized around goals. Display of specific goal-related information is in separate displays, e.g. System Status, Disease Trends, and Source Trends. These displays carry goal-specific information in an integrated fashion. The information is geared to support all 3 SA levels. A global view of the system is available at all times. The main dashboard offers system, source, quality and program vitals, which determine the priority goals. Salience and explicit importance is provided to the key vitals. The breadcrumbs and links approach allows easy switching from one goal to another. The annotated timeline allows users to look at events that are relevant and contextual to understanding the trends. The timelines also include relevant historical and performance data for project support. Auto filtering of data has been limited and the back link allowed users to switch back to previous goals or to the home at any given point. The 19 core principles were applied in the final design of the interface, leading to an implementation that greatly reduces workload and improves SA.

Evaluation Study

In the specific aim 2, I propose to experimentally evaluate the situation awareness of expert users in the signal characterization task by conducting studies to empirically measure the

expert's situation awareness when using the implemented prototype and the current system, using situational awareness measurement strategies in a laboratory setting. Here I present the implementation of a direct measurement technique using the Situation Awareness Global Assessment Technique (SAGAT). A study design to measure the SA using SAGAT implementation and a measurement instrument is presented in this section.

Evaluation Instrument

My general goal in this evaluation was to determine how useful the system is at helping users achieve SA while performing the signal characterization task. Specifically, I tested the claim that users had higher levels of SA when utilizing the interface developed by applying contextualized design, as described earlier in this chapter. The users were measured for timeliness in responding to questions and also for level of confidence in responding to the questions. The questions are based on the GDTA that was conducted and are goal oriented. The questions could be classified as perception, interpretation or forecasting questions. With the help of two SMEs, a set of 3 unique scenarios was developed. Table 6 describes the three scenarios that were used in the study. Each scenario was checked by an external SME for validity. With input from the SME, sample data sets were created for both the old and the new interfaces. A set of 30 questions per scenario was created, based on information requirements identified during GDTA (10 for each level of SA).

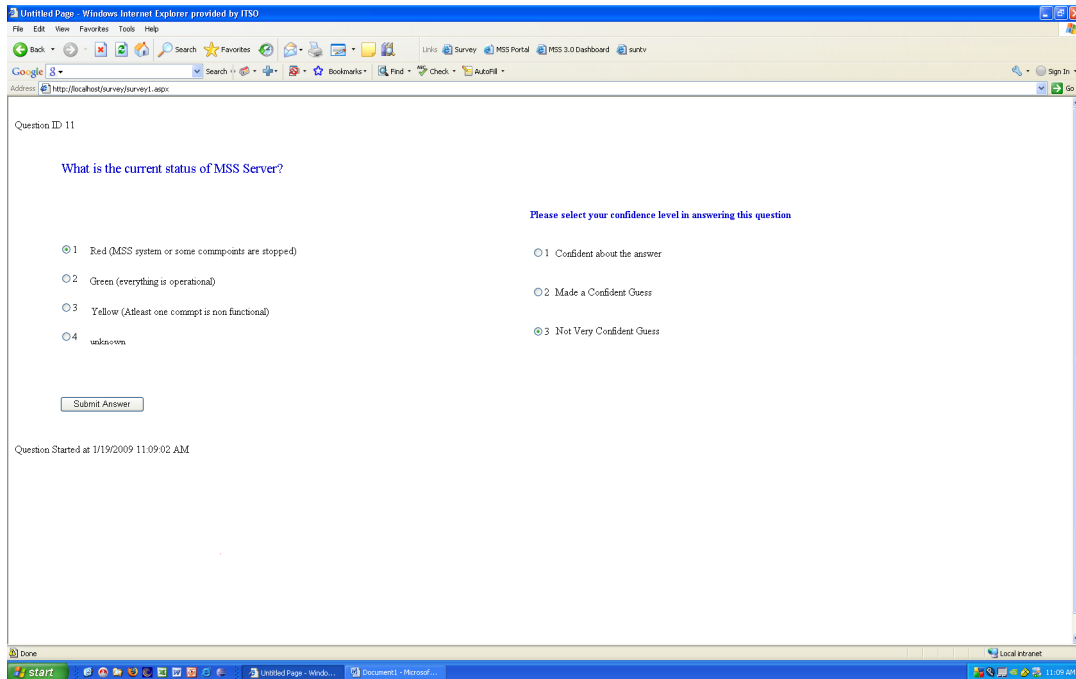


Figure 28: Web-Based Survey Instrument

A web-based survey instrument was developed to collect data. The web portal allowed users to login securely, and the system automatically assigned each user a random scenario. The subjects were instructed to attend to their tasks as they normally would, with the SAGAT queries regarded as secondary. The screen was frozen (blacked out) at random times, except the first time that was at the three-minute mark. The user was treated to a random interface (old or new) following which they were required to respond to a randomly selected 2 to 4 questions for a freeze interval with no more than 3 freezes. The total number of questions for a scenario was 9 (3 for each SA level). Other variables collected included time for response and confidence level. The duration of the freeze lasted until they had entered the answers for the SAGAT queries. No displays or visual aids were provided when the subjects were answering the queries. If subjects did not know the answer, they were encouraged to make their best guess. If they did not feel comfortable making a guess, they were allowed to go to the next question. In this study, simultaneous testing of multiple subjects was not performed. The study lasted for 15 minutes for

each scenario (freeze duration not accounted for), with an average of 4 freezes. The study results are discussed in detail in Chapter 4.

	Pilot Scenario	Scenario 1	Scenario 2
Issue	Unusual increasing trend in reporting for one lab over a week's time. Messages in error queue vocabulary & duplicates	High volume of reporting on a midweek from 2 of the 3 major clinical reference labs in the reporting jurisdiction	Increase trend of Arboviral diseases, FB diseases in one jurisdiction
Vignettes	<ul style="list-style-type: none"> Resending entire message history while sending updates Scheduled system update at a lab previous week 	<ul style="list-style-type: none"> Outbreak in the area 3rd lab with the highest market presence sending wrong codes (error messages) Seasonal bump expected for flu 	<ul style="list-style-type: none"> Heavy rain, flooding along the coast of river. Outbreak of salmonella No seasonal bump expected

Table 6: Scenarios for Evaluation Study

Chapter Summary

In this chapter, I described the method of implementing an interface to deliver higher levels of SA to the end user. I detailed the method, including two CTA techniques, Goal Directed Task Analysis (GDTA) and Concept Mapping (CM). GDTA was primarily done to identify the user goals and the information needed to successfully meet them. Concept Mapping was done to explicate the domain and the operational rules that are applied by the users to understand the data. I have detailed the implementation of a prototype for testing the SA levels delivered using the new method. I also introduced a SAGAT testing instrument developed to evaluate the user SA levels when using the new and the old interface.

In Chapter 4, I describe the user study that compares the SA levels of users when using the old and the new interface. The results are presented and discussed in Chapter 5.

CHAPTER 4- RESULTS

Overview & Objectives

My general goal in this study was to determine if the interface system built using the hybrid design is helping users to characterize the messaging trends and events. To measure the effectiveness of the interface, the mental construct of situational awareness was identified as a key component to evaluate among the users. Measuring the situational awareness of the user when performing a signal characterization task is done using the SAGAT technique as described in chapter 2.

Specifically, in this study I tested the claim that the users SA levels namely perception (level 1), interpretation (level 2) and for projection (level 3) will be higher when utilizing the interface built using the hybrid approach. To evaluate the levels, the user's performance was compared against the SA levels when they were using an interface built using traditional technology centered approach.

Comparison Systems

In this study, I compared user performance in a newly designed message brokering engine dashboard developed by applying the hybrid method (as discussed in chapter 3) against a dashboard of a message broken system that is used real world at over 26 public health jurisdiction. National Electronic Disease Surveillance System-Message Subscription Service (NEDSS-MSS) (Srinivasan, Abellera, Danos, & McNabb, 2008) is a CDC developed message brokering tool used by over 26 public health jurisdictions to receive and send real time electronic lab results, lab orders and case notification messages. For the study the NEDSS MSS dashboard was selected for evaluation due to the following reasons.

- The MSS dashboard is a critical component of electronic lab results monitoring task.

- The dashboard plays a critical role in informing the proper functioning of the public health reporting process (Case Notification).
- The dashboard is built using the traditional approach of technology-centered design.
- The system documentation, design and implementation are publicly available.

Although the system design included input from the end users, technology adoption e.g. BIRT reports, Jfreecharts, richfaces etc. has significantly changed the representation. For this study, the dashboard of the NEDSS MSS was simulated using BIRT reports, Jfreecharts in a webpage. The other technology that was used in implementing the dashboard was JavaScript.

The second interface is a newly designed prototype based on the concepts defined in chapter 3. The new interface is a UCD based design that utilizes domain knowledge to represent contextual data. The technology framework used in determining the contextual data includes Semantic Web solutions (e.g. Jena API, SPARQL) Rule engines (Drools). For the graphical representations, Google visualization API, Javascript, database and web technologies were utilized. Refer to Appendix D for a sample screenshot, code and data used for populating the dashboard. The process of identifying the data for presentation was explained in chapter 3.

Pilot Study

As part of the study, I conducted a pilot exercise to determine the following: (1) sample size needed to significantly detect a difference in correctness of response when using either interfaces, (2) whether any of the questions, instructions, or tasks were confusing; (3) validate if the SAGAT framework negatively impacted the user activity. I recruited 4 volunteers from the pilot study. All the volunteers were very familiar with the research. Two had hands on experience with the NEDSS MSS system, while two were trained in the use of NEDSS-MSS. One user was knowledgeable about the new contextualized system. This subject provided a

number of suggestions for clarifying the instructions and the demonstration section. These feedbacks were used to revise the demonstration section of the two interfaces. There was no changes effected on either interface, hence all 4 subjects were randomly treated to the two interfaces. The data was used in the power calculation to determine the appropriate number of subjects necessary to obtain significant results. The mean correctness score was 77.5% for the old interface questions and 82.5% for the new interface questions.

Using the software from the Biostatistics Primer (Glantz 1997), I determined that I would need between 51 - 54 subjects to significantly detect a difference of at least 5%points in the mean correctness scores. Sample size was calculated for comparing means of a paired sample with 80% power and 5% type-1 error. I decided to try to recruit 60 subjects to account for dropouts and no shows.

As a result of the pilot study, I made a few more changes to the wording of the SAGAT questions (changes the tense of the questions), and to the tutorials on each of the tools. The second, third and fourth subjects used the revised tutorial version and indicated that all instructions and questions were clear. None of the tools were modified during or after the pilot study.

Final Study

To evaluate the role of contextualized information representation in the domain of public health, I implemented the SAGAT for evaluating the SA of experts while performing the signal characterization task. The experts were to be tested for SA using both the current systems and the prototype implementation. In this section, I describe the final study and explain the evaluation methods, and report the study results.

Method

For this evaluation, I used methods from the field of human-computer interaction and situational awareness. Conventional summative and outcome-based evaluations, which are common in the health information systems evaluation domain, lack in their ability to describe the potentially important effects of computer systems on human cognitive processes (Kushniruk & Patel, 2004). So I decided to use methods of evaluation emerging from cognitive and usability engineering to measure the use's situational awareness domain. The application of empirical method like SAGAT was inspired by the high degree of validity, sensitivity and reliability across a range of studies when a broad range of queries were used in measuring SA (Vidulich, 1992). In the following sections, I outline the methods for this evaluation. I describe the subjects of the study and the procedure that these subjects followed.

Study Subjects

The subjects for this study are public health informaticians working in the domain of electronic laboratory and public health case reporting. Typically there are 2 or 3 such resources in every public health jurisdiction (state or city department of health) that has the authority to collect the reports sent by participating providers to meet the legal reporting requirements. Due to the very limited number of subjects, the inclusion criteria also seek any user with operational knowledge of the electronic message brokering systems. This allowed experts in the consulting industry working on projects to build, expand and deploy solutions at the public health entities also participate this study. Eligible subjects for this study should have prior experience working with public health message exchange systems (Rhapsody, MSS, Cloverleaf, Mirth, Biosense Integrator or a home grown system). The subject has attended one of the following trainings. The Rhapsody training with Orion health or the NEDSS Message Subscription Service training

sponsored by CDC. Based on the above criteria, I was able to recruit 64 subjects, out of which I had 4 opt out of the study after participating in one of the interfaces. I had to make the choice of dropping one subject due to scheduling conflicts. So data from 59 subjects were used in the study. The obtained sample consisted of 59 subjects having the demographic characteristics as reported in Table 2 below. The average age of the subjects was 40.71 years with a range of

Table 7 Demographic characteristics of the sample

Characteristic	Category	Frequency	Percent
Educational background	IT/Informatics	28	47.5
	Public Health	23	39.0
	Other	8	13.6
	Total	59	100.0
Experience	1 Year	14	23.7
	2 Years	18	30.5
	3 Years	14	23.7
	4 or more years	13	22.0
	Total	59	100.0

29 to 60 years. The percent of subjects who were 40 or older was 49.2%. Each subject signed a written consent form before participating in the study.

Procedure

Each subject willing to participate in the study was scheduled for two 45 min appointments. A web based scheduling tool (doodle.com) was used to identify a 45 min slot for meeting which had a minimum of 30 days apart. The average difference between the two

appointments was 47 days. The participant was requested to sign and fax or email a scanned copy of the signed consent form before the first appointment. The appointments are all web based and did not require physical presence. During the appointment, the subject was again asked to verify their training and experience with integration engines and similar technologies. After that each subject was assigned a unique ID, which will be used to track their record in the future sessions.

Upon completion of the above activity, the subject was presented a demo version of a message exchange solution dashboard (for both new and old interfaces). The demo was based on a real world situation (duplicate reporting scenario). After that the investigator led demo, the subject was asked to spend 10 minutes assessing the dashboard and familiarize with various data elements presented on the dashboard. During this period, the study investigator was available over phone or live meeting software to answer any questions pertaining to the demo page and the representations contained within. The subject was also provided a demonstration of the questions page and the steps in submitting a response. Once the subject familiarized and felt comfortable with the questions format and the survey layout, the core activity was then started. The subject was queried one last time before the core activity to answer any questions on the process or the tools.

The core activity consisted of a simulated scenario in a public health message exchange dashboard. See Appendix B for the two scenarios. The subject was asked to perform his/her daily tasks of analyzing the content of the dashboard. The screen was frozen blank at randomly selected time and the survey page was displayed. The subject was asked to answer a random set of questions based on the display he/she had just analyzed. Three measurements are done in the survey page, subject's response to the question, time taken to come up with an answer and the

level of confidence while responding to the question. Upon answering a random number of questions, the dashboard screen will reappear. The subject was asked to continue the analysis. The process of freezing the dashboard and questioning the user happened multiple times for random intervals based on SAGAT requirements. During each freeze session no more than 9 questions were queried and the total number of questions did not exceed 30.

At the end of the survey activity the subject was requested to comment about his/her general experiences and provide suggestions for improvement. If the subject was interested, they had the opportunity to receive feedback on their performance at the end of the second appointment (after completing both the interfaces).

Analysis of Results

In this section, I discuss the results from the final study. The results of the study have been grouped into three sets based on the hypothesis tested. In the analysis section, the newly designed interface is referred as “new interface” and the NEDSS MSS interface is referred as “Conventional Interface”. The first set of 7 hypotheses was characterized as response-related, meaning related to whether or not the response was correct. The second set was characterized as response time related and the last set is on user confidence.

Response Type Related Analysis

In the study, all the responses were captured as a correct or incorrect response. The following analysis tests the following 7 hypotheses:

Responding to Perception, Interpretation and Forecasting Questions

Hypothesis 1 states that the rates of correct responses for perception questions are different between the conventional and new interface. This hypothesis was tested by means of a McNemar test of the dependent 2-way contingency table shown in Table 8

Table 8 Cross-tabulation of correct and incorrect responses to perception questions

		New Interface		
		Correct	Incorrect	Total
Conventional Interface	Correct	380	65	445
	Incorrect	77	9	86
	Total	457	74	531

The McNemar test for the frequencies in Table 8 produced a p -value of .356. Consequently, the null hypothesis of no difference between the two interfaces in the rates of correct responses to perception questions cannot be rejected. This showed that there is no difference in the level of awareness between the two interfaces when responding to perception questions.

Hypothesis 2 states that the rates of correct responses for interpretation questions are different between the conventional and new interface. This hypothesis was tested by means of a McNemar test of the dependent 2-way contingency table shown in Table 9.

Table 9 Cross-tabulation of correct and incorrect responses to interpretation questions

		New Interface		
		Correct	Incorrect	Total
Conventional Interface	Correct	152	285	437
	Incorrect	32	62	94
	Total	184	347	531

The McNemar test for the frequencies in Table 9 produced a p -value of $<.0001$. Consequently, the null hypothesis of no difference between the two interfaces in the rates of correct responses to interpretation questions is rejected. The rate of correct responses was significantly greater in the new interface condition.

Hypothesis 3 states that the rates of correct responses for forecasting questions are different between the conventional and new interface. This hypothesis was tested by means of a McNemar test of the dependent 2-way contingency table shown in Table 10.

Table 10 Cross-tabulation of correct and incorrect responses to forecasting questions

		New Interface		
		Correct	Incorrect	Total
Conventional Interface	Correct	107	308	415
	Incorrect	29	87	116
	Total	136	395	531

The McNemar test for the frequencies in Table 10 produced a p -value of $<.0001$.

Consequently, the null hypothesis of no difference between the two interfaces in the rates of correct responses to forecasting questions is rejected. The rate of correct responses was significantly greater in the new interface condition.

Impact of Age, Experience and Edu Background to Perception Questions

Hypothesis 4 states that there are differences between age groups, occupational experience levels, and educational background in the rates of correct responses to the perception questions. Age groups were defined as under 40 and 40 years old and older for the purposes of this and all subsequent hypotheses addressing age groups. The test of this hypothesis with

respect to age groups was conducted by means of separate chi-square analyses of the 2-way contingency tables for the two interfaces shown in Table 11.

Table 11 Cross-tabulation of levels of correct responses to perception questions for the two age groups

Interface	Age group	Response correctness		Total	Chi-square	<i>p</i>
		Correct	Incorrect			
New	Under 40	230	40	270	.772	.380
	40 or Older	215	46	261		
	Total	445	86	531		
Conventional	Under 40	234	36	270	.166	.683
	40 or Older	223	38	261		
	Total	457	74	531		

The chi-squares for the frequencies in Table 11 did not reach the threshold for statistical significance. Consequently, the null hypothesis of no difference between the two age groups in their rates of correct responses to perception questions cannot be rejected for either interface.

The test of this hypothesis with respect to occupational experience levels was conducted by means of separate chi-square analyses of the 4 X 2 contingency tables for the two interfaces shown in Table 12.

Table 12 Cross-tabulation of levels of correct responses to perception questions for the four occupational experience levels

Interface	Experience level	Response correctness		Total	Chi-square	<i>p</i>
		Correct	Incorrect			
New	1 Year	110	16	126	1.495	.883
	2 Years	134	28	162		
	3 Years	104	22	126		
	4 or more years	97	20	117		
	Total	445	86	531		
Conventional	1 Year	108	18	126	.188	.980
	2 Years	141	21	162		
	3 Years	108	18	126		
	4 or more years	100	17	117		
	Total	457	74	531		

The chi-squares for the frequencies in Table 12 did not reach the threshold for statistical significance. Consequently, the null hypothesis of no difference between the four occupational experience levels in their rates of correct responses to perception questions cannot be rejected for either interface.

For the purposes of this hypothesis and all subsequent ones addressing educational background groups, a dichotomous representation of educational background was used consisting of the following two categories: IT/Informatics and Other. The test of this hypothesis

with respect to educational background was conducted by means of separate chi-square analyses of the 2-way contingency tables for the two interfaces shown in Table 13.

Table 13 Cross-tabulation of levels of correct responses to perception questions for the two educational background groups

Interface	Educational background	Response correctness		Total	Chi-square	P
		Correct	Incorrect			
New	IT/Informatics	214	38	252	.440	.507
	Other	231	48	279		
	Total	445	86	531		
Conventional	IT/Informatics	212	40	252	1.500	.221
	Other	245	34	279		
	Total	457	74	531		

The chi-squares for the frequencies in Table 13 did not reach the threshold for statistical significance. Consequently, the null hypothesis of no difference between the two educational background groups in their rates of correct responses to perception questions cannot be rejected for either interface.

Impact of Age, Experience and Edu Background to Interpretation Questions

Hypothesis 5 states that there are differences between age groups, occupational experience levels, and educational background in the rates of correct responses to the interpretation questions. The test of this hypothesis with respect to age groups was conducted by

means of separate chi-square analyses of the 2-way contingency tables for the two interfaces shown in Table 14.

Table 14 Cross-tabulation of levels of correct responses to interpretation questions for the two age groups

Interface	Age group	Response correctness		Total	Chi-square	<i>p</i>
		Correct	Incorrect			
New	Under 40	224	46	270	.167	.683
	40 or Older	213	48	261		
	Total	437	94	531		
Conventional	Under 40	91	179	270	.218	.641
	40 or Older	93	168	261		
	Total	184	347	531		

The chi-squares for the frequencies in Table 14 did not reach the threshold of statistical significance for either interface. Consequently, the null hypothesis of no difference between the two age groups in their rates of correct responses to interpretation questions cannot be rejected for either interface.

The test of the fifth hypothesis with respect to occupational experience levels was conducted by means of separate chi-square analyses of the 4 X 2 contingency tables for the two interfaces shown in Table 15.

Table 15 Cross-tabulation of levels of correct responses to interpretation questions for the four occupational experience levels

Interface	Experience level	Response correctness		Total	Chi-square	<i>p</i>
		Correct	Incorrect			
New	1 Year	112	14	126	6.367	.095
	2 Years	127	35	162		
	3 Years	100	26	126		
	4 or more years	98	19	117		
	Total	437	94	531		
Conventional	1 Year	51	75	126	6.957	.073
	2 Years	59	103	162		
	3 Years	32	94	126		
	4 or more years	42	75	117		
	Total	184	347	531		

The chi-squares for the frequencies in Table 15 did not reach the threshold for statistical significance for either interface. Consequently, the null hypothesis of no difference between the four occupational experience levels in their rates of correct responses to interpretation questions cannot be rejected for either interface.

The test of Hypothesis 5 with respect to educational background was conducted by means of separate chi-square analyses of the 2-way contingency tables for the two interfaces shown in Table 16.

Table 16 Cross-tabulation of levels of correct responses to interpretation questions for the two educational background groups

Interface	Educational background	Response correctness		Total	Chi-square	P
		Correct	Incorrect			
New	IT/Informatics	205	47	252	.296	.586
	Other	232	47	279		
	Total	437	94	531		
Conventional	IT/Informatics	83	169	252	.623	.430
	Other	101	178	279		
	Total	184	347	531		

The chi-squares for the frequencies in Table 16 did not reach the threshold for statistical significance. Consequently, the null hypothesis of no difference between the two educational background groups in their rates of correct responses to interpretation questions cannot be rejected for either interface.

Impact of Age, Experience and Edu Background to Forecasting Questions

Hypothesis 6 states that there are differences between age groups, occupational experience levels, and educational background in the rates of correct responses to the forecasting questions. The test of this hypothesis with respect to age groups was conducted by means of separate chi-square analyses of the 2-way contingency tables for the two interfaces shown in Table 17.

Table 17 Cross-tabulation of levels of correct responses to forecasting questions for the two age groups

Interface	Age group	Response correctness		Total	Chi-square	<i>p</i>
		Correct	Incorrect			
New	Under 40	209	61	270	.180	.672
	40 or Older	206	55	261		
	Total	415	116	531		
Conventional	Under 40	72	198	270	.321	.571
	40 or Older	64	197	261		
	Total	136	395	531		

The chi-squares for the frequencies in Table 17 did not reach the threshold for statistical significance. Consequently, the null hypothesis of no difference between the two age groups in their rates of correct responses to forecasting questions cannot be rejected for either interface.

The test of the sixth hypothesis with respect to occupational experience levels was conducted by means of separate chi-square analyses of the 4 x 2 contingency tables for the two interfaces shown in Table 18.

Table 18 Cross-tabulation of levels of correct responses to forecasting questions for the four occupational experience levels

Interface	Experience level	Response correctness		Total	Chi-square	<i>p</i>
		Correct	Incorrect			
New	1 Year	112	14	126	6.397	.095
	2 Years	127	35	162		
	3 Years	100	26	126		
	4 or more years	98	19	117		
	Total	437	94	531		
Conventional	1 Year	33	93	126	.578	.901
	2 Years	40	122	162		
	3 Years	35	91	126		
	4 or more years	28	89	117		
	Total	136	395	531		

The chi-squares for the frequencies in Table 18 do not reach the threshold for statistical significance. Consequently, the null hypothesis of no difference between the four occupational experience levels in their rates of correct responses to forecasting questions cannot be rejected for either interface.

The test of Hypothesis 6 with respect to educational background was conducted by means of separate chi-square analyses of the 2-way contingency tables for the two interfaces shown in Table 19.

Table 19 Cross-tabulation of levels of correct responses to forecasting questions for the two educational background groups

Interface	Educational background	Response correctness		Total	Chi-square	<i>p</i>
		Correct	Incorrect			
New	IT/Informatics	192	60	252	1.084	.298
	Other	223	56	279		
	Total	415	116	531		
Conventional	IT/Informatics	71	181	252	1.653	.199
	Other	65	214	279		
	Total	136	395	531		

The chi-squares for the frequencies in Table 19 do not reach the threshold for statistical significance. Consequently, the null hypothesis of no difference between the two educational background groups in their rates of correct responses to forecasting questions cannot be rejected for either interface.

Relationship between SA level to Perception & Interpretation Questions

Hypothesis 7 proposes that the relationship between the correctness of one’s responses to the perception and interpretation questions is stronger for the new interface than for the conventional interface. This was tested by computing the Pearson correlations between the correctness scores (i.e., 1 = correct, 2 = incorrect) for perception and interpretation performance using the new interface and using the conventional interface. The difference between the two non-overlapping correlated correlations was tested using the Pearson-Filon test, with significance

being assessed on the basis of one-tailed p-value. The correlations between all four scores are reported in Table 20.

Table 20 Correlations between response correctness scores on perception and interpretation questions for the two interfaces

Interface	Task	Condition	2	3	4
New	Perception	1	.144***	-.044	.009
New	Interpretation	2		-.001	.006
Conventional	Perception	3			.064
Conventional	Interpretation	4			

***. Correlation is significant at the 0.001 level (2-tailed). N = 531 for all correlations

The correlations between perception and interpretation response correctness were .144 for the new interface and .064 for the conventional interface. The Pearson-Filon z-value for the difference between these correlations was 1.315, which has a 1-tailed p-value of .0945.

This result indicates that the null hypothesis of no difference between the interfaces in the relationship between the perception and interpretation questions cannot be rejected. The use of the new interface did not result in a significantly stronger positive relationship between the correctness of subjects' responses to the perception and interpretation questions.

In order to interpret the Hypothesis 7 results better, further analysis was done to see if the proportion of subjects who achieved correct responses to both the perception and interpretation questions was significantly higher for the new interface than for the conventional interface. This was tested by the application of the McNemar test for correlated proportions. The contingency table on which the McNemar test was computed is presented in Table 21.

Table 21 Contingency Table for Both Correct vs. One or More Incorrect Responses to the Perception and Interpretation Questions by Interface Type

New Interface: Perception & Interpretation Responses	Conventional Interface: Perception & Interpretation Responses		Total
	Incorrect on one or both	Correct on both	
Incorrect on one or both	107	47	154
Correct on both	260	117	377
Total	367	164	531

The McNemar test produced a p-value of <.0001, 1-tailed. This result indicates that the null hypothesis of no difference between the interfaces in the proportions of the subjects who correctly responded to both the perception and interpretation questions can be rejected. The use of the new interface resulted in a significantly higher proportion (i.e., .7099) of subjects who responded correctly to both types of questions than occurred among the same subjects when they used the conventional interface (i.e., .3094).

Relationship between SA level to Interpretation & Forecasting Questions

Hypothesis 8 proposes that the relationship between the correctness of one’s responses to the interpretation and forecasting questions is stronger for the new interface than for the conventional interface. This was tested by computing the Pearson correlations between the correctness scores (i.e., 1 = correct, 2 = incorrect) on the interpretation and forecasting tasks using the new interface and using the conventional interface. The difference between the two non-overlapping correlated correlations was tested using the Pearson-Filon test, with significance being assessed on the basis of one-tailed p-value. The correlations between all four scores are reported in Table 22.

Table 22 Correlations between response correctness scores on perception and interpretation questions for the two interfaces

Interface	Task	Condition	2	3	4
New	Interpretation	1	.209***	.006	.046
New	Forecasting	2		-.017	.007
Conventional	Interpretation	3			.062
Conventional	Forecasting	4			

***. Correlation is significant at < 0.001 level (2-tailed). N = 531 for all correlations

The correlations between interpretation and forecasting response correctness scores were .209 for the new interface and .062 for the conventional interface. The Pearson-Filon z-value for the difference between these correlations was 2.437, which has a 1-tailed p-value of .0075.

This result indicates that the null hypothesis of no difference between the interfaces in the relationship between the interpretation and forecasting questions can be rejected. The use of the new interface resulted in a significantly stronger positive relationship between the correctness of subjects' responses to the interpretation and forecasting questions.

In order to interpret the Hypothesis 8 results better, further analysis was done to see if the proportion of subjects who achieved correct responses to both the interpretation and forecasting questions was significantly higher for the new interface than for the conventional interface. This was tested by the application of the McNemar test for correlated proportions. The contingency table on which the McNemar test was computed is presented in Table 23.

Table 23 Contingency Table for Both Correct vs. One or More Incorrect Responses to the Perception and Interpretation Questions by Interface Type

		Conventional Interface: Interpretation & Forecasting Responses		
		Incorrect on one or both	Correct on both	Total
New Interface: Interpretation & Forecasting Responses	Incorrect on one or both	154	18	172
	Correct on both	323	36	359
Total		477	54	531

The McNemar test produced a p-value of <.0001, 1-tailed. This result indicates that the null hypothesis of no difference between the interfaces in the proportions of the subjects who correctly responded to both the interpretation and forecasting questions can be rejected. The use of the new interface resulted in a significantly higher proportion (i.e., .6761) of subjects who responded correctly to both types of questions than occurred among the same subjects when they used the conventional interface (i.e., .1017).

Response Time Related Analysis

Hypotheses 9 to 19 focused on the dependent variable of response time as a measure of the efficacy of the interfaces being compared.

Response Time Analysis by SA Level

Hypothesis 9 holds that user response time for perception questions differs between the two new and conventional interfaces. This hypothesis was tested by means of a paired t-test on the response times for the perception question achieved through the new and conventional interfaces. The results of this test are presented in Table 24.

Table 24 Results of t-test of difference between the two interfaces in their mean response times to the perception questions

Response time means		Paired Differences					
Conventional interface	New interface	Mean	Std. deviation	Std. error of mean difference	t	df	Sig. (2-tailed)
26.5574	16.8475	-9.71	3.24	.42	-23.054	58	<.001

The results in Table 24 indicate that the null hypothesis that there is no difference in mean response time between the two interfaces should be rejected. The mean response time for perception questions was significantly lower for the new interface than for the conventional one.

Hypothesis 10 holds that user response time for interpretation questions is shorter for the new interface than for the conventional interface. This one-tailed hypothesis was tested by means of a paired t-test on the response times for the interpretation question achieved through the new and conventional interfaces. The results of this test are presented in Table 25.

Table 25 Results of t-test of difference between the two interfaces in their mean response times to the interpretation questions

Response time means		Paired Differences					
Conventional interface	New interface	Mean	Std. deviation	Std. error of mean difference	t	df	Sig. (2-tailed)
94.2316	34.6403	-59.59	17.00	2.34	-25.436	58	<.001

The results in Table 25 indicate that the null hypothesis that there is no difference in mean response time between the two interfaces should be rejected. The mean response time for interpretation questions was significantly lower for the new interface than for the conventional one.

Hypothesis 11 holds that user response time for forecasting questions is shorter for the new interface than for the conventional interface. This one-tailed hypothesis was tested by means of a paired t-test on the response times for the forecasting question achieved through the new and conventional interfaces. The results of this test are presented in Table 26.

Table 26 Results of t-test of difference between the two interfaces in their mean response times to the forecasting questions

Response time means		Paired Differences					
Conventional interface	New interface	Mean	Std. deviation	Std. error of mean difference	T	df	Sig. (2-tailed)
155.7363	42.7137	-113.02	20.52	2.67	-42.306	58	<.001

The results in Table 26 indicate that the null hypothesis that there is no difference in mean response time between the two interfaces should be rejected. The mean response time for forecasting questions was significantly lower for the new interface than for the conventional one.

Response Time Analysis by Type of Response for Level 1 Questions

Hypothesis 12 states that response times differ between correct and incorrect responses for perception questions in each interface. This hypothesis was tested by conducting two ANOVAs, one for each interface, using response time for perception questions as the dependent variable and correctness of response and subject as the independent variables. Subject was included as a factor in order to remove the effect of differences between subjects due to differences in the scenario, session, and particular question used in eliciting responses to the perception questions. The ANOVA results are summarized in Table 27.

Table 27 Results of ANOVA of response time for perception questions by correctness of response and subject for each interface

Interface	Source	df	F	Sig.	η^2
New	Response correctness	1	1.754	.186	.003
	Subject	58	1.669	.003	.161
	Response correctness * subject	41	1.756	.004	.120
	Error	430	(26.041)		
Conventional	Response correctness	1	6.003	.015	.012
	Subject	58	.505	.999	.061
	Response correctness * subject	42	.385	1.000	.034
	Error	429	(58.077)		

Note: Values in parentheses are mean square errors.

The results in Table 27 indicate that the null hypothesis of no difference in response times between correct and incorrect responders using either interface must be rejected. Correct responders using the new interface exhibited a significantly lower mean response time than incorrect responders (16.73 vs. 17.43). Correct responders using the conventional interface exhibited a significantly higher mean response time than incorrect responders. (26.92 vs. 24.30). Thus, the direction of the difference was inconsistent between the two interfaces and the degree of difference within each interface, although significant, was very small in practical terms. Of more interest, perhaps, was the wide difference in response times between correct responders using the two interfaces. Considering only the 360 instances of subjects giving correct responses

to a perception question using both interfaces, the mean response times were 16.67 using the new interface 26.74 using the conventional interface (t for difference = 21.419, $p < .001$).

Response Time Analysis by Types of Response for Level 2 Questions

Hypothesis 13 states that response times differ between correct and incorrect responses for interpretation questions in each interface. This hypothesis was tested by conducting two ANOVAs, one for each interface, using response time for interpretation questions as the dependent variable and correctness of response and subject as the independent variables. Subject was included as a factor in order to remove the effect of differences between subjects due to differences in the scenario, session, and particular question used in eliciting responses to the interpretation questions. The ANOVA results are summarized in Table 28.

Table 28 Results of ANOVA of response time for interpretation questions by correctness of response and subject for each interface

Interface	Source	df	F	Sig.	η^2
New	Response correctness	1	568.119	<.001	.454
	Subject	58	2.289	<.001	.106
	Response correctness * subject	40	3.010	<.001	.096
	Error	431	(176.081)		
Conventional	Response correctness	1	872.904	<.001	.639
	Subject	58	.630	.984	.027
	Response correctness * subject	52	.716	.931	.027
	Error	419	(490.382)		

Note: Values in parentheses are mean square errors.

The results in Table 28 indicate that the null hypothesis of no difference in mean response times to interpretation questions between correct and incorrect responders using either interface must be rejected. Correct responders using the new interface exhibited a significantly lower mean response time than incorrect responders (27.07 vs. 69.82). Correct responders using the conventional interface also exhibited a significantly lower mean response time than incorrect responders. (49.05 vs. 118.19). Thus, the direction of the difference in mean response times between correct and incorrect responders was consistent between the two interfaces, and the degree of difference within each interface was both statistically significant and very large in practical terms. Again, there was also wide difference in response times between correct responders using the two interfaces. Considering only the 152 instances of subjects giving correct responses to a interpretation question using both interfaces, the mean response times were 26.76 using the new interface 49.03 using the conventional interface (t for difference = 14.656, $p < .001$).

Response Time Analysis by Types of Response for Level 3 Questions

Hypothesis 14 states that response times differ between correct and incorrect responses for forecasting questions in each interface. This hypothesis was tested by conducting two ANOVAs, one for each interface, using response time for forecasting questions as the dependent variable and correctness of response and subject as the independent variables. Subject was included as a factor in order to remove the effect of differences between subjects due to differences in the scenario, session, and particular question used in eliciting responses to the forecasting questions. The ANOVA results are summarized in Table 29.

Table 29 Results of ANOVA of response time for forecasting questions by correctness of response and subject for each interface

Interface	Source	df	F	Sig.	η^2
New	Response correctness	1	161.702	<.001	.208
	Subject	58	1.646	.003	.123
	Response correctness * subject	43	2.111	<.001	.117
	Error	428	(317.867)		
Conventional	Response correctness	1	355.412	.000	.421
	Subject	58	.615	.988	.042
	Response correctness * subject	55	.664	.969	.043
	Error	416	(2208.683)		

Note: Values in parentheses are mean square errors.

The results in Table 29 indicate that the null hypothesis of no difference in response times to forecasting questions between correct and incorrect responders using either interface must be rejected. Correct responders using the new interface exhibited a significantly lower mean response time than incorrect responders (36.46 vs. 65.09). Correct responders using the conventional interface also exhibited a significantly lower mean response time than incorrect responders. (84.96 vs. 180.11). Thus, the direction of the difference in mean response times between correct and incorrect responders was consistent between the two interfaces and the degree of difference within each interface was both statistically significant and very large in practical terms. Again, there was also wide difference in response times between correct

responders using the two interfaces. Considering only the 107 instances of subjects giving correct responses to a forecasting question using both interfaces, the mean response times were 38.93 using the new interface 87.06 using the conventional interface (t for difference = 18.275, $p < .001$).

Impact of Age, Experience and Edu on Response Time for Level 1 Questions

Hypothesis 15 posited that within each interface, user response time for perception questions is related to age, years of experience, and educational background. The test of this hypothesis with respect to age was conducted by means of a t-test of the difference in mean response times to the perception questions averaged within subject for subjects under 40 years of ages and for subjects 40 or more years of age. The results of this test are shown in Table 30.

Table 30 Results of t-test of difference between under 40 and 40 or older age groups in mean response times for perception questions

Interface	Response time means		Difference		t	df	Sig. (2-tailed)
	Under 40	40 or older	Mean	Std. error of mean difference			
New	16.41	17.30	-.895	.490	-1.828 ^a	45.82	.074
Conventional	26.88	26.22	.659	.622	1.059 ^b	57	.294

^a Levene’s test significant at $p = .018$, equal variance not assumed.

^b Levene’s test not significant, equal variances assumed.

The results in Table 30 indicate that the differences between age groups in mean response times to perception questions did not reach statistical significance in either interface. Consequently, the null hypothesis of no difference between age groups in their mean response times to perception questions in either interface cannot be rejected.

The test of hypothesis 15 with respect to occupational experience levels was conducted by means of an analysis of variance of the responses obtained with each interface using a 4-level categorization of occupational experience (i.e., 1, 2, 3, and 4 or more years of experience) as the ANOVA factor and response times to perception questions averaged within subject as the dependent variable. The result of these ANOVAs are reported in Table 31.

Table 31 Results of analysis of variance of perception question response times by occupational experience levels for the new and conventional interfaces

Interface	Source	df	F	Sig.	η^2
New	Occupational experience level	3	.922	.436	.048
	Error	55	(3.637)		
Conventional	Occupational experience level	3	.637	.594	.034
	Error	55	(5.832)		

Note: Values in parentheses are mean square errors.

The results in Table 31 indicate that the differences between occupational experience levels in mean response times to perception questions did not reach statistical significance in either interface. Consequently, the null hypothesis of no difference between occupational experience levels in their mean response times to perception questions in either interface cannot be rejected.

The test of Hypothesis 15 with respect to educational background was conducted by means of a t-test of the difference in mean response times to the perception questions for subjects having an IT/Informatics educational background and those having a Public Health or other type of educational background. The results of this test are shown in Table 32.

Table 32 Results of t-test of difference between IT/Informatics vs. other educational background groups in mean response times for perception questions

Interface	Response time means		Difference		T	df	Sig. (2-tailed)
	IT/Informatics education	Other education	Mean	Std. error of mean difference			
New	17.17	16.56	.615	.494	1.245 ^a	57	.218
Conventional	26.05	27.02	-.970	.616	-1.576 ^a	57	.121

Note: sample sizes for IT/Informatics and Other educational backgrounds were 28 and 31, respectively for both the new and conventional interfaces.

^a Levene's test not significant, equal variances assumed.

The results in Table 32 indicate that the differences between the two educational background groups in mean response times to perception questions did not reach statistical significance in either interface. Consequently, the null hypothesis of no difference between the two educational background groups examined in this study in their mean response times to perception questions in either interface cannot be rejected.

Impact of Age, Experience and Edu on Response Time for Level 2 Questions

Hypothesis 16 posited that within each interface, user response time for interpretation questions is related to age, years of experience, and educational background. The test of this hypothesis with respect to age was conducted by means of a t-test of the difference in mean response times to the interpretation questions averaged within subject for subjects under 40 years of ages and for subjects 40 or more years of age. The results of this test are shown in Table 33.

Table 33 Results of t-test of difference between under 40 and 40 or older age groups in mean response times for interpretation questions

Interface	Response time means		Difference		t	df	Sig. (2-tailed)
	Under 40	40 or older	Mean	Std. error of mean difference			
New	34.54	34.74	-.195	2.143	-.091 ^a	57	.928
Conventional	94.51	93.94	.576	.4.062	.142 ^a	57	.888

Note: sample sizes for under 40 and 40 or older were 30 and 29, respectively for both the new and conventional interfaces.

^a Levene's test not significant, equal variances assumed.

The results in Table 33 indicate that the differences between age groups in mean response times to interpretation questions did not reach statistical significance in either interface. Consequently, the null hypothesis of no difference between age groups in their mean response times to interpretation questions in either interface cannot be rejected.

The test of hypothesis 16 with respect to occupational experience levels was conducted by means of an analysis of variance of the responses obtained with each interface using a 4-level categorization of occupational experience (i.e., 1, 2, 3, and 4 or more years of experience) as the ANOVA factor and response times to interpretation questions averaged within subject as the dependent variable. The result of these ANOVAs are reported in Table 34.

Table 34 Results of t-test of difference between under 40 and 40 or older age groups in mean response times for interpretation questions

Interface	Source	df	F	Sig.	η^2
New	Occupational experience level	3	.500	.684	.027
	Error	55	(68.309)		
Conventional	Occupational experience level	3	1.305	.282	.066
	Error	55	(235.472)		

Note: Values in parentheses are mean square errors.

The results in Table 34 indicate that the differences between occupational experience levels in mean response times to interpretation questions did not reach statistical significance in either interface. Consequently, the null hypothesis of no difference between occupational experience levels in their mean response times to interpretation questions in either interface cannot be rejected.

The test of Hypothesis 16 with respect to educational background was conducted by means of a t-test of the difference in mean response times to the interpretation questions for subjects having an IT/Informatics educational background and those having a Public Health or other type of educational background. The results of this test are shown in Table 35.

Table 35 Results of t-test of difference between IT/Informatics vs. other educational background groups in mean response times for interpretation questions

Interface	Response time means		Difference		t	df	Sig. (2-tailed)
	IT/Informatics education	Other education	Mean	Std. error of mean difference			
New	34.38	34.88	-.501	2.144	-.234 ^a	57	.816
Conventional	94.86	93.66	1.198	4.064	.295 ^a	57	.769

Note: sample sizes for IT/Informatics and Other educational backgrounds were 28 and 31, respectively for both the new and conventional interfaces.

^a Levene's test not significant, equal variances assumed.

The results in Table 35 indicate that the differences between the two educational background groups in mean response times to interpretation questions did not reach statistical significance in either interface. Consequently, the null hypothesis of no difference between the two educational background groups examined in this study in their mean response times to interpretation questions in either interface cannot be rejected.

Impact of Age, Experience and Edu on Response Time for Level 3 Questions

Hypothesis 17 posited that within each interface, user response time for forecasting questions is related to age, years of experience, and educational background. The test of this hypothesis with respect to age was conducted by means of a t-test of the difference in mean response times to the forecasting questions averaged within subject for subjects under 40 years of ages and for subjects 40 or more years of age. The results of this test are shown in Table 36.

Table 36 Results of t-test of difference between under 40 and 40 or older age groups in mean response times for forecasting questions

Interface	Response time means		Difference		t	df	Sig. (2-tailed)
	Under 40	40 or older	Mean	Std. error of mean difference			
New	43.34	42.06	1.283	2.283	.616 ^a	57	.540
Conventional	157.76	153.64	4.116	4.927	.835 ^a	57	.407

Note: sample sizes for under 40 and 40 or older were 30 and 29, respectively for both the new and conventional interfaces.

^a Levene's test not significant, equal variances assumed.

The results in Table 36 indicate that the differences between age groups in mean response times to forecasting questions did not reach statistical significance in either interface. Consequently, the null hypothesis of no difference between age groups in their mean response times to forecasting questions in either interface cannot be rejected.

The test of hypothesis 16 with respect to occupational experience levels was conducted by means of an analysis of variance of the responses obtained with each interface using a 4-level categorization of occupational experience (i.e., 1, 2, 3, and 4 or more years of experience) as the ANOVA factor and response times to forecasting questions averaged within subject as the dependent variable. The result of these ANOVAs are reported in Table 37.

Table 37 Results of analysis of variance of forecasting question response times by occupational experience levels for the new and conventional interfaces

Interface	Source	df	F	Sig.	η^2
New	Occupational experience level	3	.440	.726	.028
	Error	55	(65.171)		
Conventional	Occupational experience level	3	.249	.862	.013
	Error	55	(370.484)		

Note: Values in parentheses are mean square errors.

The results in Table 37 indicate that the differences between occupational experience levels in mean response times to forecasting questions did not reach statistical significance in either interface. Consequently, the null hypothesis of no difference between occupational experience levels in their mean response times to forecasting questions in either interface cannot be rejected.

The test of Hypothesis 16 with respect to educational background was conducted by means of a t-test of the difference in mean response times to the forecasting questions for subjects having an IT/Informatics educational background and those having a Public Health or other type of educational background. The results of this test are shown in Table 38.

Table 38 Results of t-test of difference between IT/Informatics vs. other educational background groups in mean response times for forecasting n questions

Interface	Response time means		Difference		t	df	Sig. (2-tailed)
	IT/Informatics education	Other education	Mean	Std. error of mean difference			
New	42.47	42.93	-.460	2.091	-.220 ^a	57	.827
Conventional	157.35	154.28	3.077	4.946	.622 ^a	57	.536

Note: sample sizes for IT/Informatics and Other educational backgrounds were 28 and 31, respectively for both the new and conventional interfaces.

^a Levene's test not significant, equal variances assumed.

The results in Table 38 indicate that the differences between the two educational background groups in mean response times to forecasting questions did not reach statistical significance in either interface. Consequently, the null hypothesis of no difference between the two educational background groups examined in this study in their mean response times to forecasting questions in either interface cannot be rejected.

Hypothesis 18 predicts that when responses are obtained by means of the new interface, the mean response times for the interpretation questions will be shorter among those responding correctly to perception questions than among those responding incorrectly to such questions, and that this difference will not be observable when responses are obtained by means of the conventional interface. This hypothesis was tested by conducting t-tests for the difference in interpretation question response times between correct and incorrect responders to the perception questions under the two interface conditions. The results of these t-tests are reported in Table 39.

Table 39 Results of t-tests of the difference in interpretation question response times between correct and incorrect perception question responders under the two interface conditions

Interface	Response time means		Difference		t	df	Sig. (2-tailed)
	Correct perception question responders	Incorrect perception question responders	Mean	Std. error of mean difference			
New	33.44	40.83	-7.381	2.911	-2.536	108.143	.013
Conventional	93.51	98.68	-5.164	4.925	-1.048	529	.295

Note: sample sizes for correct responders and incorrect responders were 445 and 86, respectively, for the new interface, and 457 and 74, respectively, for the conventional interface.
^a Levene's test not significant, equal variances assumed.

The results in Table 39 support the prediction of Hypothesis 18 and justify the rejection of the null hypothesis. The predicted lower response time to interpretation questions among those responding correctly to the perception questions in comparison to those responding incorrectly was observed to occur under the new interface condition but not under the conventional interface condition. While testing, hypothesis 18, I looked at alternative way of testing. This hypothesis predicts that there will be a significant correlation between response times to the perception and interpretation questions under the new interface condition, and that this correlation will be significantly higher than that between the corresponding response times obtained under the conventional interface condition. This hypothesis was tested by comparing the correlations between response times to the perception and interpretation questions obtained under each interface condition. The null hypothesis of no difference between the correlations obtained for two screen types could not be rejected. The correlation for the new interface was -.009 (n = 531, p = .830). The correlation for the conventional interface was .054 (n = 531, p = .214). The difference between the correlations (.063) was not statistically significant (p = .153, 1-tailed).

Hypothesis 19 predicts that when responses are obtained by means of the new interface, the mean response times for the forecasting questions will be shorter among those responding correctly to interpretation questions than among those responding incorrectly to such questions, and that this difference will not be observable when responses are obtained by means of the conventional interface. This hypothesis was tested by conducting t-tests for the difference in forecasting question response times between correct and incorrect responders to the interpretation questions under the two interface conditions. The results of these t-tests are reported in Table 40.

Table 40 Results of t-tests of the difference in forecasting question response times between correct and incorrect interpretation question responders under the two interface conditions

Interface	Response time means		Difference			df	Sig. (2-tailed)
	Correct interpretation question responders	Incorrect interpretation question responders	Mean	Std. error of mean difference	t		
New	42.19	45.14	-2.946	2.519	-1.170 ^a	529	.243
Conventional	155.28	155.98	-.703	5.679	-.124 ^a	529	.902

Note: sample sizes for correct responders and incorrect responders were 437 and 94, respectively, for the new interface, and 184 and 347, respectively, for the conventional interface.

^a Levene's test not significant, equal variances assumed.

The results in Table 40 do not support the prediction of Hypothesis 19 and do not justify the rejection of the null hypothesis. The predicted lower response time for forecasting questions among those responding correctly to the interpretation questions in comparison to those responding incorrectly was not observed to occur under the new interface condition. As an alternative way to hypothesis 19, I tested this in a different approach. This alternative approach predicts that there will be a significant correlation between response times to the interpretation

and forecasting questions under the new interface condition, and that this correlation will be significantly higher than that between the corresponding response times obtained under the conventional interface condition. This hypothesis was tested by comparing the correlations between response times to the interpretation and forecasting questions obtained under each interface condition. The null hypothesis of no difference between the correlations obtained for two screen types could not be rejected. The correlation for the new interface was $-.002$ ($n = 531$, $p = .970$). The correlation for the conventional interface was $.027$ ($n = 531$, $p = .528$). The difference between the correlations ($.029$) was not statistically significant ($p = .319$, 1-tailed).

Confidence Analysis

Hypotheses 20 to 30 focused on the dependent variable of confidence in one's responses to the questions posed by means of the new and conventional interfaces.

Confidence Related Analysis for All SA Levels

Hypothesis 20 holds that users' confidence in their responses to perception questions differs according to the interface used to elicit their responses. This hypothesis was tested by conducting a 1-within, 1-between ANOVA on confidence ratings, where the within-factor was the interface type and the between factor was the user i.d. The user i.d. was included as a factor to remove the effect of differences between subjects due to differences in the scenario, session, and particular question used in eliciting responses to the perception questions. The results of this ANOVA are reported in Table 41.

Table 41 Results of ANOVA of confidence ratings of responses to perception questions by interface type and user

Source	df	F	Sig.	η^2
User	58	1.185	.176	.127
Interface	1	5.219	.023	.010
Interface * user	58	.808	.841	.089
Between-subjects error	472	(5.703)		
Within-subjects error	472	(5.153)		

Note: Values in parentheses are mean square errors.

The results in Table 41 indicate that the null hypothesis of no difference between interface types in mean confidence rating of responses to perception questions must be rejected. Subjects using the new interface exhibited a significantly higher mean confidence rating of their responses to the perception questions than did subjects using the conventional interface (6.695 vs. 6.377).

Hypothesis 21 holds that users' confidence in their responses to interpretation questions differs according to the interface used to elicit their responses. This hypothesis was tested by conducting a 1-within, 1-between ANOVA on confidence ratings, where the within-factor was the interface type and the between factor was the user i.d. The user i.d. was included as a factor in order to remove the effect of differences between subjects due to differences in the scenario, session, and particular question used in eliciting responses to the interpretation questions. The results of this ANOVA are reported in Table 42.

Table 42 Results of ANOVA of confidence ratings of responses to interpretation questions by interface type and user

Source	Df	F	Sig.	η^2
user	58	1.160	.207	.125
Interface	1	882.919	.023	.617
Interface * user	58	1.301	.841	.053
Between-subjects error	472	(4.415)		
Within-subjects error	472	(4.517)		

Note: Values in parentheses are mean square errors.

The results in Table 42 indicate that the null hypothesis of no difference between interface types in mean confidence rating of responses to interpretation questions must be rejected. Subjects using the new interface exhibited a significantly higher mean confidence rating of their responses to the interpretation questions than did subjects using the conventional interface (6.944 vs. 3.068).

Hypothesis 22 holds that users' confidence in their responses to forecasting questions differs according to the interface used to elicit their responses. This hypothesis was tested by conducting a 1-within, 1-between ANOVA on confidence ratings, where the within-factor was the interface type and the between factor was the user i.d. The user i.d. was included as a factor in order to remove the effect of differences between subjects due to differences in the scenario, session, and particular question used in eliciting responses to the forecasting questions. The results of this ANOVA are reported in Table 43.

Table 43 Results of ANOVA of confidence ratings of responses to forecasting questions by interface type and user

Source	Df	F	Sig.	η^2
user	58	.785	.873	.088
Interface	1	1309.676	<.001	.716
Interface * user	58	.811	.838	.0004
Between-subjects error	472	(3.970)		
Within-subjects error	472	(3.991)		

Note: Values in parentheses are mean square errors.

The results in Table 43 indicate that the null hypothesis of no difference between interface types in mean confidence rating of responses to forecasting questions must be rejected. Subjects using the new interface exhibited a significantly higher mean confidence rating of their responses to the forecasting questions than did subjects using the conventional interface (6.821 vs. 2.384).

Confidence Related Analysis by response type for SA Level 1

Hypothesis 23 states that mean confidence ratings differ between correct and incorrect responses for perception questions in each interface. This hypothesis was tested by conducting two ANOVAs, one for each interface, using confidence ratings of responses to perception questions as the dependent variable and correctness of response and subject as the independent variables. Subject was included as a factor in order to remove the effect of differences between subjects due to differences in the scenario, session, and particular question used in eliciting responses to the perception questions. The ANOVA results are summarized in Table 44.

Table 44 Results of ANOVA of confidence ratings of responses to perception questions by correctness of response and subject for each interface

Interface	Source	df	F	Sig.	η^2
New	Response correctness	1	83.249	<.001	.137
	Subject	58	.904	.674	.086
	Response correctness * subject	41	1.098	.318	.074
	Error	430	(4.378)		
Conventional	Response correctness	1	140.055	<.001	.218
	Subject	58	.932	.618	.084
	Response correctness * subject	42	.493	.997	.032
	Error	429	(4.356)		

Note: Values in parentheses are mean square errors.

The results in Table 44 indicate that the null hypothesis of no difference in response times between correct and incorrect responders using either interface must be rejected. Correct responders using the new interface exhibited a significantly higher mean confidence rating than that for incorrect responders (7.118 vs. 4.536 on a scale of 1 to 10). Correct responders using the conventional interface also exhibited a significantly higher mean confidence rating than that for incorrect responders. (6.856 vs. 3.374 on a scale of 1 to 10). Thus, the direction of the difference was consistent between the two interfaces and the degree of difference within each interface was both statistically significant and quite large in practical terms. Considering only the 380 instances of subjects giving correct responses to a perception question using both interfaces, the mean

confidence ratings were 7.13 using the new interface and 6.92 using the conventional interface, which was a nonsignificant difference (t for difference = 1.450, $p = .148$).

Confidence Related Analysis by response type for SA Level 2

Hypothesis 24 states that mean confidence ratings differ between correct and incorrect responses for interpretation questions within each interface. This hypothesis was tested by conducting two ANOVAs, one for each interface, using confidence ratings of responses to interpretation questions as the dependent variable and correctness of response and subject as the independent variables. Subject was included as a factor in order to remove the effect of differences between subjects due to differences in the scenario, session, and particular question used in eliciting responses to the interpretation questions. The ANOVA results are summarized in Table 45.

Table 45 Results of ANOVA of confidence ratings of responses to interpretation questions by correctness of response and subject for each interface

Interface	Source	df	F	Sig.	η^2
New	Response correctness	1	218.250	<.001	.283
	Subject	58	1.138	.237	.086
	Response correctness * subject	40	1.415	.053	.073
	Error	431	(3.119)		
Conventional	Response correctness	1	451.128	<.001	.440
	Subject	58	1.364	.047	.077
	Response correctness * subject	52	1.457	.026	.078

Interface	Source	df	F	Sig.	η^2
	Error	419	(1.902)		

Note: Values in parentheses are mean square errors.

The results in Table 45 indicate that the null hypothesis of no difference in response times between correct and incorrect responders using either interface must be rejected. Correct responders using the new interface exhibited a significantly higher mean confidence rating than that for incorrect responders (7.588 vs. 4.179 on a scale of 1 to 10). Correct responders using the conventional interface also exhibited a significantly higher mean confidence rating than that for incorrect responders. (5.076 vs. 1.987 on a scale of 1 to 10). Thus, the direction of the difference was consistent between the two interfaces and the degree of difference within each interface was both statistically significant and quite large in practical terms. Considering only the 152 instances of subjects giving correct responses to an interpretation question using both interfaces, the mean confidence ratings were 7.67 using the new interface and 5.09 using the conventional interface, which was a significant difference (t for difference = 11.227, $p < .001$).

Confidence Related Analysis by response type for SA Level 3

Hypothesis 25 states that mean confidence ratings differ between correct and incorrect responses for forecasting questions within each interface. This hypothesis was tested by conducting two ANOVAs, one for each interface, using confidence ratings of responses to forecasting questions as the dependent variable and correctness of response and subject as the independent variables. Subject was included as a factor in order to remove the effect of differences between subjects due to differences in the scenario, session, and particular question used in eliciting responses to the forecasting questions. The ANOVA results are summarized in Table 46.

Table 46 Results of ANOVA of confidence ratings of responses to forecasting questions by correctness of response and subject for each interface

Interface	Source	df	F	Sig.	η^2
New	Response correctness	1	61.863	<.001	.100
	Subject	58	1.212	.148	.114
	Response correctness * subject	43	1.358	.071	.094
	Error	428	(4.111)		
Conventional	Response correctness	1	1503.568	<.001	.716
	Subject	58	1.537	.010	.042
	Response correctness * subject	55	1.633	.004	.043
	Error	416	(.664)		

Note: Values in parentheses are mean square errors.

The results in Table 46 indicate that the null hypothesis of no difference in response times between correct and incorrect responders using either interface must be rejected. Correct responders using the new interface exhibited a significantly higher mean confidence rating than that for incorrect responders (7.247 vs. 5.367 on a scale of 1 to 10). Correct responders using the conventional interface also exhibited a significantly higher mean confidence rating than that for incorrect responders (4.916 vs. 1.486 on a scale of 1 to 10). Thus, the direction of the difference was consistent between the two interfaces and the degree of difference within each interface was both statistically significant and quite large in practical terms. Considering only the 107 instances of subjects giving correct responses to a forecasting question using both interfaces, the mean

confidence ratings were 6.92 using the new interface and 4.92 using the conventional interface, which was a significant difference (t for difference = 9.001, $p < .001$).

Impact of Age, Experience and Edu on Confidence for SA Level 1 Questions

Hypothesis 26 posited that within each interface, confidence ratings of responses to perception questions are related to age, years of experience, and educational background. The test of this hypothesis with respect to age was conducted by the application of a t-test of the difference in mean confidence ratings of responses to the perception questions averaged within subject for subjects under 40 years of ages and for subjects 40 or more years of age. This test was performed separately for the two interfaces. The results of these tests are shown in Table 47.

Table 47 Results of t-test of difference between under 40 and 40 or older age groups in mean confidence ratings of responses to perception questions

Interface	Confidence rating means		Difference		t	df	Sig. (2-tailed)
	Under 40	40 or older	Mean	Std. error of mean difference			
New	6.94	6.44	.493	.182	2.714 ^a	57	.009
Conventional	6.37	6.38	-.013	.216	-.059 ^a	57	.953

^a Levene’s test not significant, equal variances assumed.

The results in Table 42 indicate that the differences between age groups in the mean confidence ratings of their responses to perception questions reached statistical significance in the new interface but not with the conventional interface. Consequently, the null hypothesis of no difference between age groups in the mean confidence ratings of their responses to perception questions in either interface is rejected only under the new interface condition. In the use of the

new interface, subjects under 40 years of age reported higher mean confidence ratings in their responses to the perception questions than did subjects 40 years of age or older.

The test of hypothesis 26 with respect to occupational experience levels was conducted by means of an analysis of variance of the responses obtained with each interface using a 4-level categorization of occupational experience (i.e., 1, 2, 3, and 4 or more years of experience) as the ANOVA factor and confidence ratings of responses to perception questions averaged within subject as the dependent variable. The result of these ANOVAs are reported in Table 48.

Table 48 Results of analysis of variance of perception question confidence ratings by occupational experience levels for the new and conventional interfaces

Interface	Source	df	F	Sig.	η^2
			3.059	.036	
New	Occupational experience level	3			.143
	Error	55	(.487)		
			.550	.650	
Conventional	Occupational experience level	3			.034
	Error	55	(.029)		

Note: Values in parentheses are mean square errors.

The results in Table 48 indicate that the differences between occupational experience levels in the mean confidence ratings of their responses to perception questions only reached statistical significance under the new interface condition. Consequently, the null hypothesis of no difference between occupational experience levels in the mean confidence ratings of their responses to perception questions in either interface can only be rejected under the new interface condition. Examination of the means for the different experience levels within the new interface condition indicated a tendency for the lowest (1 year) and highest (4 or more years) experience

levels to have higher mean confidence ratings, but none of the pairwise comparisons of experience levels reached statistical significance in the post hoc tests.

The test of Hypothesis 26 with respect to educational background was conducted by the application of a t-test of the difference in mean confidence ratings of responses to the perception questions for subjects having an IT/Informatics educational background and those having a Public Health or other type of educational background. This test was performed separately for the two interfaces. The results of these tests are shown in Table 49.

Table 49 Results of t-test of difference between IT/Informatics vs. other educational background groups in mean confidence ratings of responses to perception questions

Interface	Response time means		Difference		t	df	Sig. (2-tailed)
	IT/Informatics education	Other education	Mean	Std. error of mean difference			
New	6.65	6.73	-.084	.193	-.436 ^a	57	.665
Conventional	6.38	6.37	.008	.216	.038 ^a	57	.970

Note: sample sizes for IT/Informatics and Other educational backgrounds were 28 and 31, respectively for both the new and conventional interfaces.

^a Levene's test not significant, equal variances assumed.

The results in Table 49 indicate that the differences between the two educational background groups in their mean confidence ratings of responses to perception questions did not reach statistical significance in either interface. Consequently, the null hypothesis of no difference between the two educational background groups examined in this study in the mean confidence ratings of their responses to perception questions in either interface cannot be rejected.

Impact of Age, Experience and Edu on Confidence for SA Level 2 Questions

Hypothesis 27 posited that within each interface, confidence ratings of responses to interpretation questions are related to age, years of experience, and educational background. The test of this hypothesis with respect to age was conducted by the application of a t-test of the difference in mean confidence ratings of responses to the interpretation questions averaged within subject for subjects under 40 years of ages and for subjects 40 or more years of age. This test was performed separately for the two interfaces. The results of these tests are shown in Table 50.

Table 50 Results of t-test of difference between under 40 and 40 or older age groups in mean confidence ratings of responses to interpretation questions

Interface	Confidence rating means		Difference		t	df	Sig. (2-tailed)
	Under 40	40 or older	Mean	Std. error of mean difference			
New	6.89	7.00	-.119	.196	-.606	57	.547
Conventional	3.06	3.08	-.025	.214	-.116	57	.908

^a Levene's test not significant, equal variances assumed.

The results in Table 50 indicate that the differences between age groups in the mean confidence ratings of their responses to interpretation questions did not reach statistical significance under either interface condition. Consequently, the null hypothesis of no difference between age groups in the mean confidence ratings of their responses to interpretation questions in either interface cannot be rejected.

The test of hypothesis 27 with respect to occupational experience levels was conducted by means of an analysis of variance of the responses obtained with each interface using a 4-level

categorization of occupational experience (i.e., 1, 2, 3, and 4 or more years of experience) as the ANOVA factor and confidence ratings of responses to interpretation questions averaged within subject as the dependent variable. The result of these ANOVAs are reported in Table 51.

Table 51 Results of analysis of variance of interpretation question confidence ratings by occupational experience levels for the new and conventional interfaces

Interface	Source	df	F	Sig.	η^2
New	Occupational experience level	3	2.643	.058	.126
	Error	55	(.515)		
Conventional	Occupational experience level	3	1.272	.293	.065
	Error	55	(.654)		

Note: Values in parentheses are mean square errors.

The results in Table 51 indicate that the differences between occupational experience levels in the mean confidence ratings of their responses to interpretation questions did not reach statistical significance under either interface condition. Consequently, the null hypothesis of no difference between occupational experience levels in the mean confidence ratings of their responses to interpretation questions in either interface cannot be rejected.

The test of Hypothesis 27 with respect to educational background was conducted by the application of a t-test of the difference in mean confidence ratings of responses to the interpretation questions for subjects having an IT/Informatics educational background and those having a Public Health or other type of educational background. This test was performed separately for the two interfaces. The results of these tests are shown in Table 52.

Table 52 Results of t-test of difference between IT/Informatics vs. other educational background groups in mean confidence ratings of responses to interpretation questions

Interface	Response time means		Difference		t	Df	Sig. (2-tailed)
	IT/Informatics education	Other education	Mean	Std. error of mean difference			
New	6.96	6.93	.032	.196	.163 ^s	57	.871
Conventional	2.97	3.15	-.182	.213	-.854 ^s	57	.396

Note: sample sizes for IT/Informatics and Other educational backgrounds were 28 and 31, respectively for both the new and conventional interfaces.

^a Levene's test not significant, equal variances assumed.

The results in Table 52 indicate that the differences between the two educational background groups in their mean confidence ratings of responses to interpretation questions did not reach statistical significance in either interface. Consequently, the null hypothesis of no difference between the two educational background groups examined in this study in the mean confidence ratings of their responses to interpretation questions in either interface cannot be rejected.

Impact of Age, Experience and Edu on Confidence for SA Level 3 Questions

Hypothesis 28 posited that within each interface, confidence ratings of responses to forecasting questions are related to age, years of experience, and educational background. The test of this hypothesis with respect to age was conducted by the application of a t-test of the difference in mean confidence ratings of responses to the forecasting questions averaged within subject for subjects under 40 years of ages and for subjects 40 or more years of age. This test was performed separately for the two interfaces. The results of these tests are shown in Table 53.

Table 53 Results of t-test of difference between under 40 and 40 or older age groups in mean confidence ratings of responses to forecasting questions

Interface	Confidence rating means		Difference		t	df	Sig. (2-tailed)
	Under 40	40 or older	Mean	Std. error of mean difference			
New	6.84	6.80	.032	.175	.186 ^a	57	.853
Conventional	2.39	2.38	.017	.135	.127 ^a	57	.899

^a Levene's test not significant, equal variances assumed.

The results in Table 53 indicate that the differences between age groups in the mean confidence ratings of their responses to forecasting questions did not reach statistical significance in either interface. Consequently, the null hypothesis of no difference between age groups in the mean confidence ratings of their responses to forecasting questions in either interface cannot be rejected.

The test of hypothesis 28 with respect to occupational experience levels was conducted by means of an analysis of variance of the responses obtained with each interface using a 4-level categorization of occupational experience (i.e., 1, 2, 3, and 4 or more years of experience) as the ANOVA factor and confidence ratings of responses to forecasting questions averaged within subject as the dependent variable. The result of these ANOVAs are reported in Table 54.

Table 54 Results of analysis of variance of forecasting question confidence ratings by occupational experience levels for the new and conventional interfaces

Interface	Source	df	F	Sig.	η^2
New	Occupational experience level	3	1.643	.190	.082
	Error	55	(.428)		
Conventional	Occupational experience level	3	.797	.501	.042
	Error	55	(.266)		

Note: Values in parentheses are mean square errors.

The results in Table 54 indicate that the differences between occupational experience levels in the mean confidence ratings of their responses to forecasting questions did not reach statistical significance under either interface condition. Consequently, the null hypothesis of no difference between occupational experience levels in the mean confidence ratings of their responses to forecasting questions in either interface cannot be rejected.

The test of Hypothesis 28 with respect to educational background was conducted by the application of a t-test of the difference in mean confidence ratings of responses to the forecasting questions for subjects having an IT/Informatics educational background and those having a Public Health or other type of educational background. This test was performed separately for the two interfaces. The results of these tests are shown in Table 55.

Table 55 Results of t-test of difference between IT/Informatics vs. other educational background groups in mean confidence ratings of responses to forecasting questions

Interface	Response time means		Difference		t	df	Sig. (2-tailed)
	IT/Informatics education	Other education	Mean	Std. error of mean difference			
New	6.65	6.73	-.084	.193	-.436 ^a	57	.665
Conventional	6.38	6.37	.008	.216	.038 ^a	57	.970

Note: sample sizes for IT/Informatics and Other educational backgrounds were 28 and 31, respectively for both the new and conventional interfaces.

^a Levene's test not significant, equal variances assumed.

The results in Table 55 indicate that the differences between the two educational background groups in their mean confidence ratings of responses to forecasting questions did not reach statistical significance in either interface. Consequently, the null hypothesis of no difference between the two educational background groups examined in this study in the mean confidence ratings of their responses to forecasting questions in either interface cannot be rejected.

Confidence analysis by response type between SA Levels 1 and 2

Hypothesis 29 predicts that when responses are obtained by means of the new interface, the mean confidence ratings for the interpretation questions will be higher among those responding correctly to perception questions than among those responding incorrectly to such questions, and that this difference will not be observable when responses are obtained by means of the conventional interface. This hypothesis was tested by conducting t-tests for the difference in the means of interpretation question confidence ratings between correct and incorrect responders to the perception questions under the two interface conditions. The results of these t-tests are reported in Table 56.

Table 56 Results of t-tests of the difference in mean confidence ratings for interpretation questions between correct and incorrect perception question responders under the two interface conditions

Interface	Confidence rating means		Difference		t	df	Sig. (2-tailed)
	Correct perception question responders	Incorrect perception question responders	Mean	Std. error of mean difference			
New	7.06	6.36	.696	.296	2.354	108.657	.020
Conventional	3.10	2.86	.236	.257	.916	529	.360

Note: sample sizes for correct responders and incorrect responders were 445 and 86, respectively, for the new interface, and 457 and 74, respectively, for the conventional interface.
^a Levene's test not significant, equal variances assumed.

The results in Table 56 support the prediction of Hypothesis 29 and justify the rejection of the null hypothesis. The predicted higher mean confidence ratings of responses to interpretation questions among those responding correctly to the perception questions in comparison to those responding incorrectly was observed to occur under the new interface condition but not under the conventional interface condition.

Hypothesis 29 predicts that there will be a significant correlation between confidence ratings for responses to the perception and interpretation questions under the new interface condition, and that this correlation will be significantly higher than that between the corresponding confidence ratings obtained under the conventional interface condition. This hypothesis was tested by comparing the correlations between confidence ratings of responses to the perception and interpretation questions obtained under each interface condition. The null hypothesis of no difference between the correlations obtained for two screen types could not be rejected. The correlation for the new interface was -.024 (n = 531, p =.579). The correlation for the

conventional interface was .024 (n = 531, p = .584). The difference between the correlations (.048) was not statistically significant (p = .218, 1-tailed).

Confidence analysis by response type between SA Levels 2 and 3

Hypothesis 30 predicts that when responses are obtained by means of the new interface, the mean confidence ratings of responses to the forecasting questions will be higher among those responding correctly to interpretation questions than among those responding incorrectly to such questions, and that this difference will not be observable when responses are obtained by means of the conventional interface. This hypothesis was tested by conducting t-tests for the difference in mean confidence ratings of responses to forecasting questions between correct and incorrect responders to the interpretation questions under the two interface conditions. The results of these t-tests are reported in Table 57.

Table 57 Results of t-tests of the difference in forecasting question response times between correct and incorrect interpretation question responders under the two interface conditions

Interface	Response time means		Difference		t	df	Sig. (2-tailed)
	Correct interpretation question responders	Incorrect interpretation question responders	Mean	Std. error of mean difference			
New	6.77	7.06	-.295	.248	-1.190	529	.235
Conventional	2.53	2.31	.219	.158	1.381	529	.168

Note: sample sizes for correct responders and incorrect responders were 437 and 94, respectively, for the new interface, and 184 and 347, respectively, for the conventional interface.
^a Levene’s test not significant, equal variances assumed.

The results in Table 57 do not support the prediction of Hypothesis 30 and do not justify the rejection of the null hypothesis. The predicted higher mean confidence rating for forecasting questions among those responding correctly to the interpretation questions in comparison to

those responding incorrectly was not observed to occur under the new interface condition. Moreover, rather than higher mean confidence ratings in the correct interpretation response condition being more evident with the use of the new interface, if anything this prediction was more nearly satisfied with the use of the conventional interface condition.

Hypothesis 30 predicts that there will be a significant correlation between confidence ratings of responses to the interpretation and forecasting questions under the new interface condition, and that this correlation will be significantly higher than that between the corresponding confidence ratings obtained under the conventional interface condition. This hypothesis was tested by comparing the correlations between confidence ratings of the interpretation and forecasting questions obtained under each interface condition. The null hypothesis of no difference between the correlations obtained for two screen types could not be rejected. The correlation for the new interface was $-.041$ ($n = 531, p = .350$). The correlation for the conventional interface was $-.013$ ($n = 531, p = .772$). The difference between the correlations ($.028$) was not statistically significant ($p = .324, 1$ -tailed).

Chapter Summary

In this chapter, I presented the pilot study and the final study that I conducted to evaluate user's situation awareness when performing a signal characterization task. From the study I demonstrated that the situational awareness levels, perception, interpretation and forecasting are higher when the user is treated with the new interface when compared to the conventional interface (Hypotheses 1-3). The response time was significantly shorter when using the new interface (Hypotheses (9-11)). It is also evident from the results that the users exhibited higher level of confidence when using the new interface compared to the conventional interface (Hypotheses 20-22). Apart from measuring and comparing the SA, further analysis was done to

see the interaction between the levels in either interface and was compared to see if the difference between the two was significantly different. It was evident that a strong association was found between perception and interpretation and interpretation and forecasting when the user was responding using the new interface when compared the conventional interface (Hypotheses 7 and 8). Similar association was found in response time (Hypotheses 18 and 19). However similar association was not found in confidence rating (Hypotheses 29 and 30). In the next chapter, I will analyze the results in detail and discuss the potential impact of these findings.

CHAPTER 5: DISCUSSION

Introduction

In Chapter 1, I stated my research aims as “Develop a method to build dashboard systems that will meet user’s SA requirements” and “Prototype a health information dashboard using the new method”. In Chapter 3, I introduced a hybrid approach to building a dashboard system. As an implementation of the design, a prototype dashboard was developed in the domain of public health electronic message exchange. An evaluation study was designed and in Chapter 4, I presented the evaluation study and analyzed the results. This chapter presents an in-depth analysis of the MeRCI design and its conceptualization. Discussions focus on the design rationale and outcomes of the evaluation in light of the desiderata put forward in Chapter 1 for the next generation information representation, the gap analysis provided in Chapter 2, and the motivations introduced in Chapter 3. Advantages and known limitations of the system are discussed, and its implications for different research areas are explained.

MeRCI – Evaluation for Situational Awareness

The purpose of this study is to develop a framework to create and present contextualized information in order to improve users’ situation awareness (ability to perceive, interpret and forecast) when using a public health information system dashboard. To measure the improvement in SA, I designed a study to evaluate SA among the users of the system while performing a specific task. To understand the results of the study, I will discuss the various hypotheses and analyze the results here.

Our first goal is to look at how users performed in each of the three levels in terms of their responses. The study finds that the user perception levels between the two interfaces were not particularly significant (See Figure 29). The high degree of correctness in responses here and the

insignificant difference between the interfaces goes to prove two major points: (a) Users are not biased by their previous experience with one of the interfaces (2) Users were able to perceive data from both representations without much difficulty.

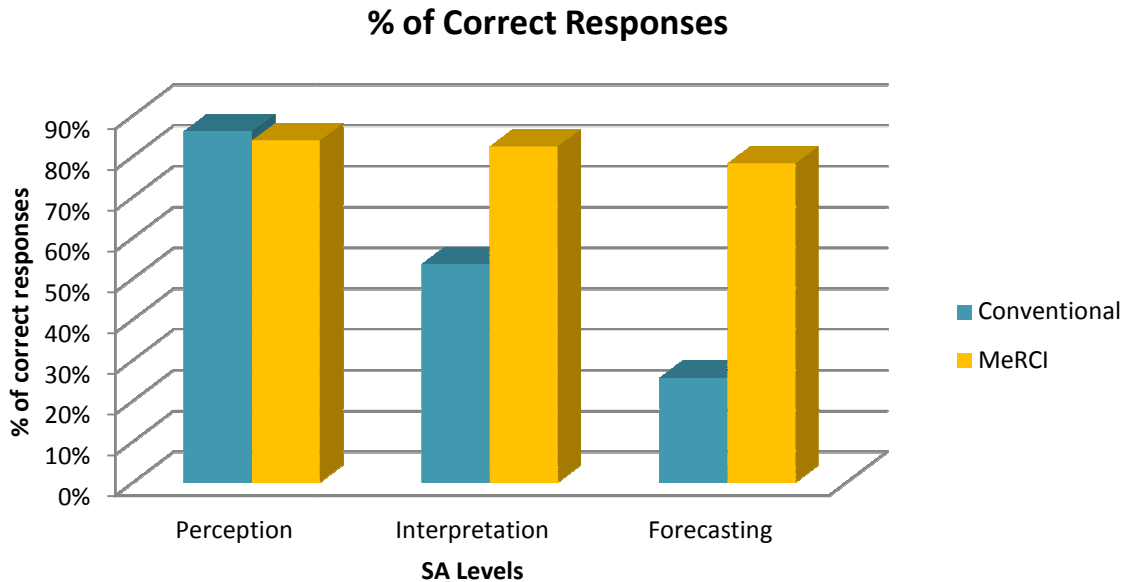


Figure 29 User SA levels measured by response type when using both the interfaces

However, there was a difference between the two interfaces in the rates of correct responses to interpretation questions. The users of the MeRCI based interface had a higher correct response rate compared to the conventional interface. A similar result was also found for forecasting questions. To confirm that these differences were not because of variables like age, years of experience and educational background, further analysis was done. The results show that there was no impact of these variables on any SA levels. These findings lead us to understand that by presenting goal-oriented information that is contextualized in terms of domain knowledge, this will improve all the 3 levels of awareness. To further strengthen this claim, an analysis was done to find the interaction between the levels (perception and interpretation) and (interpretation and

forecasting). The test was conducted by verifying whether the relationship between the correctness of responses to the perception and interpretation questions was stronger for the new interface than for the conventional one (Hypothesis 7). It was found that, with the new interface, a significantly stronger positive relationship was detected between the correctness of subjects' responses to the perception and interpretation questions. Similarly, the use of the new interface resulted in a significantly stronger positive relationship between the correctness of subjects' responses to the interpretation and forecasting questions (Hypothesis 8). This finding can be attributed directly to the philosophy of supporting Level 2 and Level 3 in representation and by aiding the mental model of the user for the particular task.

In order to understand the role of the contextualized information design approach, another variable was also analyzed. Total time taken by the study participants in responding to each question was collected and analyzed. This variable can help us look into how soon the user was able to map the information to a schema. Figure 30 describes the results of the analysis on response time variable for all three SA levels. It was found that the user response time for perception, interpretation and forecasting questions was significantly shorter when using the new interface. This could be attributed to the availability of goal-specific information for both level 2 and level 3 in the new interface.

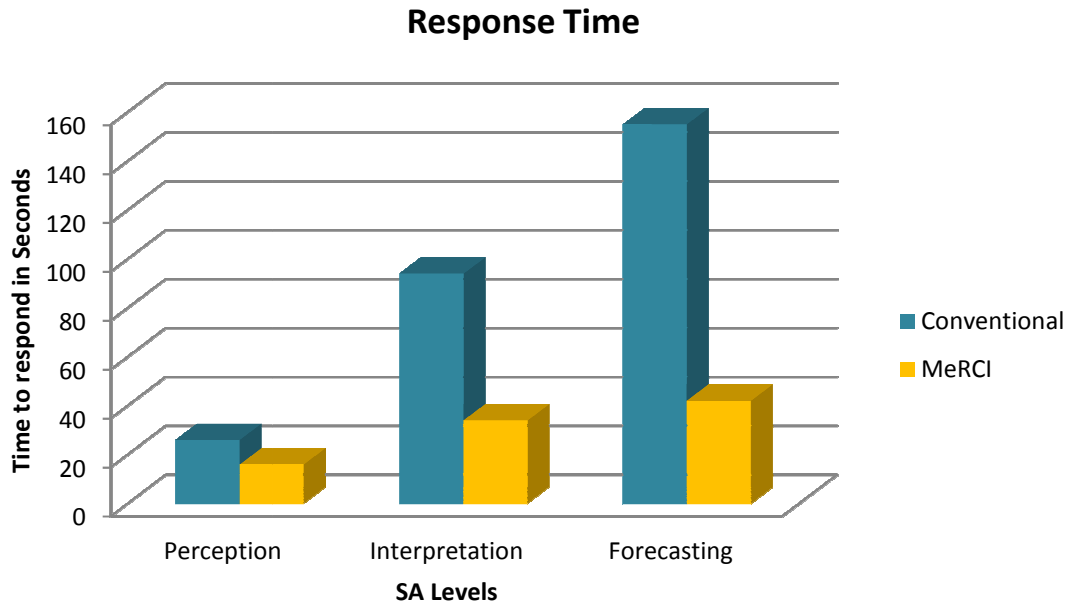


Figure 30 User Response Time in Seconds for all SA Levels

In order to understand if there is any relationship between responding to question correctly or incorrectly and response time, I conducted some analysis on all three SA levels. The results showed that correct responses to perception questions using the new interface exhibited a significantly lower mean response time than incorrect ones; however the correct responses using

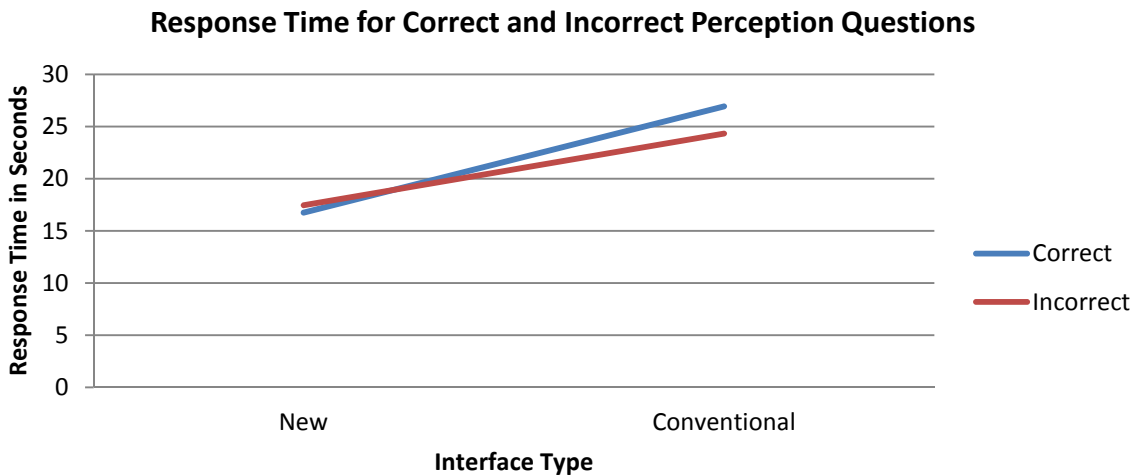


Figure 31 Comparison of Response Time for Correct and Incorrect Responses for Perception questions

the conventional interface had a significantly higher mean response time than incorrect ones. While the new interface response time was generally shorter, what was interesting was to find that in the conventional interface, the users were faster to respond incorrectly. This could be a detriment to achieving user goals. A review of the conventional interface showed that there was some semantic misunderstanding among many users on information that was displayed.

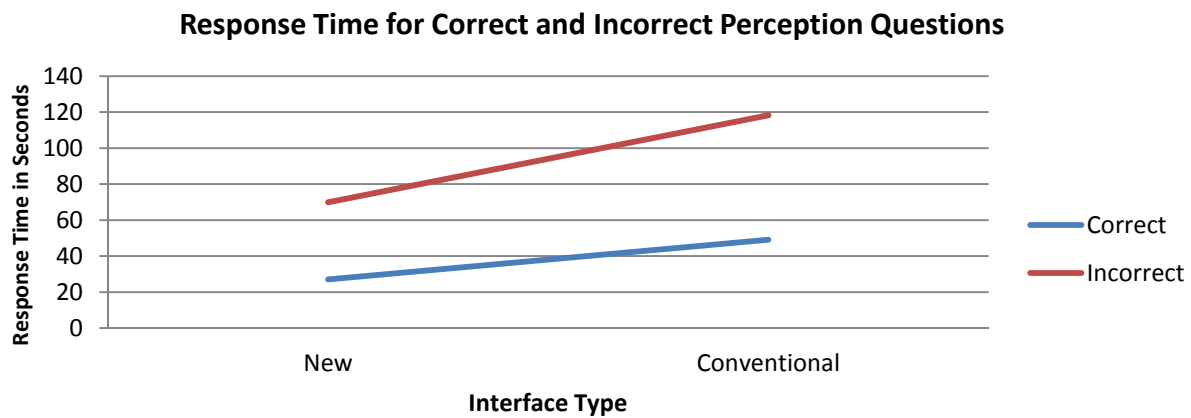


Figure 32 Comparison of Response Time for Correct and Incorrect Responses for Interpretation questions

Figures 32 and 33 shows that the response time for incorrect responses was significantly longer than correct responses in both interpretation and forecasting questions. To ensure that there was no effect of external variables like age, experience and educational background, I studied the impact on response time of these variables for all SA levels.

Results indicated that there was no interference of age or experience or educational background on the response time.

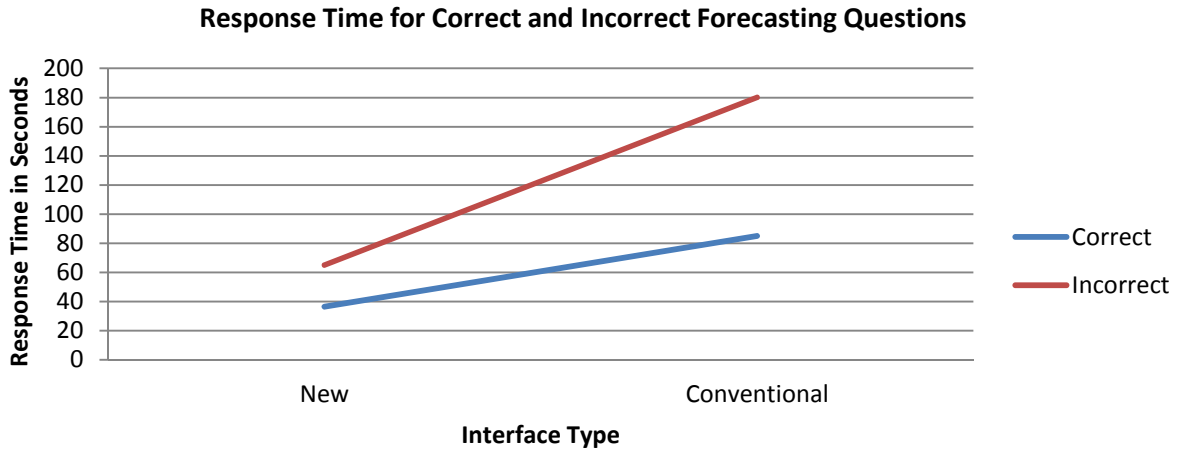


Figure 33 Comparison of Response Time for Correct and Incorrect Responses for forecasting questions

Confidence is a critical link between SA and performance (Christ & McKeever, 1994), If SA is good and confidence in that SA is high, a person is more likely to achieve a good outcome. So in this study the third variable measured was confidence level when the user is performing the task.

Figure 34 lays out the confidence levels of users when responding to

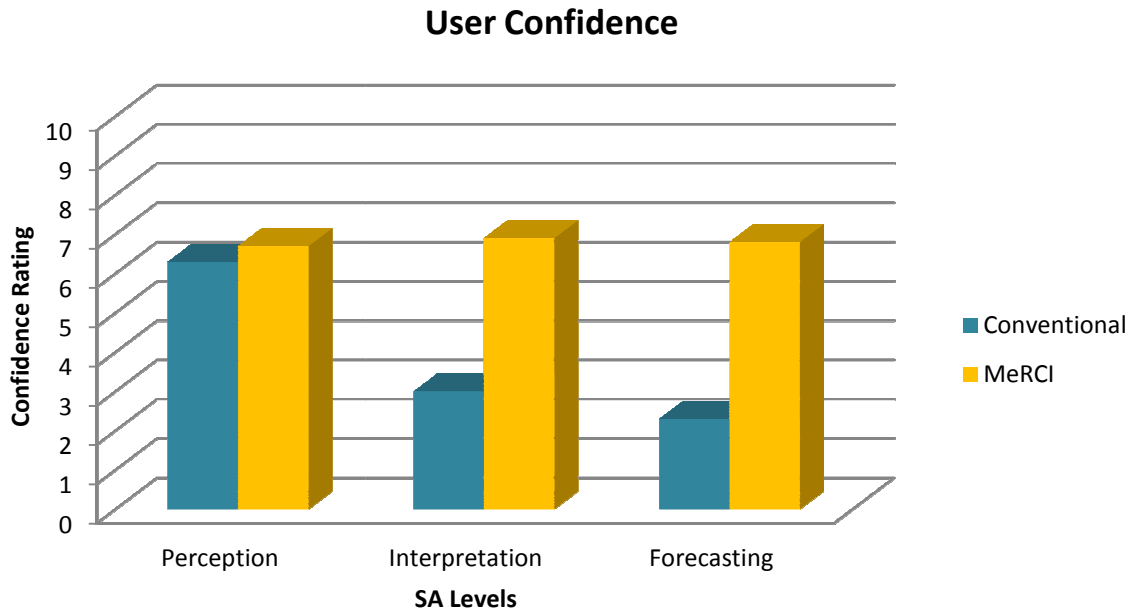


Figure 34 User Confidence Level when responding to different SA levels

SA levels 1, 2 and 3. It was evident that the users had a higher level of confidence when working with the new interface than with the conventional interface. This also maps back to the rate of correctness among users. For perception, users expressed a higher level of confidence and the rate of correctness was also noticeably higher. In terms of interpretation and forecasting, the level of confidence also maps well with the rate of correctness. In order to understand whether there is any relationship between responding to question correctly or incorrectly and response confidence level, I conducted some analyses on all three SA levels. Refer to Figure 35 for user confidence when responding to perception question. It was found that correct responders using the new interface had a significantly higher mean confidence rating than incorrect ones.

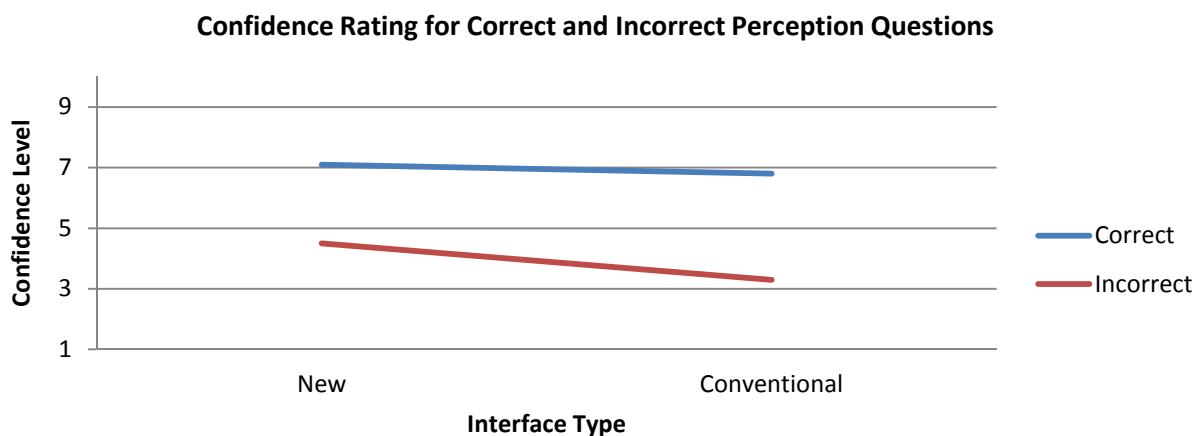


Figure 35 Confidence Rating for Correct and Incorrect responses to Perception Questions

Similarly, correct responders using the conventional interface also had a significantly higher mean confidence rating than incorrect responders. The degree of difference within each interface was both statistically significant and quite large in practical terms. Similar results were found in interpretation and forecasting questions. See Figures 36 and 37.

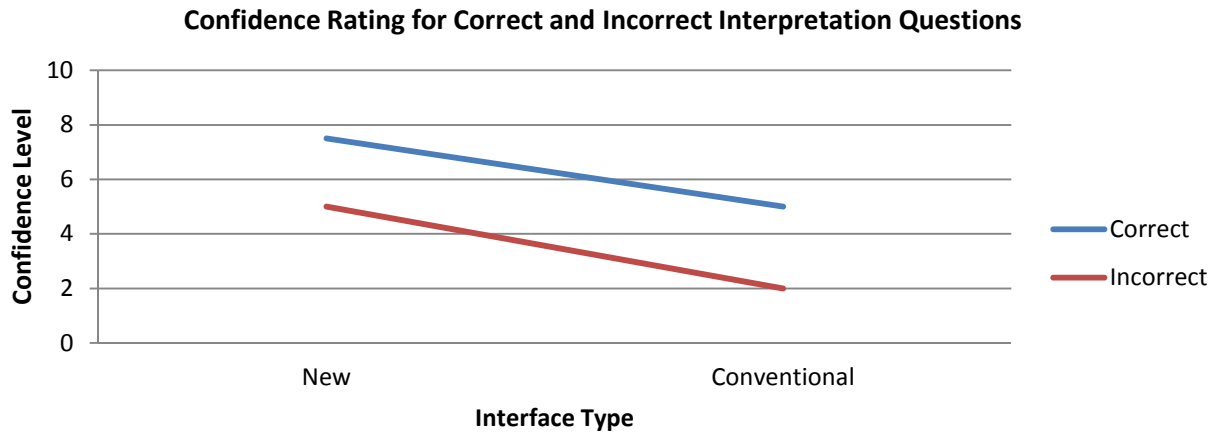


Figure 36 Confidence Rating for Correct and Incorrect responses to Interpretation Questions

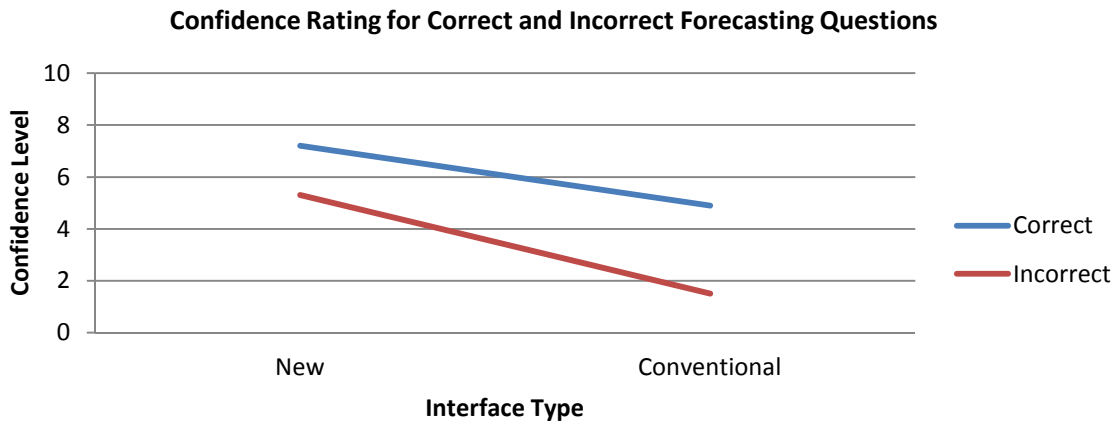


Figure 37 Confidence Rating for Correct and Incorrect Forecasting Questions

To find the correlation between the confidence levels when responding to perception and forecasting questions, further analysis was conducted. It was found that the predicted higher mean confidence ratings of responses to interpretation questions among those responding correctly to the perception questions in comparison to those responding incorrectly was observed for the new interface but not for the conventional interface. This justifies the expectation that a user who responded correctly to perception questions was confident in responding to interpretation questions. This relation was shown only in the new interface. Incidentally, no strong relation was found between those responding correctly to interpretation question and their

confidence level while responding to forecasting questions. Similarly no such relationship exists in the conventional interface.

Summary

In this section, I discussed the results of the evaluation study conducted to evaluate user's SA levels. In these studies, the results demonstrated that the new interface designed using MeRCI principles was significantly better than the conventional interface for all SA levels. In the following sections, I will discuss the strategies and lessons learnt in implementing a dashboard interface that will improve user SA during specific tasks with the system.

Guidelines for Building towards Situational Awareness

The way information is presented in the system will greatly influence the SA of the user. It directly influences the user's ability to acquire information in a limited time and the accuracy with which it can be gathered. It is also critical to match the users' mental schemata, thus reducing their cognitive and physical overload. In this section, I have discussed the findings of my research and developed guidelines for interface design. There are 20 principles proposed for designing a dashboard to deliver better SA.

- 1. Design around user goals:** Interfaces should aim at delivering information to support the user's goals. All information related to the goal should be co-located and should directly assist the goals of the user. The SA requirements identified using a CTA method should provide input about the information required to address each goal.
- 2. Account for changing goals:** Interfaces should allow users to switch from one goal to another. Designing to be goal driven will provide all information to aid user goals but when the goals of the user change, it must be possible to allow users to switch from one goal to another. This can be supported by providing appropriate location information on

the screen, using breadcrumbs, pagination buttons etc. Some researchers also call for a balance between goal-driven and data-driven approaches, allowing the value of the data to be taken into consideration while designing the interface.

- 3. Organize information available to support Levels 2 & 3:** When human data-processing capabilities are limited, undue load on human memory could lead to reduced efficiency in processing information. Hence interfaces should be designed to deliver all the information needed for perception and forecasting, for example by making historical trends and known correlated elements available on the interface instead of forcing the users to perform some tasks elsewhere to get to the data.
- 4. Make information explicit for Level 2:** In some cases calculations need to be done with the data (e.g. averaging), while in others the data need simply to be verified with another set of data (e.g. allowable range). In each of these situations, the information should be made more explicit instead of forcing the user to make the interpretation. This conversion from level 2 to level 1 can drastically improve awareness of the situation.
- 5. Reduce data overload:** Quite often designers find themselves caught up in a question of real estate on the screen versus information needed. The GDTA approach does not do a particularly good job in helping with filtering of information. The onus is on the designer to carefully filter the information needed for the goal. The information filtering activity is risky because it deprives users of SA, directly affecting global SA by affecting their proactive assessment of the situation. Only information that is truly not needed and redundant should be removed. All other information pertinent to the goal, regardless of its level of impact, should be carefully and meaningfully organized on the screen for the user to review.

- 6. Integrate information:** Bringing all information required for the user goal into a single frame or co-location is critical, removing the need for users to move from one page to another searching for information. Identify ways to have all information on a single screen. Often it is common to find only partial information for the goal is available in one system. All required information, even if it was originally in multiple systems needs to be pieced together to provide a complete picture. Lack of integration could involve unneeded workload and loss of awareness.
- 7. Explicitly indicate missing information:** It is common to find that sometimes information required is not available in a given situation. It is critical to signal this lack of data instead of hiding the unavailability. In some cases support information rather than critical information may be missing, e.g. allowable ranges of values. Alerts and values are unusable when such information is missing, often leading to delay in actively responding to information, thus affecting the SA of the user. People tend to operate under the assumption that operations are proceeding normally if information is not indicated as missing. This assumption could prove costly for their SA.
- 8. Include domain knowledge in the representation:** It is common among system designers to leave the users to apply operational and domain knowledge to make meaningful use of the data and interpret them. Presenting domain knowledge in a clear and easily digestible manner will allow users to meaningfully compare data with domain knowledge. This strategy will allow support of the SA levels, instead of the system making decisions for the user. Presenting domain knowledge will also help users refer to the information and be aware of changes as knowledge evolves. Allowing alert

mechanisms makes end users aware of new knowledge, and can be applied to the data in an effective manner.

- 9. Present critical information needed to trigger mental schemata:** During the GDTA it is critical to identify the cues that SMEs use to determine or assess the environment. These cues are generally used to activate the users' mental models and schemata. Designers should pay increased attention to identifying these cues and making them salient on the screen. Careful consideration needs to be given to the salience feature and its impact on the overall SA level.
- 10. Consider semantics of data to support SA:** This is a new paradigm shift that is proposed by this research where the value of the data and its context is taken into consideration while presenting the information. For example, events surrounding a source are presented only when the user goal deals with that source. The semantics of the data itself helps to filter some of the information. Appropriate measures needed to be taken to ensure that this information is available to support global SA and not lost during information filtering.
- 11. Provide SA support and avoid making decisions:** The overall effectiveness of an interface is measured by the level of SA available to the user and not by the quality of the decisions made by the user. Caution must be exercised in providing any forms of decision by the system. The system should provide users with all information necessary to affect their goals and allow them to make informed decisions. Studies have shown that user performance among SA delivery systems were found to be more effective than decision making systems, as the users were more aware of the environment and the domain and

could make more meaningful decisions. Questions about correct vs. incorrect decisions are beyond the scope of this research work.

12. Support Global SA at all times: One of the common problems noted in interface design is the loss of global SA. Often designers get lost in the design phase, especially when designing for specific goals, and lose track of keeping the end user globally aware of the environment. This is often referred to as attention narrowing. Designers should discourage this from happening while using the system and focus on providing global cues that will keep the user aware of the environment at any given time.

13. System functionality should address all user goals: The system designed should support all the user goals. The system should not allow the user to perform a goal with partial or wrong information if it was not designed for it. Often designers find systems being used for functions that they were not designed for. A complete assessment needs to be done to identify what other new functions the system needs to fully support. Appropriate design work needs to be included. A clear mapping should exist between the user goals and the system functions.

14. Adopt consistent representation across all goals: Information representation in the dashboard across goals should remain consistent. The user goals, decisions and the information types should determine the representation model. There needs to be standardization across the system in information representation. Having different forms of representation for the same dataset while using different goals will strain the users' cognitive ability. This could prove costly and redundant at times. Unless there is a need to view the data differently, consistency is to be maintained in all screens.

- 15. Allow systematic querying of information:** In a complex interface, it is essential that the users be allowed when necessary to further interact and ask for more information. The interface should allow systematic querying instead of tolerating an open box. This will lead to failure in SA by allowing user to get drowned in a wealth of information, which may not be critical for the goal.
- 16. Minimize the numbers of levels in logical branches:** It is crucial that the system does not enforce complexity on the users by allowing chaining of logical rules e.g. ‘if x then y unless z and n’. These complex operators will force the user to expend time and effort and challenge the user’s mental models. Sometimes it has been found that even if users develop a mental model with all the operators, applying it to real world information and interpreting it correctly was really difficult.
- 17. Implement an interface requiring less cognitive effort:** Excessive need to analyze data will slow down the search and retrieval of information. While extraneous data should be eliminated, caution should be implemented in designing an interface that is consistent with user goals. It is evident from studies that significant cost is included when users jump from one screen to other in search of relevant information. The solution is to maximize the organization of information so that dense information can be made readable using salience features, and group information for coherence for an in-depth review when required.
- 18. Interface should support uncertainty & higher levels of complexity:** It is only logical to design for uncertainty, but often system designers do not allow for this, as there is very little guidance for designing a system to support it. For SA levels to be higher, it is simply recommended that the system should continue to provide information to help

users determine whether the information confirms or negates the situation. Contextual data, as well as the semantics of the data, will play a critical role in providing background information. Designers should carefully review the process by which users handle uncertainty in the domain, as it varies by domain.

19. Use data salience property with caution: Use of salience enters into the realm of data-driven processing of information. A trade-off needs to be achieved between being data-driven and goal-driven, because there is a need for a data-driven model to direct the user to focus on high priority goals as they evolve. Use of salience in representing these cues is critical but caution needs to be established in choosing the level of salience. Especially when alerts are created, it is important to utilize salience in describing the severity of the alert instead of burying them in a list. These alerts need substantial physically salient features e.g. bright colors, trend markers, sounds etc. They are critical in activating certain goals of the users.

20. Minimize Task Complexity: The designers should consider reducing the number of steps required by the users to perform a desired activity with the system, such as the number of clicks needed to access information. Reducing the number of steps could improve user's experience with the interface and also improve performance by requiring the users to have to remember fewer steps and thereby reduce their cognitive workload. Reduced task complexity will lead to simpler mental models, which can be accessed easily later. Hence designers should consider building systems with lower task complexity.

Summary

In this section, I presented a general process and some key guidelines to consider while designing an interface. As a key driver for decision-making, it is really crucial for system designers to adopt these principles to build a system that will deliver SA. These suggestions are a good starting point but certainly there is a lot to be learnt about their impact on performance and decision-making. With the exponential growth of information and information systems supporting users, it is becoming critical that interface systems evolve to support the awareness levels of the users and guide them to appropriate elements in the environment needing attention. The SA design process and principles described here are based on the scientific evidence, theoretical and practical lessons learnt through the progress of the study. These principles are expected to provide guidelines to the system designer in creating interfaces that effectively support SA. In the next section, I describe the experiences in choosing a rule-based system.

Guidelines for choosing a Rule based System

The outcome of the concept mapping process described in Chapter 3 was the domain knowledge and operational rules that are applicable in a domain. In order to effectively utilize these rules and knowledge, a human and machine processable formal representation was required. In this section, I describe some of the rule-based systems that were reviewed and some guidelines to be considered in choosing a system.

A rules engine helps resolve (or at least reduce) the issues and difficulties inherent in the development and maintenance of an application's business logic, and can be considered as a framework for implementing complex business logic. Most rules engines allow users to use declarative programming to express the consequences that result from some information or knowledge. Users can concentrate on facts that are known to be true and their associated

outcomes — that is, on an application's business logic. If the business logic code includes a group of if-else statements, a rules engine should be considered. Maintaining complex Boolean logic can be a difficult task, and a rules engine can help organize this. Changes are significantly less likely to produce errors when the logic is expressed using a declarative rather than an imperative programming language. In my research I evaluated a few rules engines for application to this project. In this section, I will discuss the lessons learnt from using the tools Drools and openRules in this research work.

Drools: Drools is an Object-Oriented Rule Engine for Java. Adapting the Rete algorithm to an object-oriented interface allows for more natural expression of business rules with regard to business objects. Drools is not just a rule engine. It provides also an application for managing rules, called the Business Rules Management System (BRMS). It allows the designer to create, modify, delete, branch and persist rules. Moreover it is possible to assign roles to users, while a login mechanism and LDAP integration makes it easy to introduce security. The JBoss Enterprise BRMS includes a fast and highly efficient rule engine and easy-to-use rules development tools, management system and repository. Drools has the following features:

- **Business Rules Engine** - The rules engine implements the full Rete algorithm with high performance indexing and optimization.
- **Rules authoring** - The authoring interface enables fast and easy rules development, change and management for process owners, administrators and business analysts.
- **Rules management** – the JBoss Enterprise BRMS includes a business rules management repository and web-based administration console that helps business analysts, developers, administrators and other users of the BRMS to manage their rules.

Drools 5 introduce the business logic integration platform that provides a unified and integrated platform for rules, workflow and event processing. Drools is split up into 4 main sub-projects:

- ***Drools Guvnor*** is a centralized repository for Drools knowledge bases, with rich web based GUIs, editors, and tools to aid in the management of large numbers of rules.
- ***Drools Flow*** provides workflow or (business) process capabilities to the Drools platform. A business process or workflow describes the order in which a series of steps needs to be executed, using a flow chart.
- ***Drools Fusion*** is the module responsible for enabling event-processing capabilities that deal with the task of processing multiple events, with the goal of identifying the meaningful events within the event cloud.
- ***Drools Planner*** optimizes automated planning by using meta-heuristic algorithms, such as *tabu search* and *simulated annealing*.
- ***Drools Expert*** provides for declarative logic programming and is flexible enough to match the semantics of any problem domain. Currently rules can be written in Java, MVEL, Python and Groovy. It is designed to allow pluggable language implementations.

A Drools rules file has one or more rule declarations. Each rule declaration is composed of one or more *conditional* elements, and one or more *consequences* or *actions* to execute. A rules file can also have multiple (that is, zero or more) import declarations, multiple global declarations, and multiple function declarations. An example of a drools implementation is described in Appendix B. The other rule engine that was considered but not used in this research is OpenRules, which is open source software for Rules-based Application Development.

It efficiently uses the power of MS Excel, Google Docs, Eclipse IDE and open source Java libraries to create, deploy, execute, and maintain different rule engines with complex business logic controlled by business analysts. OpenRules goes beyond the traditional BRMS, covering not only business logic but also presentation logic. OpenRules supports rules-based interaction processes with a quick and intuitive GUI generation. Additionally, OpenRules integrates Business Rules with popular Machine Learning and Optimization tools. Table 58 provides a view of how some of the rules frameworks can support the formal representation requirements prescribed by MeRCI.

Frameworks	Features				
	Algorithm	Rules Engine	Rules Language	Rules Repository	Rules GUI
Jboss Drools	Rete algorithm	Yes	Yes	Yes	Yes
OpenRules	Optimization techniques: Constraint & Linear Programming	Yes	No	Yes	No
OpenL Tablets	Optimized sequential forward chaining algorithm	Yes	No	Yes	Yes
OpenCyc	Natural Language Processing	Yes	Yes	Yes	Yes
KAON	Uses web ontology language and frame logic	Yes	Yes	Yes	Yes

Table 58: Comparison of Rules Engine Features

Summary

In this section I have presented some of the research done in the domain of knowledge representation and rule engines. Representation languages like OWL and RDF could be used to express the domain knowledge in a very descriptive manner. In this research, however, I used

rules engines instead of representing the domain due to the nature of the problem for which features like inferencing are not required but complex rules are allowed. Table 58 provides a very high level comparison of some of the rule and knowledge representation frameworks that are used in the industry.

Chapter Summary

In this chapter, I discussed the outcomes of the evaluation study. Using the study results, I demonstrated that the interface developed using the new framework with contextualized information did in fact improve users' situation awareness (ability to perceive, interpret and forecast) during a signal characterization task. The users' SA levels, particularly interpretation and forecasting, were significantly high while using the new interface. I also demonstrated that the users were quicker to respond correctly with the new interface in all 3 SA levels and were more confident while responding to questions. In the later sections, I presented some of the principles for building a system interface that would improve awareness levels. I also presented guidelines for choosing a presentation framework for domain knowledge and rules. In the next chapter, I summarize the contributions of my thesis research, the limitations of my current approach and the possibilities of building on this research in the future.

CHAPTER 6: Summary & Conclusion

In this chapter, I summarize the design concepts for building a system interface to improve the users' situational awareness, discuss the contributions of my research, report on the limitations of my current approach, and present avenues for future work.

Contextualized Information Representation for Situational Awareness

This dissertation offers a new method for implementing a health information dashboard that is focused on delivering situational awareness to system users. I presented an implementation of a public health integration broker dashboard prototype using the MeRCI method. I explained how the MeRCI approach provides information about the system users' goals and how the information required for them are made available to users in a meaningful manner. The approach utilizes domain and operational knowledge to help users focus on the goals and presents information pertinent to these goals. The approach focuses on providing SA support rather than on decision-making.

An empirical study conducted on a prototype interface developed using MeRCI concepts demonstrated higher levels of situational awareness among users for all three SA levels, Perception, Interpretation and Forecasting. The study compared the user's response time and confidence when using the traditional system in comparison with the prototype system. It was demonstrated that the user's level of confidence, response time, as well as the correctness of response were all significantly higher when users pursued their goals using the prototype as compared to the old interface.

Users of the system consistently indicated after the study that the prototype interface provided them information relevant to their tasks and provided appropriate contextual data that was helpful in interpreting information or forecasting future impacts. They thought that the

MeRCI based prototype helped them find the appropriate information quickly and easily due to its relevance, and its clear, easy and organized representation.

Because the studies involved a small number of scenarios in a single domain, more evaluation is needed to justify broader claims. Nevertheless, these initial results suggest that, by using domain and operational knowledge to provide context specific to the goal of the users, the SA levels can be substantially improved.

Contributions

The primary contribution of my work is to the interdisciplinary field of health informatics. My work expands on ideas from the contributing fields of human-centered information visualization, cognitive engineering and knowledge-based information systems to create a useful method that can be applied to the domain of public health. Specifically my contribution includes

- Specification of a method for building user-centered health information dashboards in complex real-time systems to promote Situational Awareness
- Document design principles for user-centered health information dashboards for improved situational awareness
- Introduce to public health a validated method of investigating the impact of HI dashboards through objective measures of awareness of system users

Though it is widely accepted that context is needed in improving the human ability to input and interact with computers in both traditional and dynamic settings, there is a vague understanding on how to apply context to systems. System designers today continue to face the expanding gap between the ability of technology to provide oceans of data and the human ability

to effectively process it. The infinite number of possibilities of presenting information has merely been recognized but is rarely used to its full potential in narrowing this gap.

This research work promulgates the principle of designing system appropriate to user goals and maximizing a person's ability to process information. Taking users' SA needs into account while building the system represents a paradigm shift from traditional UCD. In addition, the use of domain and operational knowledge to trigger the appropriate mental model aids in delivering better situational awareness. This fundamental change in the focus of system design has produced very high levels of confidence and timeliness in acquiring the awareness. The approach places importance on the context of the data as it applies to user goals instead of on presentation and visual elements. By focusing on the meaning and role of the data as it relates to user goals, the approach delivers a platform for dynamic information presentation that is matched to the users' goals, tasks and SA requirements. Consequently the user's brain processes the information that activates schemata leading the user to focus attention on critical environmental cues that are relevant, anticipate future states based on past models and classify situations at a rapid rate enabling faster decisions.

Although the study has been implemented in the domain of public health, the method can be easily applied to other domains. The adoption of accepted and validated CTA process and a validated evaluation methodology has made this approach applicable to domains of health informatics, finance, systems engineering and others. With its key focus on decision-making in dynamic systems, the design approach provides a way to overcome the barriers to technology-centered design and create user-centered systems. The design principles discussed in this study are capable of guiding the implementation of system interfaces that are applicable beyond public

health. Undoubtedly, the principles are not finalized and as more research is conducted on SA in a multitude of domains, the factors affecting SA will serve to augment these design principles.

Limitations and Future Work

In this section, I present the limitations of my research in its scope and some of the potential areas where further avenues for research exist. This research was conducted in a controlled environment with a small number of scenarios covering a single domain. The study goals focused on evaluating the impact of system design on users' SA levels, response time and level of confidence. The outcomes of the study, although very revealing, open up avenues for further research that need to be conducted to substantiate the general claim of the role of MeRCI design. Here I have discussed some of the key areas where further research is proposed.

An aspect of this research is the support provided for users to rapidly classify and adapt to information perceived in a previous situation. As psychological studies in other domains suggest, perfect schema matching is not required to support the interpretation of a current situation. Experts and users with longer experience can utilize the knowledge and the operation triggers to match the schemata at a faster pace, even when the situation is not very similar. Studies also show that users who are novices or less experienced require far more information to build and match a schema. In my research, the outcome did not differentiate this role of experience and background knowledge, potential reasons for this including the limited set of study samples which narrow the study groups to two based on experience and three based on age. Further study is recommended to better understand changes in SA needs as experience, age and background vary.

Uncertainty is a common feeling among users operating in a complex environment. The interaction between uncertainty and users' SA has been one of the areas studied but lacks full

guidance for system designers in building systems that can instill a higher level of confidence among users when facing uncertainty. Endsley and Jones have pointed out that the accuracy of one's decision is based on the actual accuracy of the individual SA. Confidence in the SA also play a critical role; users with a higher level of SA and higher confidence are expected to formulate and execute actions that will result in good outcomes, whereas users who have high SA but low confidence will delay action and try to gather more information to achieve better SA. In contrast, a user who does not have sufficient SA but high confidence could end up with a bad outcome. Thus it is important to understand that there is not only a need to ensure that users have as good a picture of the situation as possible, but that they are also able to reliably apply the correct amount of confidence to the situation. In this study, due to its limited number of samples, the concept of confidence and its linkages to SA were not studied in detail. Further work will help expand the methods to other domains.

Shared situational awareness is a common attribute required in a team environment. Today, organizations are highly matrix-driven and individuals' roles are focused on specific functions in the system. The concept of "team SA" defined by Endsley focuses on "the degree to which every team member possesses the SA required for his or her responsibilities." The success and failure of the team depends on the success and failure of each of its members. Although the studied domain, PH surveillance, is managed by multiple resources, the overall goal of surveillance practice and monitoring requires team SA. The study, however, did not consider the team SA requirements delivered by this information system. This is a potential area for investigating and expanding the guidelines for MeRCI.

Concluding Remarks

Today, the scope of surveillance has increased from that involved in keeping a vigilant watch on a single borough in London to whole national and global communities. We have transformed the tools of data collection from shoe leather and a dog-eared notebook to electronic messaging that travels at the speed of telecommunications signals. We have access to electronic data from sensors in all walks of life. The growth of data is exponential and we have analytic engines with tools that can perform more calculations in a day than people used to perform in their lifetime. But we have a number of problems to solve around how to effectively access this data. This research has worked on addressing a fundamental problem of how to present information to users in a way that will improve their awareness of the situation.

Like any complex domain, public health user needs for routine or PH event surveillance are beyond the scope of any one single system. Dashboards are critical system components that bring together information from multiple systems so as to deliver awareness to the user about a particular situation. Current dashboards leave the users to adapt to environmental situations in utilizing the data rather than providing context for them. This also leads to failure in providing situational awareness among the users of the system. Public health practitioners need more contexts to interpret the data presented to them. Although the literature stresses the importance of using context, to date there is very little work advocating how to use and represent context to improve awareness.

This research introduced a method of implementing contextualized information representation called MeRCI, in which the user goals, domain knowledge and operational knowledge are used to present information for processing in a clear and easy way. A follow-up study to measure awareness showed increased levels of awareness in all three SA levels among

users of the interface developed using the MeRCI method. Users also exhibited higher levels of confidence when using the system as compared to traditionally built systems. This research has provided a better understanding of the contextualized information representation in public health systems and will provide a model for design and evaluation of information representation in health information systems.

I would like to finish by presenting four reasons why I believe that understanding and applying the SA construct is important in the design of strategically important information systems and why we should continue research into identifying methods to achieve and continue to maintain SA. It is important to mention:

- a) *SA can be linked to performance*: The higher the SA, the better the performance and outcomes.
- b) *Inadequate SA may be associated with errors*: Improve SA to reduce human errors.
- c) *SA may be related to expertise*: Deliver SA to bridge the gap in expertise.
- d) *SA is the basis for decision making in most cases*: Improve SA for aiding decision-making.

APPENDIX A – GOAL DIRECTED TASK ANALYSIS

Goal Directed Task Analysis

In this section, all the GDTA hierarchy tables are presented

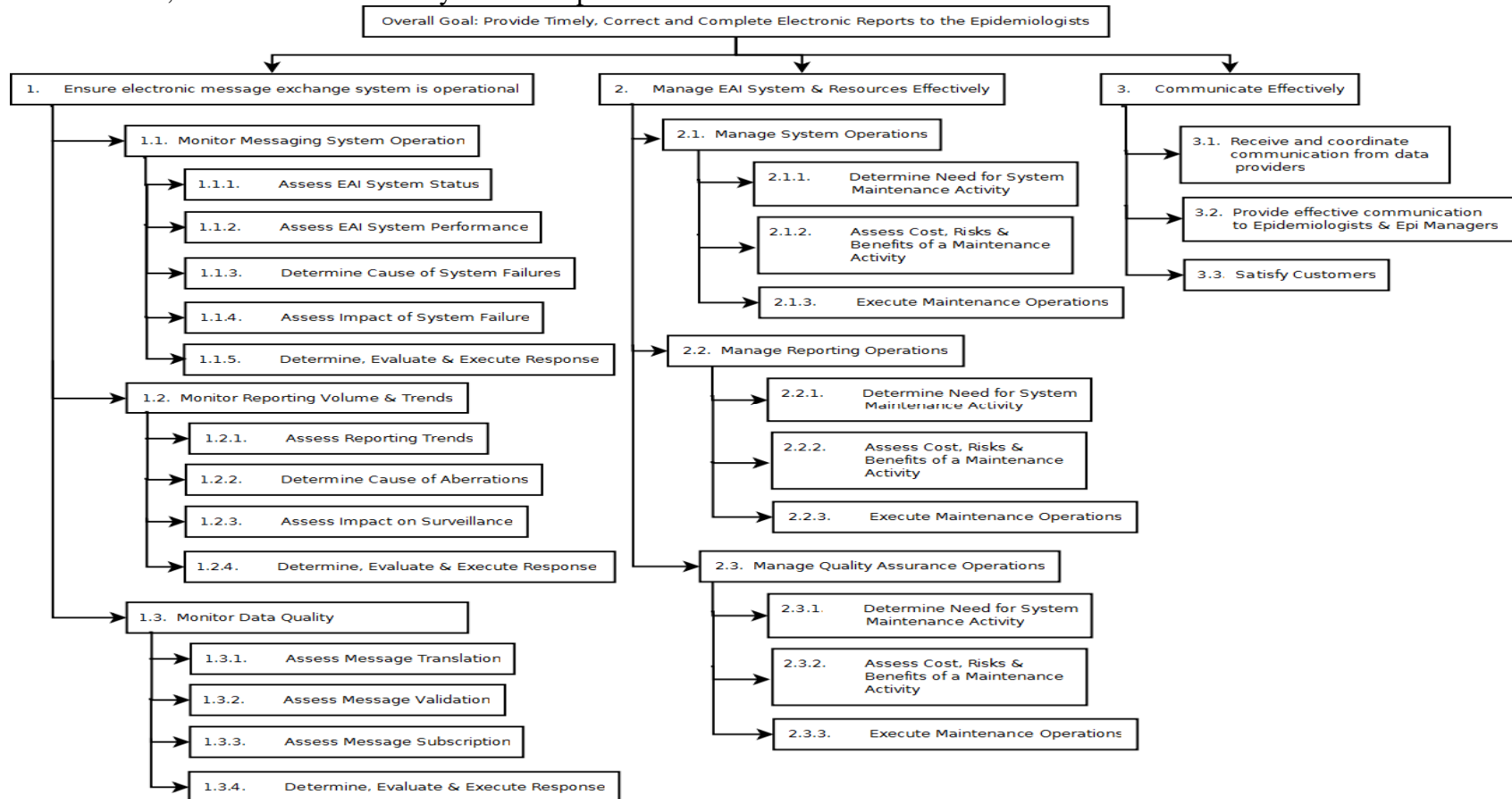


Figure 38: GDTA - Overall Goal

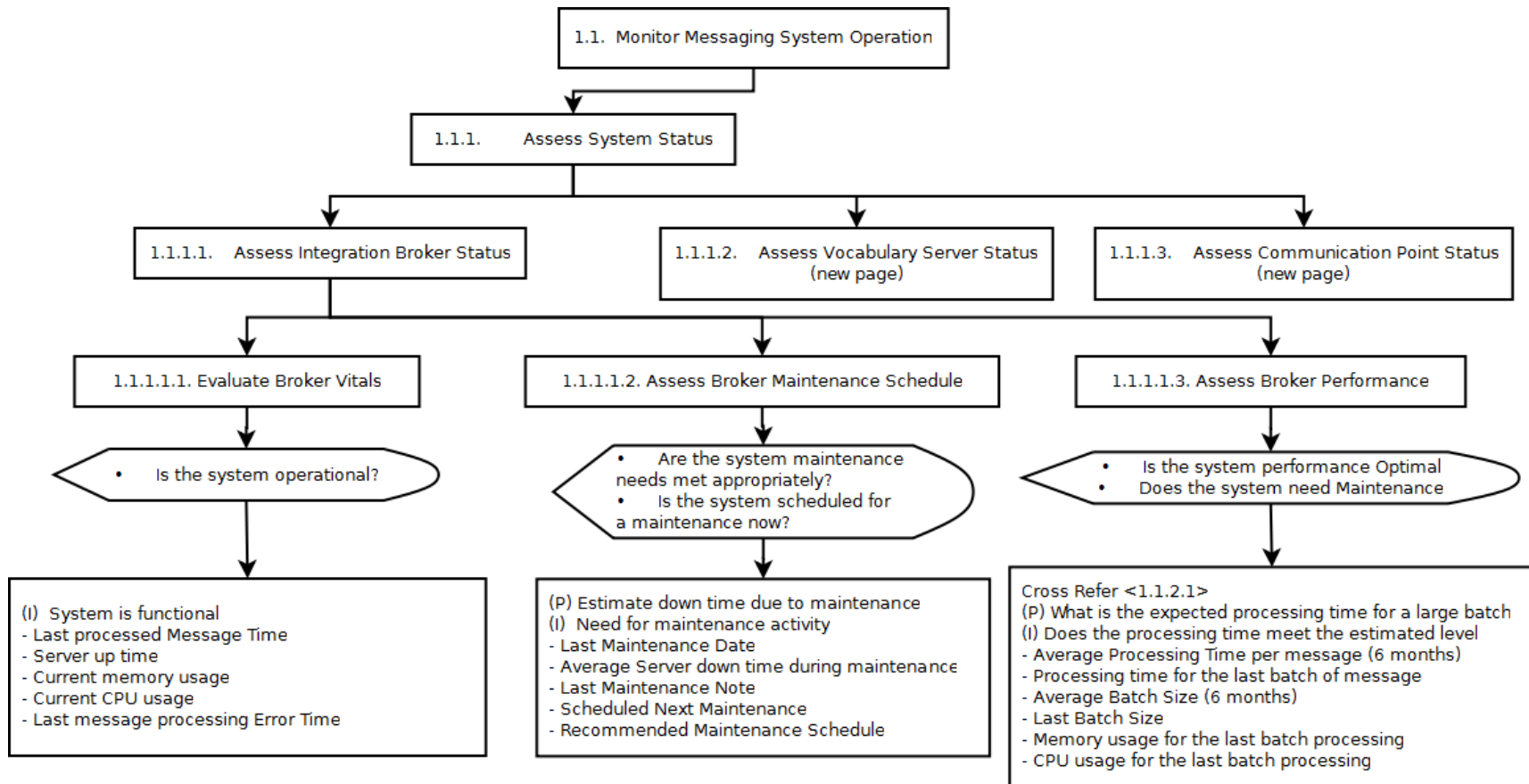


Figure 39: Monitoring System Operation

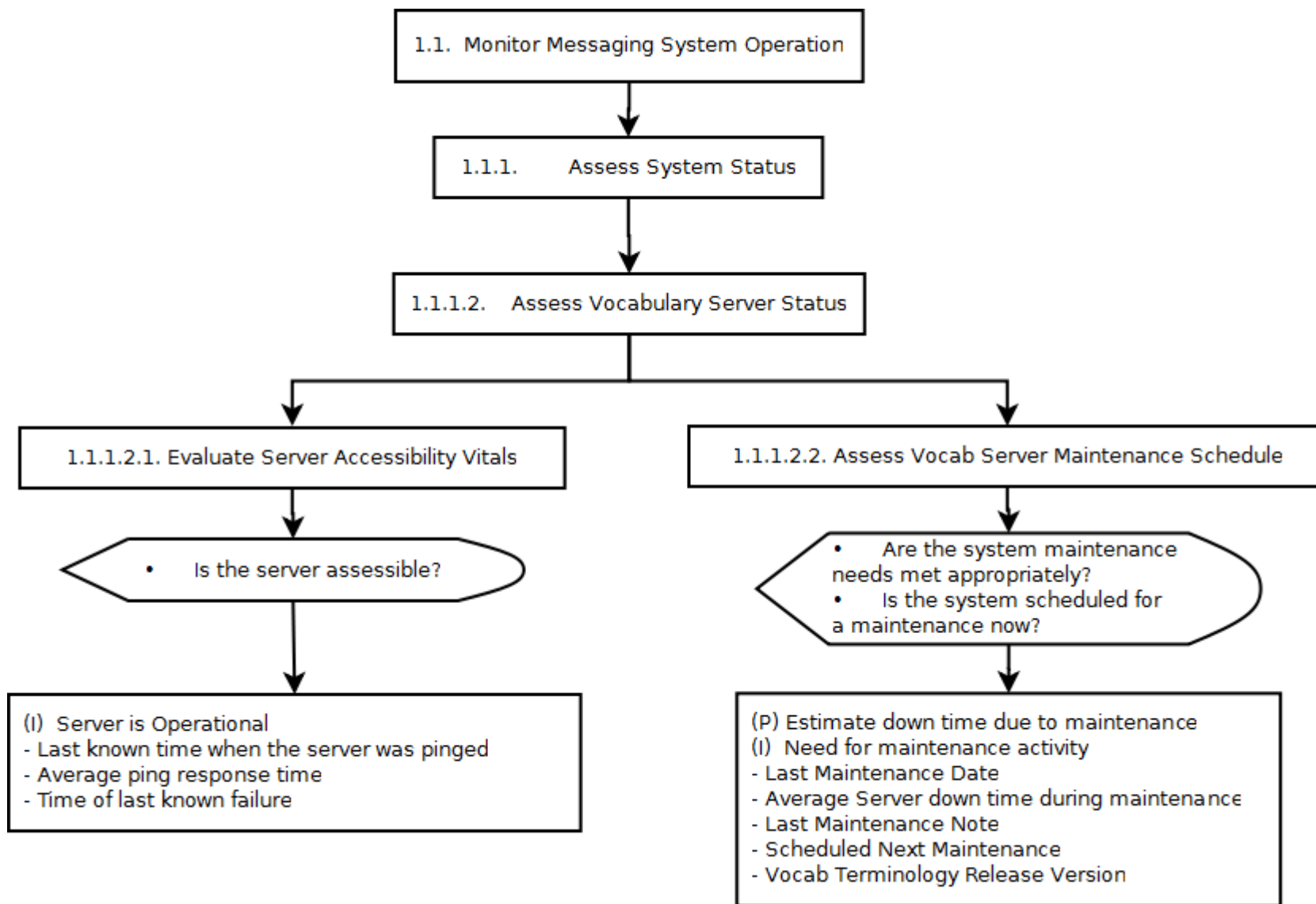


Figure 40: Assess System Status

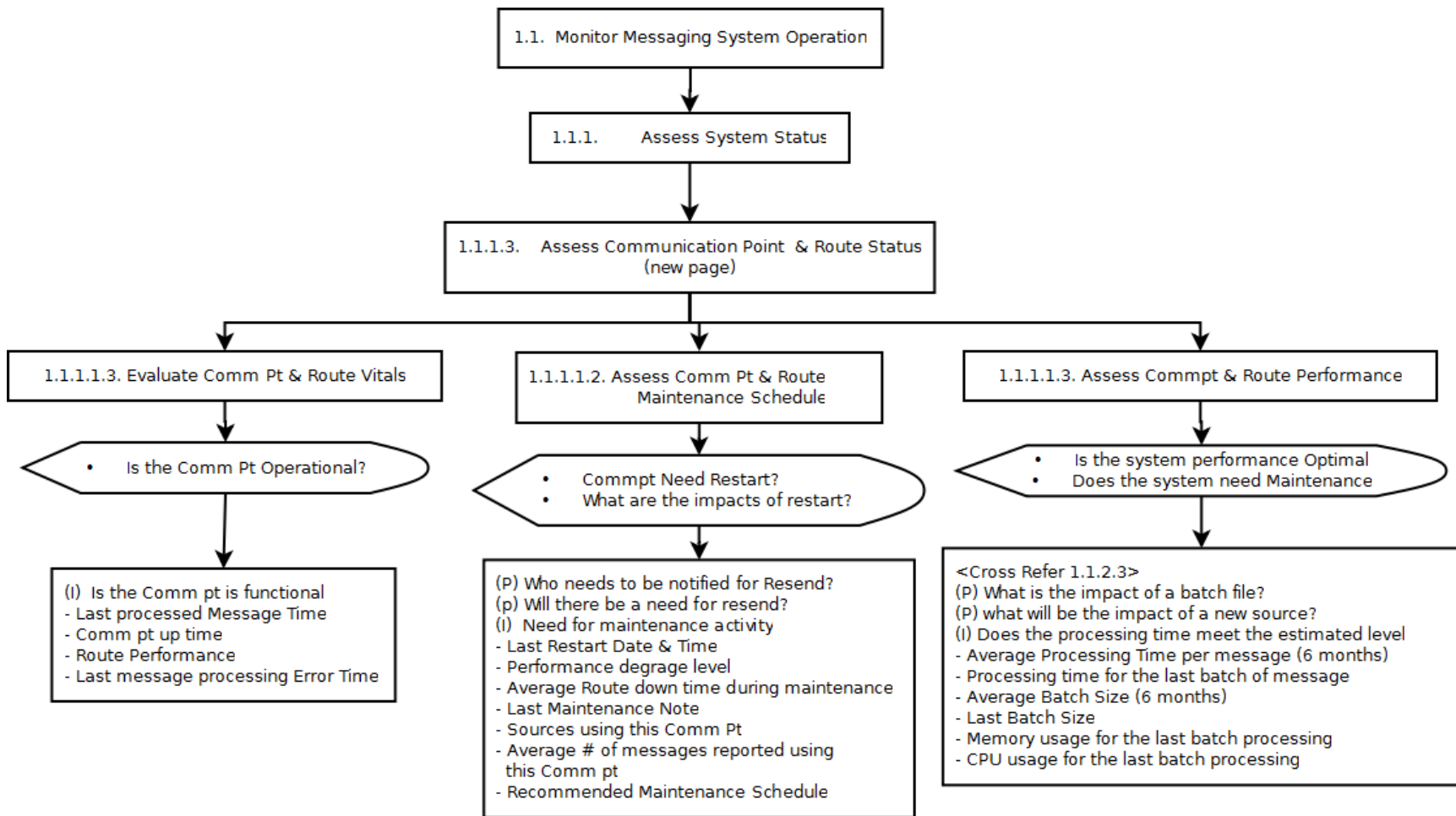


Figure 41: Assess Comm pt & Route Status

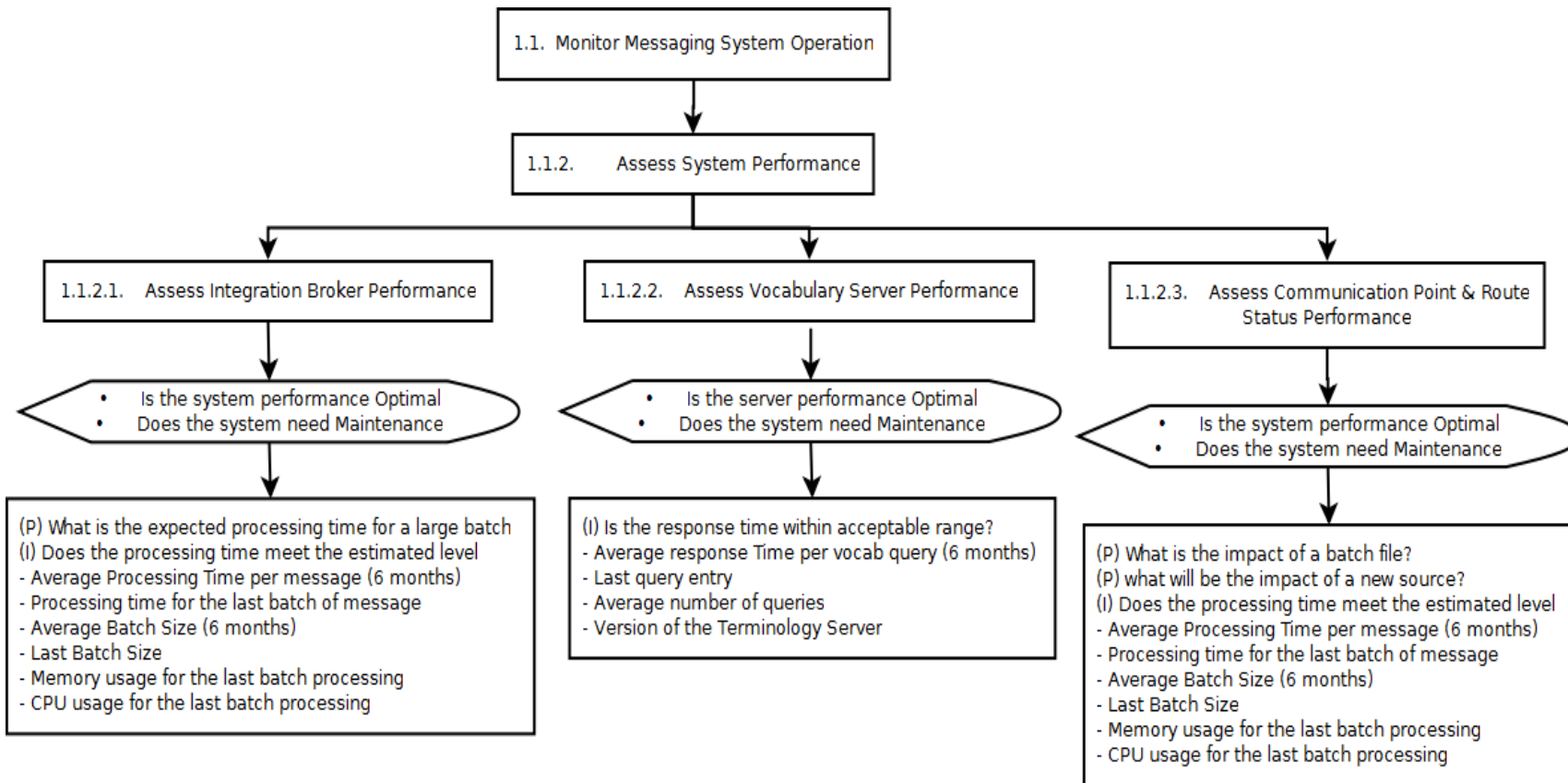


Figure 42: Assess System Performance

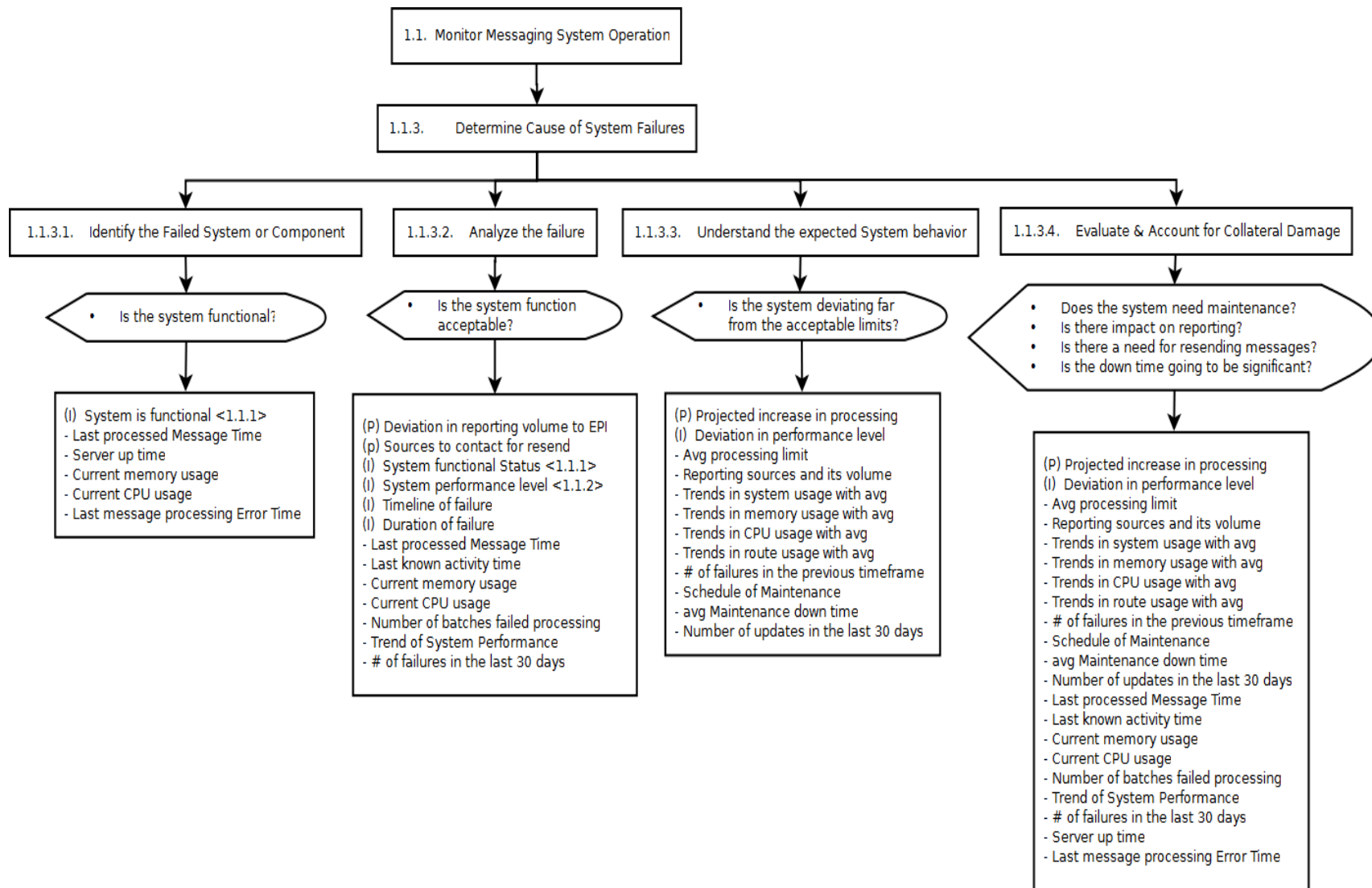


Figure 43: Determine Cause of System Failures

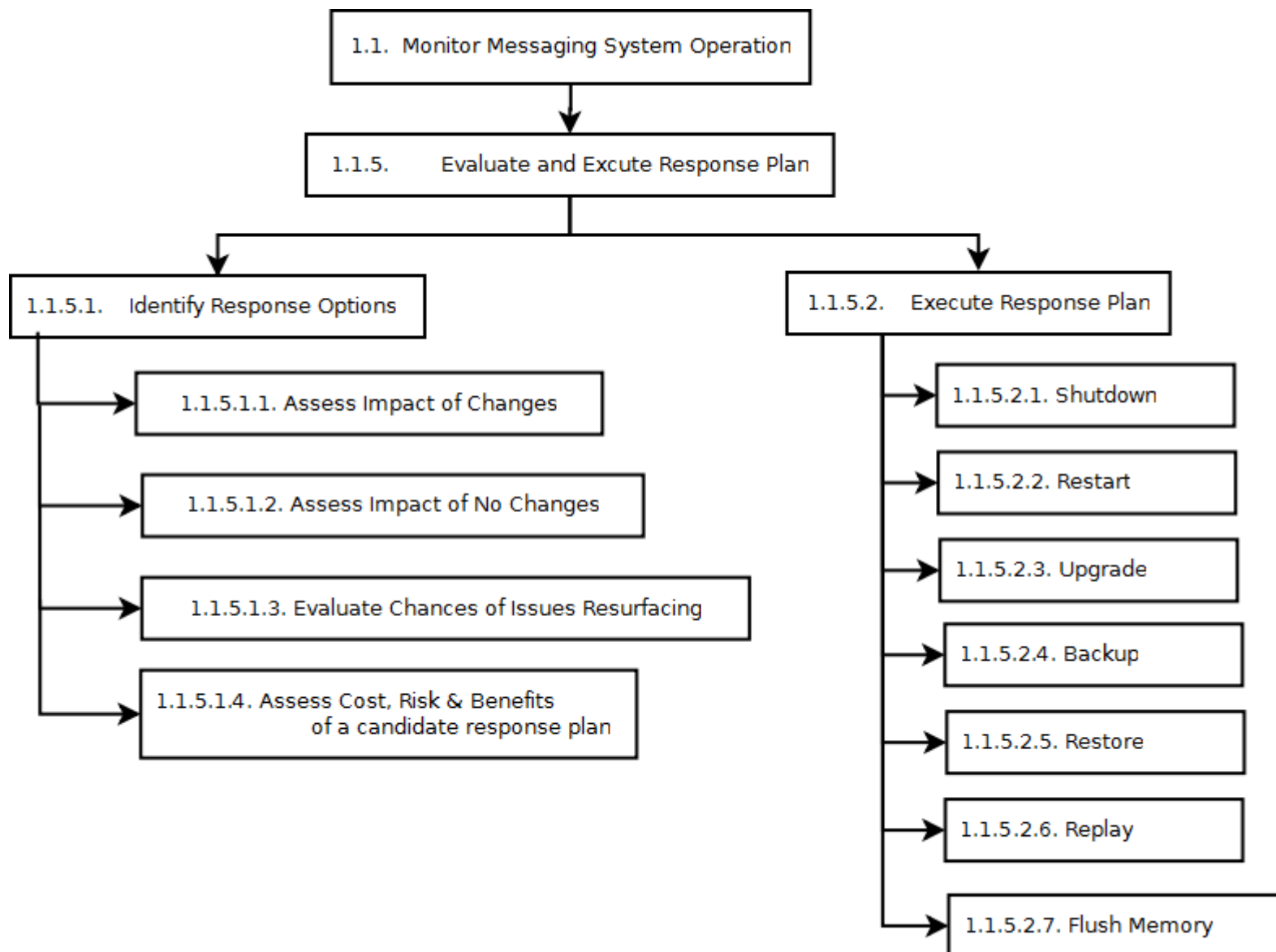


Figure 44: Evaluate and Execute Response Plan

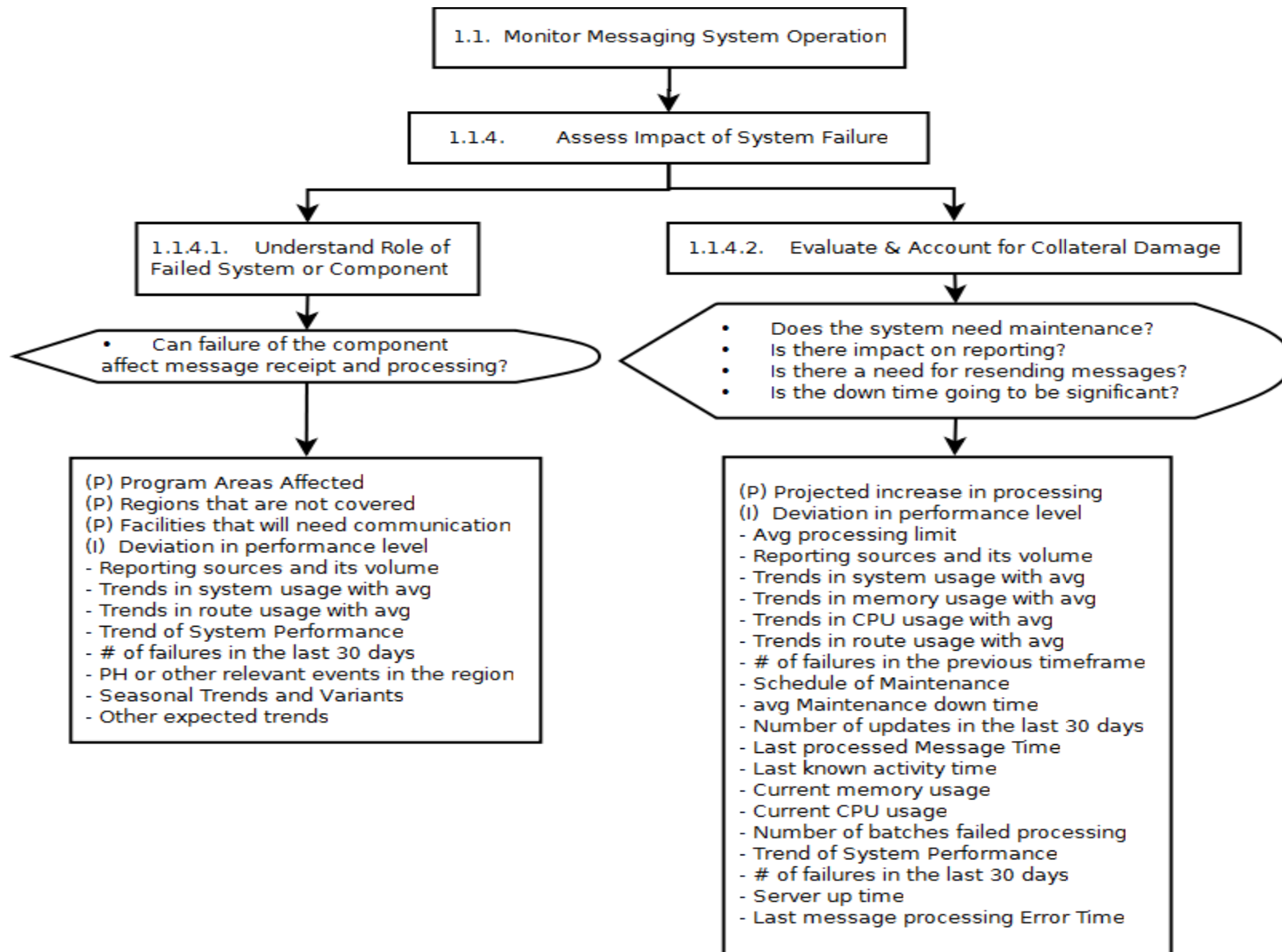


Figure 45: Assess Impact of Failures

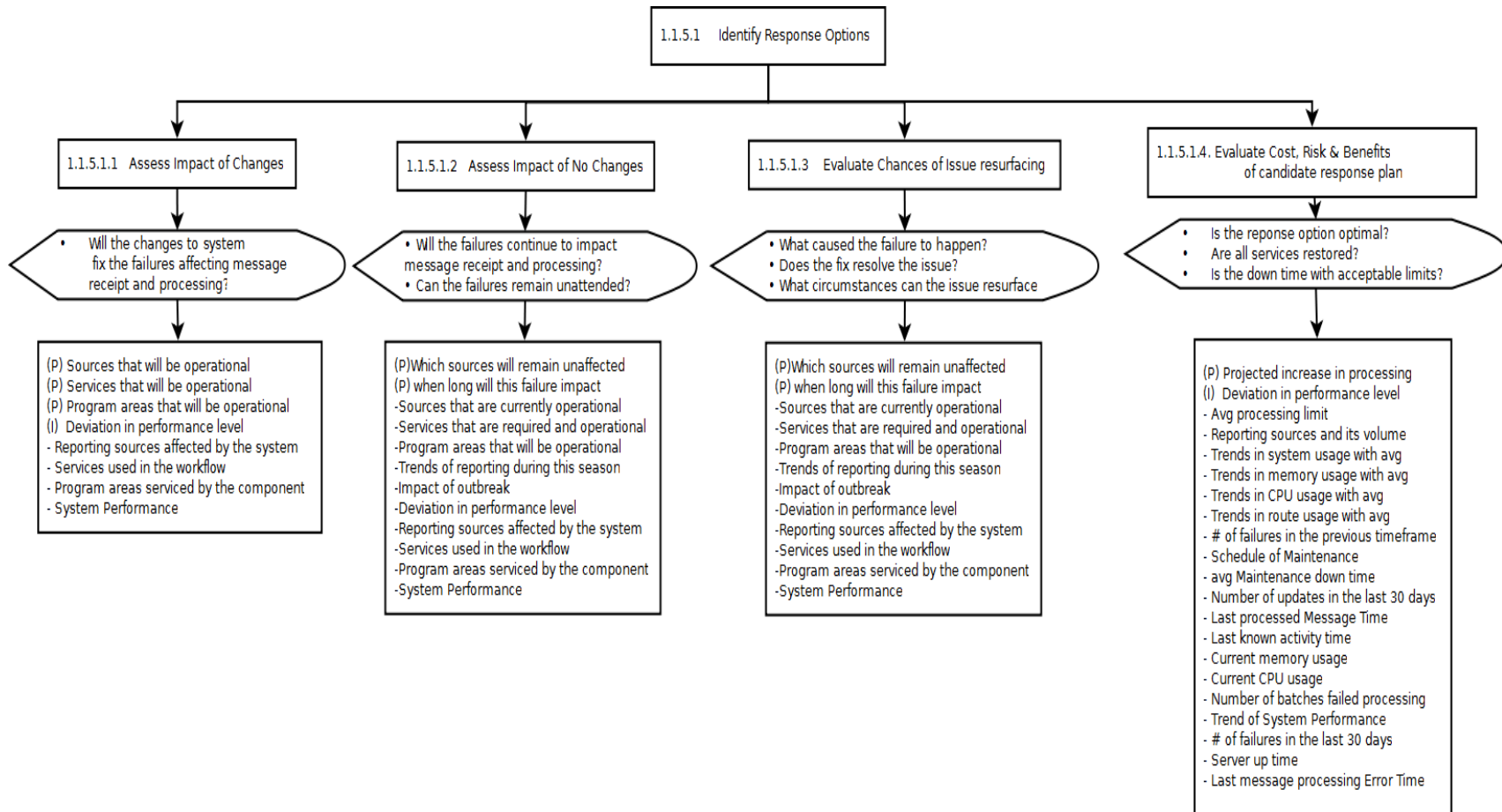


Figure 46: Identify Response Options

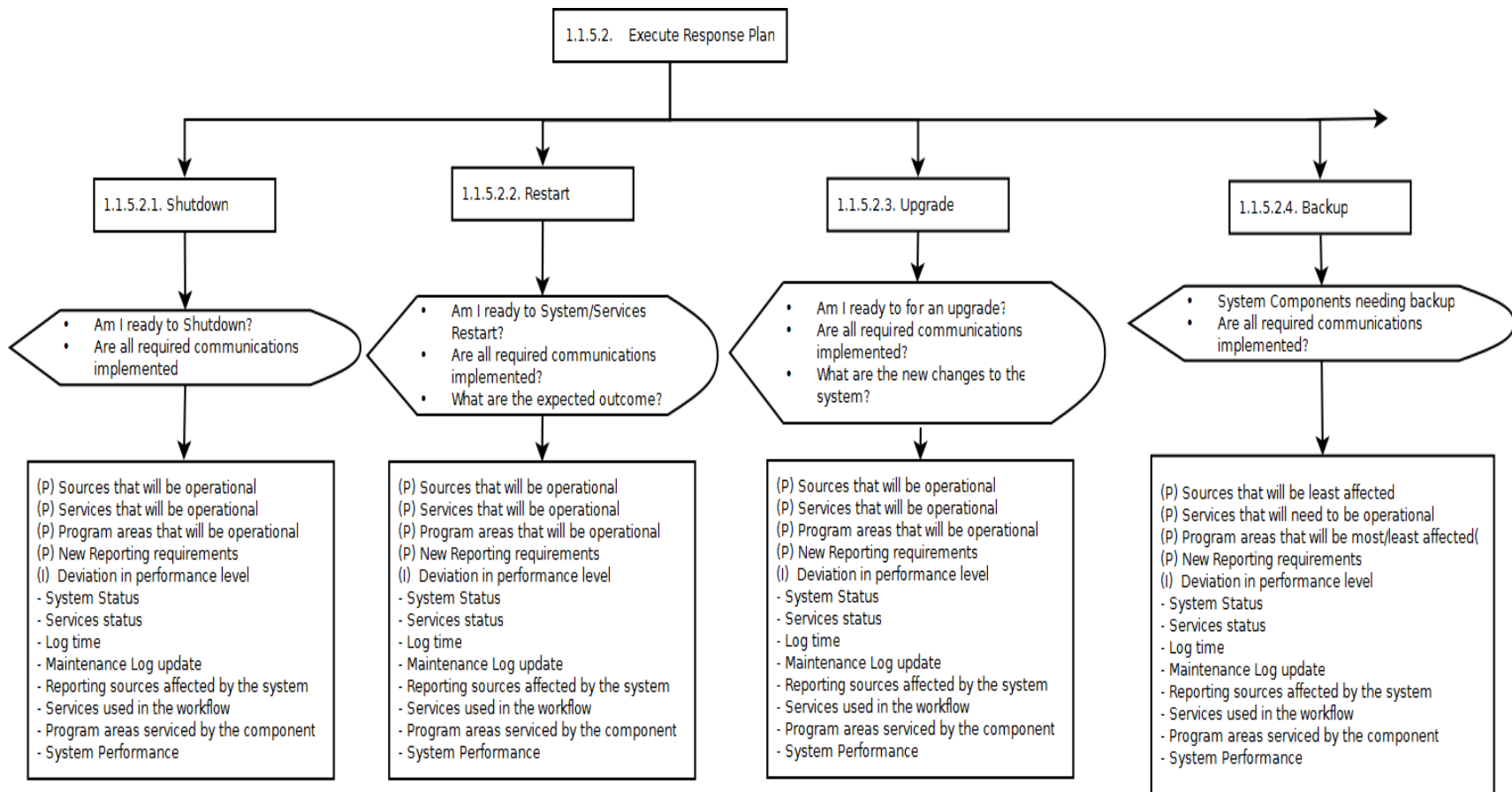


Figure 47: Execute Response Options

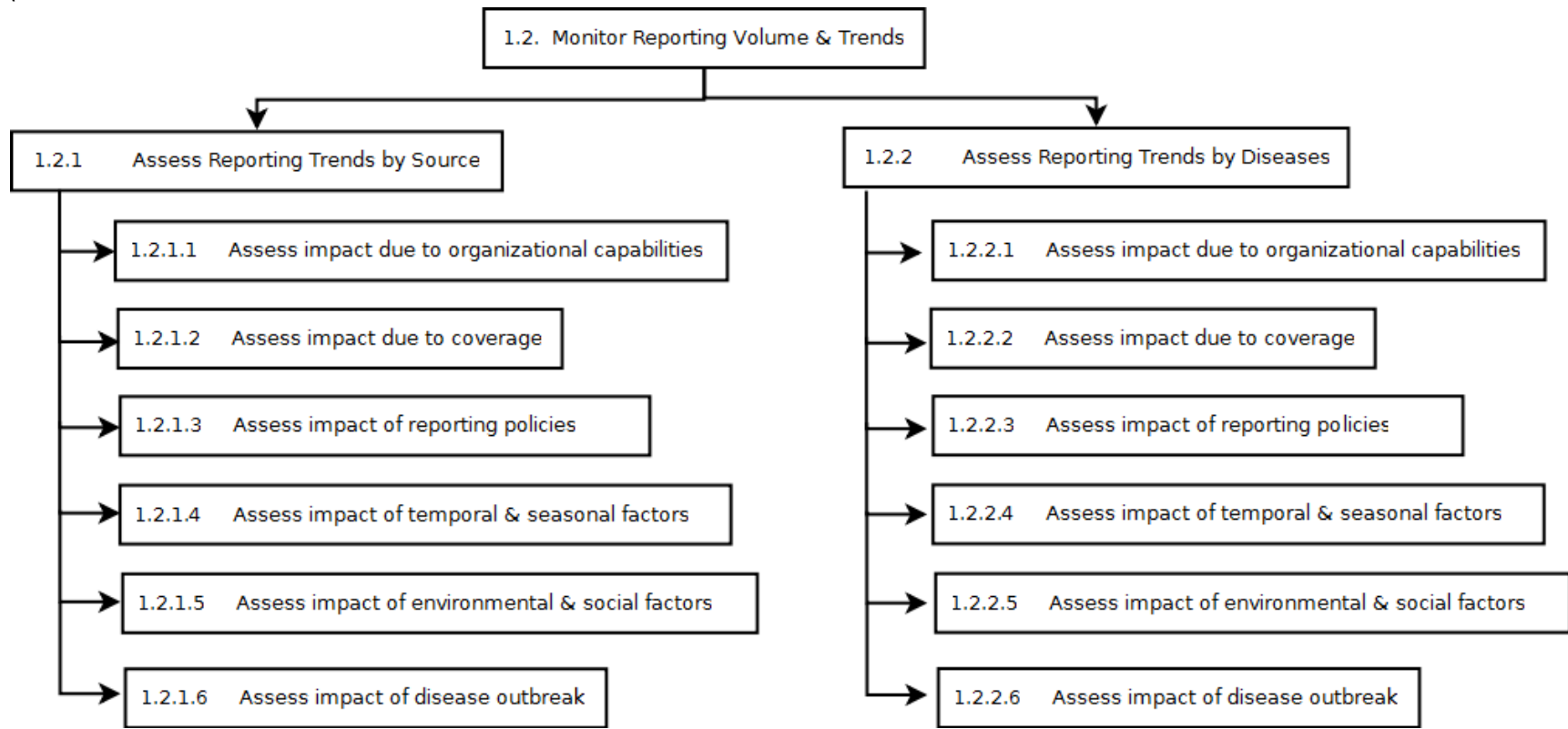


Figure 48: Monitor Reporting Volume and Trends

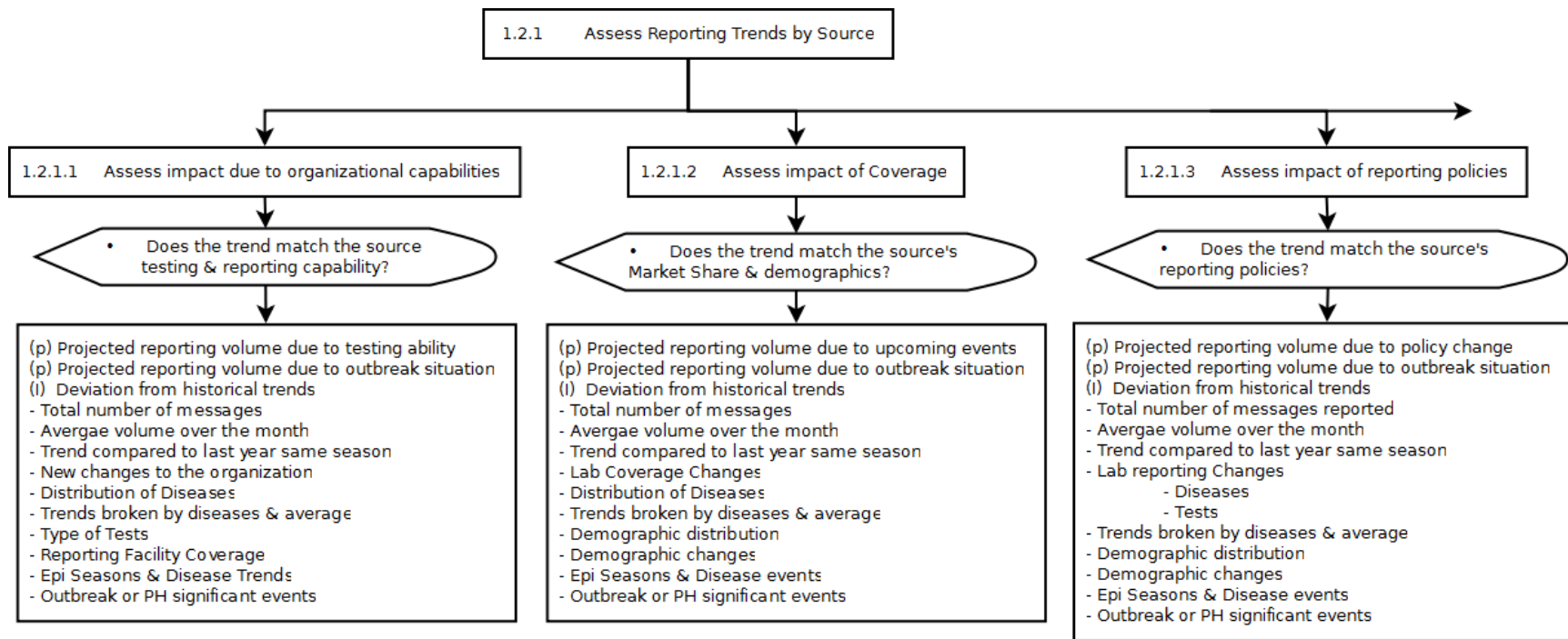


Figure 49: Assess Reporting Trends by Source

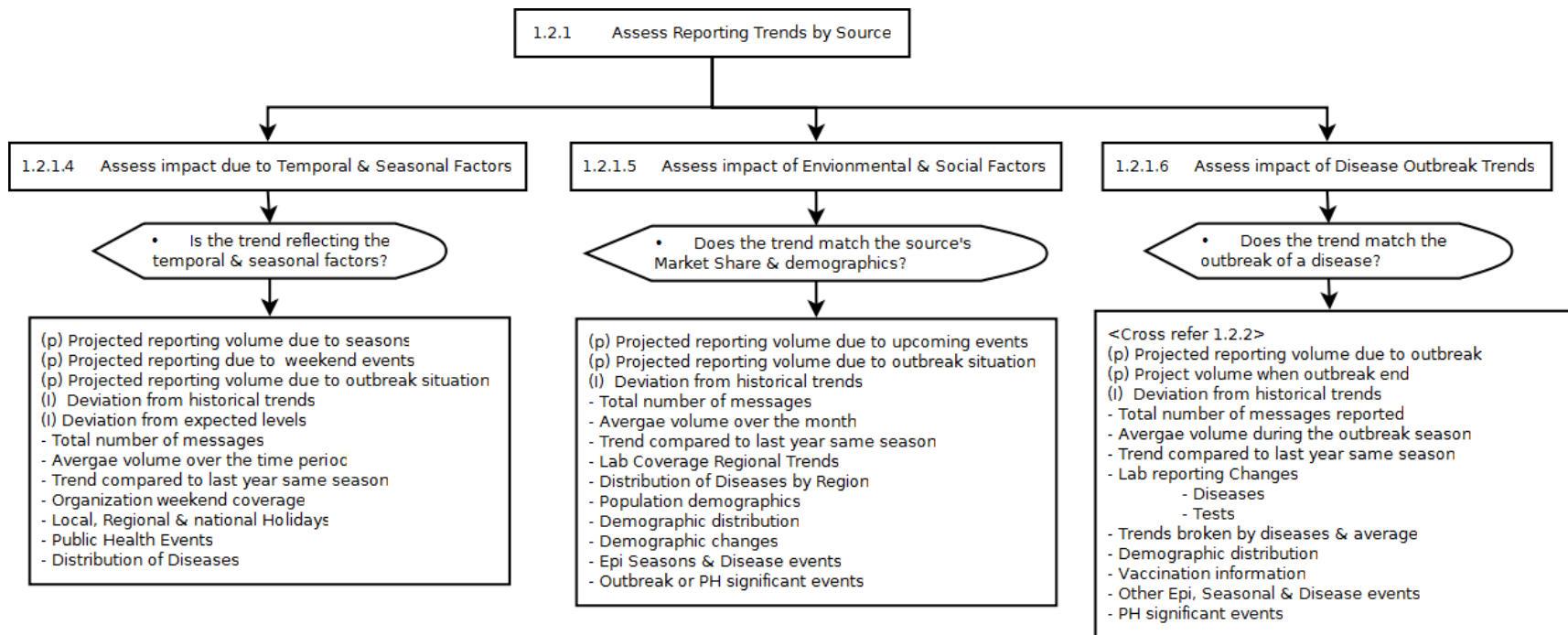


Figure 50: Assess Reporting Trends by Source (continued)

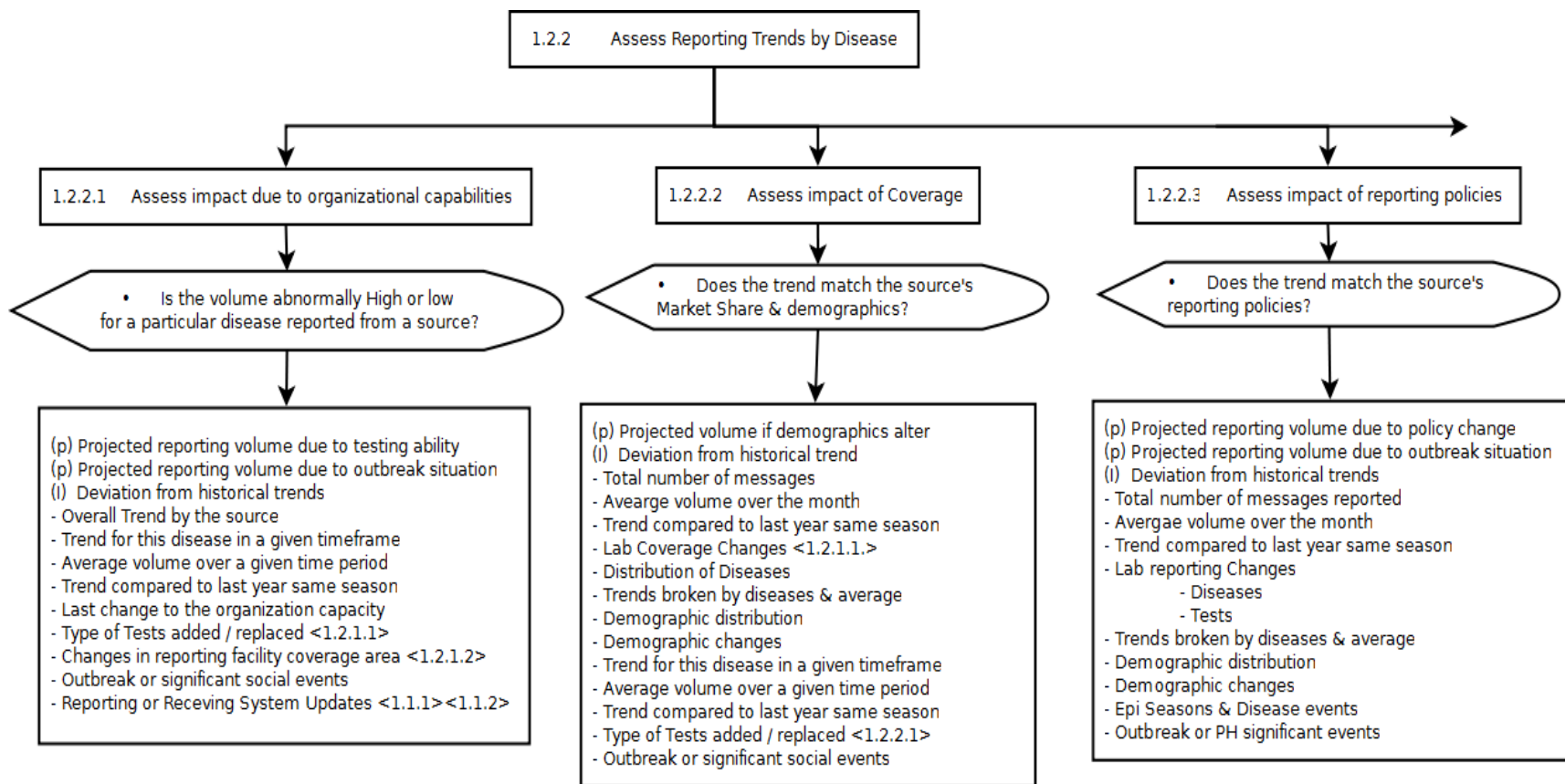


Figure 51: Assess Reporting Trends by Diseases

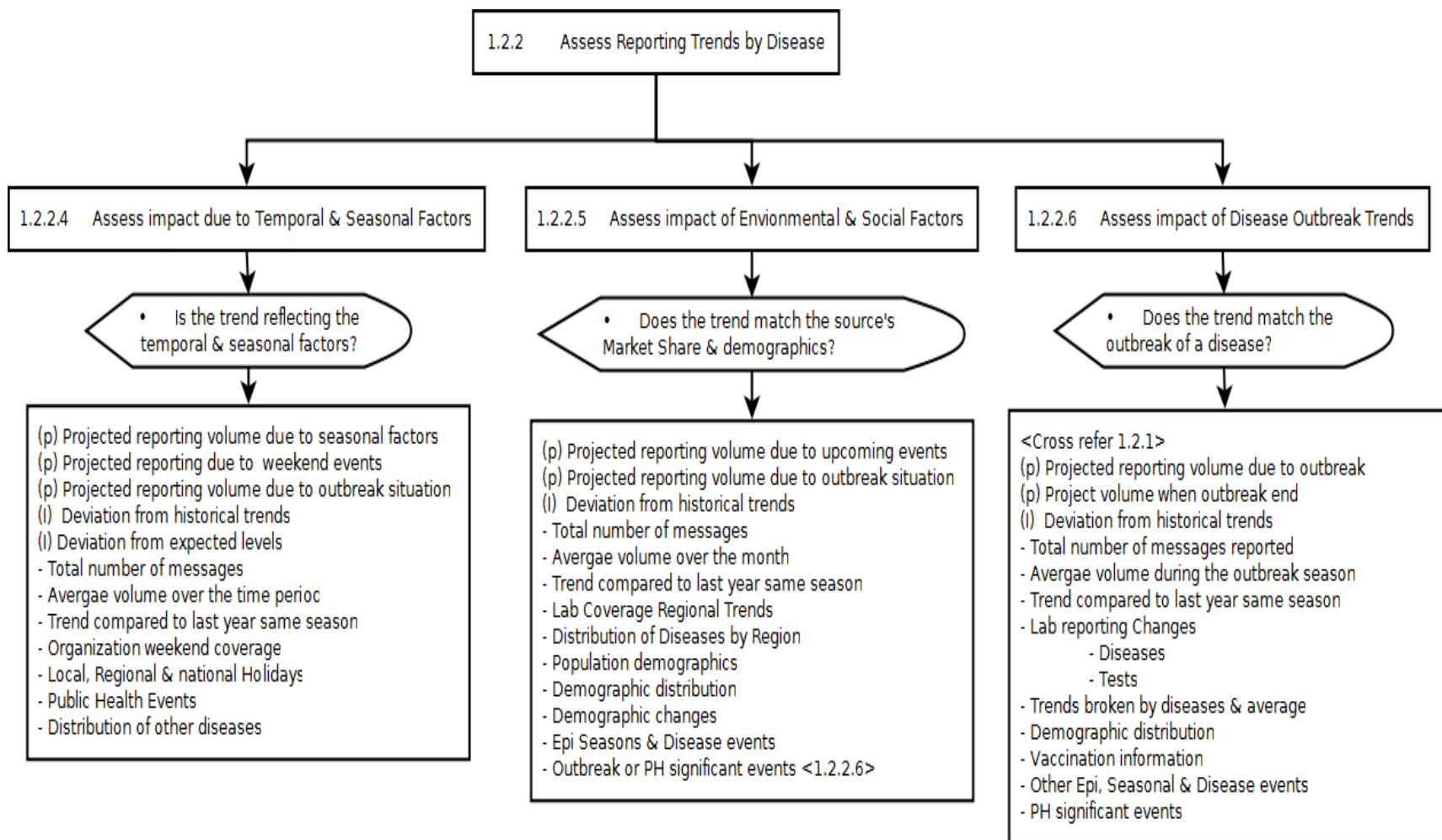


Figure 52: Assess Reporting Trends by Diseases (continued)

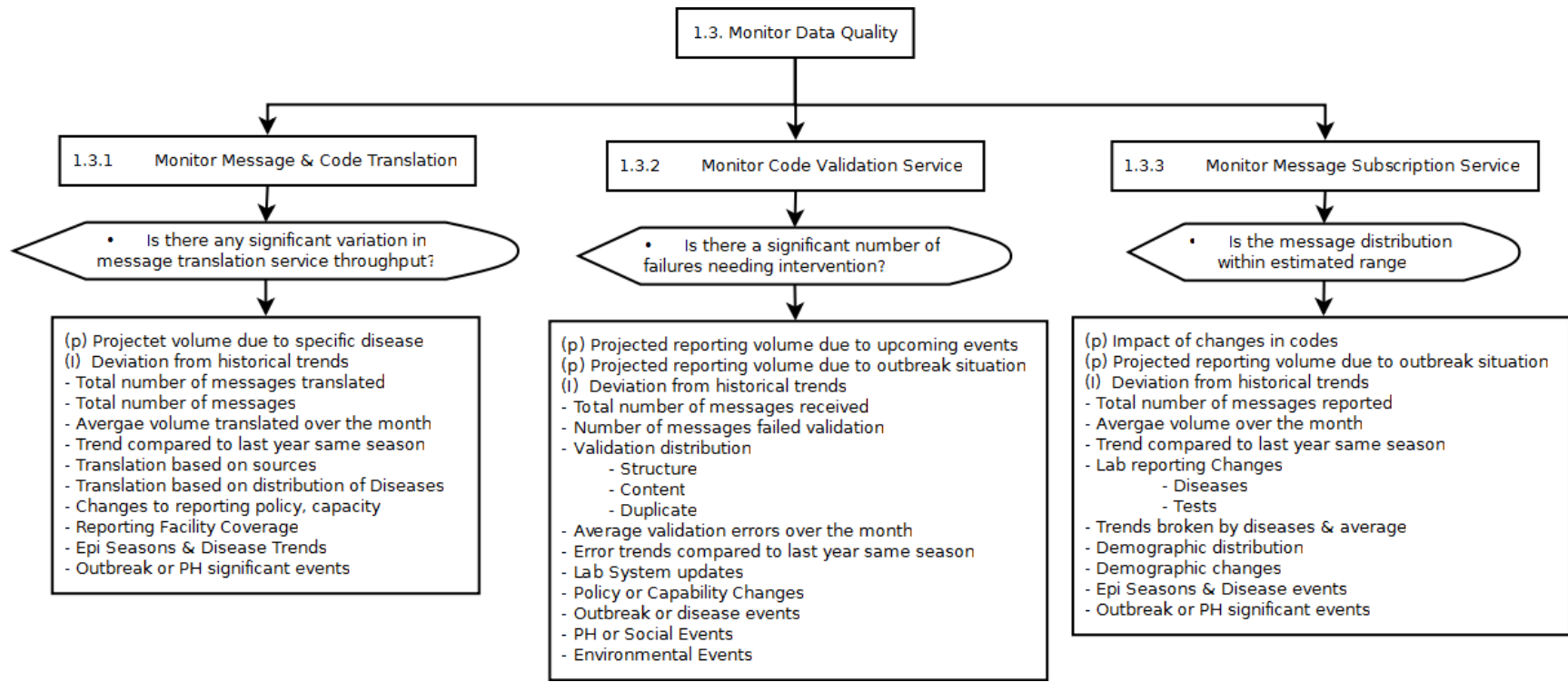


Figure 53: Monitor Data Quality

APPENDIX B – CONCEPT MAP FORMAL REPRESENTATION

In this section, I have included the drool rule for 3 scenarios described in chapter 3 and the java code that will invoke the rules.

File: rule.drl

```
package com.msg

import com.msg.model.LabMessage;
import com.msg.utils.DroolsMessageTrendHelper;
import com.msg.utils.MessageTrendConstants;
import com.msg.utils.MessageTrendUtils;
import java.util.Calendar

//-----//
rule "Rule_HolidayEffect"
dialect "mvel"
    when
        m: LabMessage ( res : result, vol : volume, date : currDate, comm :
labCommSystem, season : season)
        eval ( vol > MessageTrendConstants.WEEKLY_AVERAGE_2011)
        eval ((DroolsMessageTrendHelper.getPrevDay(date) == Calendar.SUNDAY) || (
DroolsMessageTrendHelper.isPrevDayFedHoliday (date)))
    then
        m.setResult(MessageTrendConstants.FALSE);
        System.out.println("Holiday Effect");
    end

//-----//

rule "Rule_SystemFailure"
dialect "mvel"
    when
        m: LabMessage ( res : result, vol : volume, date : currDate, comm :
labCommSystem, season : season)
        eval ( comm ==
MessageTrendConstants.LAB_COMMUNICATION_SYSTEM_FTP &&
!DroolsMessageTrendHelper.checkFtpComm())
        eval ( vol ==0 || vol <
MessageTrendConstants.CURRENT_SEASON_WEEKLY_AVERAGE)
    then
        m.setResult(MessageTrendConstants.FALSE);
        System.out.println("Comm pt failure Effect");
```

end

//-----//

```
rule "Rule_HighVolumeWednesday"
dialect "mvel"
    when
        m: LabMessage ( res : result, vol : volume, date : currDate, comm :
labCommSystem, season : season)
        eval ( vol > MessageTrendConstants.WEEKLY_AVERAGE_2011)
        eval ( DroolsMessageTrendHelper.getCurrDay(date) ==
Calendar.WEDNESDAY)
        eval ( season == MessageTrendConstants.SEASON_SUMMER)
    then
        m.setResult(MessageTrendConstants.FALSE);
        System.out.println("Lab PCR test day");
end
```

//-----//

File : LabMessage.java

```
package com.msg.model;

public class LabMessage {
    private String msgId;
    private String result;
    private Integer volume;
    private String currDate;
    private String labCommSystem;
    private String season;

    public String getMsgId() {
        return msgId;
    }
    public void setMsgId(String msgId) {
        this.msgId = msgId;
    }
    public String getResult() {
        return result;
    }
    public void setResult(String result) {
        this.result = result;
    }
    public Integer getVolume() {
```

```

        return volume;
    }
    public void setVolume(Integer volume) {
        this.volume = volume;
    }
    public String getCurrDate() {
        return currDate;
    }
    public void setCurrDate(String currDate) {
        this.currDate = currDate;
    }
    public String getLabCommSystem() {
        return labCommSystem;
    }
    public void setLabCommSystem(String labCommSystem) {
        this.labCommSystem = labCommSystem;
    }
    public String getSeason() {
        return season;
    }
    public void setSeason(String season) {
        this.season = season;
    }
}

```

File : MessageTrendMain.java

```

package com.msg.service;

import java.io.BufferedReader;
import java.io.FileNotFoundException;
import java.io.FileReader;
import java.io.IOException;

public class MessageTrendMain {
    public static final void main(String[] args) {
        MessageTrendServiceImpl messageTrendServiceImpl = new
MessageTrendServiceImpl();
        String s = readFileAsString("LabMessageDetails.xml");
        messageTrendServiceImpl.detectMessageTrend(s);
    }

    private static String readFileAsString(String filePath){
        StringBuffer fileData = new StringBuffer(1000);
        BufferedReader reader;
        try {

```

```

        reader = new BufferedReader(new FileReader(filePath));
        char[] buf = new char[1024];
        int numRead=0;
        while((numRead=reader.read(buf)) != -1){
            String readData = String.valueOf(buf, 0, numRead);
            fileData.append(readData);
        }
        buf = new char[1024];
        reader.close();
    } catch (FileNotFoundException e) {
        e.printStackTrace();
    } catch (IOException e) {
        e.printStackTrace();
    }
    return fileData.toString();
}
}

```

File: MessageTrendServiceImp.java

```

package com.msg.service;

import java.io.IOException;
import java.io.StringReader;

import org.drools.KnowledgeBase;
import org.drools.KnowledgeBaseFactory;
import org.drools.builder.KnowledgeBuilder;
import org.drools.builder.KnowledgeBuilderError;
import org.drools.builder.KnowledgeBuilderErrors;
import org.drools.builder.KnowledgeBuilderFactory;
import org.drools.builder.ResourceType;
import org.drools.io.ResourceFactory;
import org.drools.logger.KnowledgeRuntimeLogger;
import org.drools.logger.KnowledgeRuntimeLoggerFactory;
import org.drools.runtime.StatefulKnowledgeSession;

import com.msg.model.LabMessage;
import com.msg.utils.MessageTrendConstants;

import javax.xml.parsers.DocumentBuilder;
import javax.xml.parsers.DocumentBuilderFactory;
import javax.xml.parsers.ParserConfigurationException;

import org.w3c.dom.*;
import org.xml.sax.InputSource;

```

```

import org.xml.sax.SAXException;

public class MessageTrendServiceImpl implements MessageTrendService {

    @Override
    public String detectMessageTrend(String xmlFile) {
        String xmlString = "";

        try {

            // load up the knowledge base
            KnowledgeBase kbase = readKnowledgeBase();
            StatefulKnowledgeSession ksession = kbase
                .newStatefulKnowledgeSession();
            KnowledgeRuntimeLogger kLogger = KnowledgeRuntimeLoggerFactory
                .newFileLogger(ksession, "test");

            LabMessage message = parseXML(xmlFile);
            message.setResult(MessageTrendConstants.FALSE);
            ksession.insert(message);
            ksession.fireAllRules();

            kLogger.close();

            ksession.dispose();

        } catch (Throwable t) {
            t.printStackTrace();
        }
        return xmlString;
    }

    private LabMessage parseXML(String XML) {
        LabMessage msg = new LabMessage();

        try {
            DocumentBuilderFactory docBuilderFactory = DocumentBuilderFactory
                .newInstance();
            DocumentBuilder docBuilder =
docBuilderFactory.newDocumentBuilder();
            Document doc = docBuilder.parse(new InputSource(new StringReader(
                XML)));
            doc.getDocumentElement().normalize();

            String value = null;
            value = getElement(doc, "MSG", "msgId");

```

```

        System.out.println("MSGID : " + value);
        msg.setMsgId(value);

        value = getElement(doc, "MSG", "volume");
        Integer val = Integer.parseInt( value);
        System.out.println("Volume:" + value);
        msg.setVolume(val);

        value = getElement(doc, "MSG", "date");
        System.out.println("Date:" + value);
        msg.setCurrDate(value);

        value = getElement(doc, "MSG", "commSystem");
        System.out.println("Comm System:" + value);
        msg.setLabCommSystem(value);

        value = getElement(doc, "MSG", "season");
        System.out.println("Season:" + value + "\n");

        msg.setSeason(value);

    } catch (ParserConfigurationException pce) {
        pce.printStackTrace();
    } catch (SAXException se) {
        se.printStackTrace();
    } catch (IOException ioe) {
        ioe.printStackTrace();
    }
}

return msg;
}

private static String getElement(Document doc, String segment1,
    String segment2) {
    String value = "";
    NodeList nlist = doc.getElementsByTagName(segment1);
    for (int s = 0; s < nlist.getLength(); s++) {
        Node node = nlist.item(s);
        if (node.getNodeType() == Node.ELEMENT_NODE) {
            Element mshElement = (Element) node;
            value = getTextValue(mshElement, segment2);
            return value;
        }
    }
    return value;
}
}

```

```

private static String getTextValue(Element ele, String tagName) {
    String textVal = null;
    NodeList nl = ele.getElementsByTagName(tagName);
    if (nl != null && nl.getLength() > 0) {
        Element el = (Element) nl.item(0);
        textVal = el.getFirstChild().getNodeValue();
    }
    return textVal;
}

private static KnowledgeBase readKnowledgeBase() throws Exception {
    KnowledgeBuilder kbuilder = KnowledgeBuilderFactory
        .newKnowledgeBuilder();
    kbuilder.add(ResourceFactory.newClassPathResource("Alert.drl"),
        ResourceType.DRL);
    KnowledgeBuilderErrors errors = kbuilder.getErrors();
    if (errors.size() > 0) {
        for (KnowledgeBuilderError error : errors) {
            System.err.println(error);
        }
        throw new IllegalArgumentException("Could not parse knowledge.");
    }
    KnowledgeBase kbase = KnowledgeBuilderFactory.newKnowledgeBase();
    kbase.addKnowledgePackages(kbuilder.getKnowledgePackages());
    return kbase;
}
}

```

File: MessageTrendService.java

```

package com.msg.service;

public interface MessageTrendService {

    public abstract String detectMessageTrend(String xmlFile);
}

```

File: DroolsMessageHelper.java

```

package com.msg.utils;

import java.util.Calendar;

public class DroolsMessageTrendHelper {

    /* Calculates and returns the current day of the week from the current date*/
}

```



```

public static Integer getCurrDay(String currDate) {
    Integer currDay;
    Calendar dt = MessageTrendUtils.dateParser(currDate);
    currDay = dt.get(Calendar.DAY_OF_WEEK);
    return currDay;
}

/* Calculates and returns the previous day of the week from the current date*/
public static Integer getPrevDay(String currDate) {
    Integer prevDay;
    Calendar dt = MessageTrendUtils.dateParser(currDate);
    dt.add(Calendar.DATE, -1);
    prevDay = dt.get(Calendar.DAY_OF_WEEK);
    return prevDay;
}

/* Checks if the prev day is a Fed Holiday */
public static boolean isPrevDayFedHoliday(String currDate) {
    //if (prev day is a Fed Holiday) {
    //    return true;
    //}
    return false;
}

/* Checks if the FTP comm pt is operational */
public static boolean checkFtpComm() {
    //if (FTP comm pt != operational) {
    //    return false;
    //}
    return false;
}
}

```

File: MessageTrendConstants.java

```
package com.msg.utils;
```

```
public class MessageTrendConstants {
```

```

    public static final Integer WEEKLY_AVERAGE_2011 = 75;
    public static final Integer HISTORICAL_WEEKLY_AVERAGE_2010 = 81;
    public static final Integer CURRENT_SEASON_WEEKLY_AVERAGE = 95;
    public static final Integer LAST_SEASON_AVERAGE = 102;
    public static final Integer AVG_SAME_WEEK_LAST_YEAR = 79;

```

```

public static final String SEASON_SUMMER = "SUMMER";

public static final String LAB_COMMUNICATION_SYSTEM_FTP = "FTP";

public static final String TRUE = "true";
public static final String FALSE = "false";

}

```

File:MessageTrendUtils.java

```

package com.msg.utils;

import java.text.DateFormat;
import java.text.ParseException;
import java.text.SimpleDateFormat;
import java.util.Calendar;
import java.util.Date;

public final class MessageTrendUtils {

    public static Calendar dateParser(String str_date) {
        java.util.Calendar cal = null;

        try {
            DateFormat formatter = new SimpleDateFormat("yyyyMMdd");
            Date date = (Date)formatter.parse(str_date);
            cal=Calendar.getInstance();
            cal.setTime(date);
        } catch (ParseException e) {
            System.out.println("Exception :"+e);
            e.printStackTrace();
        }
        return cal;
    }
}

```

Main.aspx.cs

```
using System;
using System.Data;
using System.Configuration;
using System.Web;
using System.Web.Security;
using System.Web.UI;
using System.Web.UI.WebControls;
using System.Web.UI.WebControls.WebParts;
using System.Web.UI.HtmlControls;
using MySql.Data.MySqlClient;

public partial class _Default : System.Web.UI.Page
{
    protected void Page_Load(object sender, EventArgs e)
    {
        if (!Page.IsPostBack)
        {
            try
            {
                string connectionString =
System.Configuration.ConfigurationManager.AppSettings.Get("survery.Con
nectionString");
                MySqlConnection connection = new
MySqlConnection(connectionString);
                MySqlDataAdapter adapter = new
MySqlDataAdapter("select * from persontable", connection);
                DataSet ds = new DataSet();
                adapter.Fill(ds, "person");
                ddlPersonId.DataSource =
ds.Tables["person"].DefaultView;
                ddlPersonId.DataTextField = "personId";
                ddlPersonId.DataValueField = "personId";
                ddlPersonId.DataBind();
                try
                {
                    string connectionString1 =
System.Configuration.ConfigurationManager.AppSettings.Get("survery.Con
nectionString");
                    MySqlConnection connection1 = new
MySqlConnection(connectionString1);
                    MySqlDataAdapter adapter1 = new
MySqlDataAdapter("SELECT * FROM interfacetable", connection1);
                    DataSet ds1 = new DataSet();
                    adapter1.Fill(ds1, "Name");
                    DDIntName.DataSource =
ds1.Tables["Name"].DefaultView;
                    // DDIntName.DataTextField = "Inteface";
                    DDIntName.DataValueField = "iName";
                    DDIntName.DataBind();
                }catch(MySqlException exc){}
```

```

        }
        catch (MySqlException ex)
        {
            //catch exception
        }
    }
}

//
protected void btnSelectPersonOnClick(object sender, EventArgs e)
{
    try{
        String connectionString =
System.Configuration.ConfigurationManager.AppSettings.Get("survery.Con
nectionString");
        MySqlConnection connection1 = new
MySqlConnection(connectionString);
        //      String x = Session["Interfaceworkingon"].ToString();
        MySqlCommand cmd1 = new MySqlCommand("UPDATE
mssdb.stagingtable SET flag ='0'", connection1);
        cmd1.Connection.Open();
        cmd1.Prepare();
        cmd1.ExecuteNonQuery();
        connection1.Close();
    }
    catch (MySqlException ex)
    {
        //catch exception
    }
    this.Session["personId"] = ddlPersonId.SelectedValue;
    this.Session["batchProcessed"] = 1;
    this.Session["Interfaceworkingon"] = DDIntName.SelectedValue;
    Response.Redirect("Summary.aspx"); //can add any url parameter
to determine what action is needed
}

protected void ddlPersonId_SelectedIndexChanged(object sender,
EventArgs e)
{
}
}

```

Survey1.aspx.cs

```

using System;
using System.Data;
using System.Configuration;

```

```

using System.Collections;
using System.Web;
using System.Web.Security;
using System.Web.UI;
using System.Web.UI.WebControls;
using System.Web.UI.WebControls.WebParts;
using System.Web.UI.HtmlControls;
using MySql.Data.MySqlClient;

public partial class survery : System.Web.UI.Page
{
    protected void Page_Load(object sender, EventArgs e)
    {
        if (!Page.IsPostBack)
        {
            try
            {
                String ss = Session["batchProcessed"].ToString();
                int ist = int.Parse(ss);
                string connectionString =
System.Configuration.ConfigurationManager.AppSettings.Get("survery.Con
nectionString");
                MySqlConnection connection = new
MySqlConnection(connectionString);
                String str = "select * from stagingtable WHERE flag =
'0' AND Batch = " + ist + " AND InterfaceID =" +
Session["Interfaceworkingon"].ToString() + " ";
                MySqlDataAdapter adapter = new MySqlDataAdapter(str,
connection);

                DataSet ds = new DataSet();
                adapter.Fill(ds, "questions");

                Label1.Text =
ds.Tables["questions"].Rows[0]["qID"].ToString();
                Label2.Text =
ds.Tables["questions"].Rows[0]["interfaceID"].ToString();
                Label3.Text =
ds.Tables["questions"].Rows[0]["qType"].ToString();
                Label4.Text =
ds.Tables["questions"].Rows[0]["qDesc"].ToString();
                Label5.Text =
ds.Tables["questions"].Rows[0]["choiceA"].ToString();
                Label6.Text =
ds.Tables["questions"].Rows[0]["choiceB"].ToString();
                Label7.Text =
ds.Tables["questions"].Rows[0]["choiceC"].ToString();
                Label8.Text =
ds.Tables["questions"].Rows[0]["choiceD"].ToString();
                Label9.Text =
ds.Tables["questions"].Rows[0]["correctAns"].ToString();
            }
            catch { }
        }
    }
}

```

```

        Label10.Text =
ds.Tables["questions"].Rows[0]["Batch"].ToString();
        Label11.Text =
ds.Tables["questions"].Rows[0]["flag"].ToString();
        Label12.Text = DateTime.Now.ToString();
        Label13.Text = "1";
        //      Label13.Text =
radCity.SelectedItem.Value.ToString();
        Label17.Text = "1";
        Label18.Text = "3";
        Label20.Text = "00";
        int ccount =
int.Parse(Session["currentCount"].ToString());
        ccount = ccount + 1;
        Session["currentCount"] = ccount;

    }
    catch (MySqlException ex)
    {
        //catch exception
    }
}

protected void Button1_Click1(object sender, EventArgs e)
{
    DateTime newTime;
    try
    {
        String x1 = Label12.Text.ToString();
        String x2 = Label20.Text.ToString();
        if (x2 == "00")
        {
            newTime = DateTime.Now; ;
        }
        else
        {
            newTime = DateTime.Parse(x2);
        }
        DateTime oldTime = DateTime.Parse(x1);

        TimeSpan ts = newTime - oldTime;
        int difference = ts.Seconds;

        String z1 = Label17.Text.ToString();
        String z2 = Label9.Text.ToString();
        String ms = "0";
        if (z1 == z2)
        {

```

```

        ms = "1";
    }else {
        ms= "0";
    }

    String connectionString =
System.Configuration.ConfigurationManager.AppSettings.Get("survery.Con
nectionString");
    MySqlConnection connection1 = new
MySqlConnection(connectionString);
    MySqlCommand cmd1 = new MySqlCommand("INSERT into
surveyresults(personID,questionID,qType,Answer,correctAnswer,Outcome,T
ime,InterfaceID,surveyDate,batchID,answerMood)
VALUES(?personID,?questionID,?qType,?Answer,?correctAnswer,?Outcome,?T
ime,?InterfaceID,?surveyDate,?batchID, ?answerMood)", connection1);
    cmd1.Connection.Open();

    MySqlParameter parama = new MySqlParameter("?personID",
int.Parse(Session["personId"].ToString()));
    cmd1.Parameters.Add(parama);

    MySqlParameter paramb = new MySqlParameter("?questionID",
Label1.Text.ToString());
    cmd1.Parameters.Add(paramb);

    MySqlParameter paramc = new MySqlParameter("?qType",
int.Parse(Label3.Text.ToString()));
    cmd1.Parameters.Add(paramc);

    MySqlParameter paramd = new MySqlParameter("?Answer",
Label17.Text.ToString());
    cmd1.Parameters.Add(paramd);

    MySqlParameter parame = new
MySqlParameter("?correctAnswer", int.Parse(Label9.Text.ToString()));
    cmd1.Parameters.Add(parame);

    MySqlParameter paramf = new MySqlParameter("?Outcome",
ms);
    cmd1.Parameters.Add(paramf);

    MySqlParameter paramg = new MySqlParameter("?Time",
difference);
    cmd1.Parameters.Add(paramg);

    MySqlParameter paramh = new MySqlParameter("?InterfaceID",
Session["Interfaceworkingon"].ToString());
    cmd1.Parameters.Add(paramh);

    MySqlParameter parami = new MySqlParameter("?surveyDate",
DateTime.Now);
    cmd1.Parameters.Add(parami);

```

```

        MySqlParameter paramj = new MySqlParameter("?batchID",
int.Parse(Label10.Text.ToString()));
        cmd1.Parameters.Add(paramj);

        MySqlParameter paramk = new MySqlParameter("?answerMood",
int.Parse(Label18.Text.ToString()));
        cmd1.Parameters.Add(paramk);

        cmd1.Prepare();
        cmd1.ExecuteNonQuery();
        connection1.Close();
    }
    catch (MySqlException ex)
    {
        Response.Write(ex.Message.ToString());
    }

    try
    {
        string connectionString =
System.Configuration.ConfigurationManager.AppSettings.Get("survey.Con
nectionString");
        MySqlConnection connection = new
MySqlConnection(connectionString);

        MySqlCommand cmd = new MySqlCommand("UPDATE stagingtable
SET flag = 1 WHERE qID = ?qID AND qType = ?qType", connection);

        MySqlParameter param = new MySqlParameter("?qID",
Label11.Text);
        cmd.Parameters.Add(param);

        MySqlParameter param1 = new MySqlParameter("?qType",
Label3.Text);
        cmd.Parameters.Add(param1);

        cmd.Connection.Open();
        cmd.Prepare();
        cmd.ExecuteNonQuery();

    }
    catch (MySqlException ex)
    {
        Response.Redirect("Summary.aspx");
    }
    int z = int.Parse(Session["currentCount"].ToString());
    int m = int.Parse(Session["batchtotal"].ToString());
    if (z < m)
    {
        Response.Redirect("survey1.aspx"); //can add any url
parameter to determine what action is needed

```



```

    }
    else
    {
        Session["currentCount"] = 0;
        Session["batchtotal"] = 0;
        int n = int.Parse(Session["batchProcessed"].ToString());
        this.Session["batchProcessed"] = n+1;
        Response.Redirect("Summary.aspx");
    }
}

protected void RadioButtonList1_SelectedIndexChanged(object
sender, EventArgs e)
{
    Label17.Text = RadioButtonList1.Selected.Value;
    Label20.Text = System.DateTime.Now.ToString();
}

protected void RadioButtonList2_SelectedIndexChanged(object
sender, EventArgs e)
{
    Label18.Text = RadioButtonList2.Selected.Value;
}
}

```

Summary.aspx.cs

```

using System;
using System.Data;
using System.Configuration;
using System.Collections;
using System.Web;
using System.Web.Security;
using System.Web.UI;
using System.Web.UI.WebControls;
using System.Web.UI.WebControls.WebParts;
using System.Web.UI.HtmlControls;
using MySql.Data.MySqlClient;

public partial class Default2 : System.Web.UI.Page
{
    void Page_Init(object sender, EventArgs e)
    {
        Label myLabel = new Label();
        int y = int.Parse(Session["batchProcessed"].ToString());

        if (y == 1)
        {

```

```

        myLabel.Text = "If you have been notified to start the
survey please ";

    }
    else
    {
        myLabel.Text = "If you have just completed the survey go
back to NEDSS Messaging Solution Dashboard Screen and continue the
task. When you are asked to continue with the survey please ";

    }
    LinkButton link = new LinkButton();
    link.Text = "Click here...";
    link.ID = "LinkButtonTest";
    link.Click += new System.EventHandler(LinkButton1_Click);

    MyPanel.Controls.Add(myLabel);
    MyPanel.Controls.Add(link);

}

```

```

protected void LinkButton1_Click(object sender, EventArgs e)
{
    try{
        String s = Session["batchProcessed"].ToString();
        int ist = int.Parse(s);
        string connectionString =
System.Configuration.ConfigurationManager.AppSettings.Get("survery.Con
nectionString");
        MySqlConnection connection = new
MySqlConnection(connectionString);
        String str = "select COUNT(*) from stagingtable WHERE flag =
'0' AND Batch = " + ist + ";";
        MySqlDataAdapter adapter = new MySqlDataAdapter(str,
connection);

        DataSet ds = new DataSet();
        adapter.Fill(ds, "questions");

        String results = ds.Tables["questions"].Rows[0][0].ToString();
        int sm = int.Parse(results);
        this.Session["batchtotal"] = sm;
        if (sm > 0)
        {
            this.Session["currentCount"] = 0;
            Response.Redirect("survey1.aspx");
        }else{

```

```
        Response.Redirect("end.aspx");
    }

} catch (MySqlException ex)
{
    //catch exception
}

}
}
```

APPENDIX C – Evaluation Study Instruments

Question Type	Question
1	Did NEDSS MSS system receive any messages today?
1	Did NEDSS MSS receive more messages than the daily average today?
1	Did Labcorp report less messages than their daily average today?
1	Did Mayo report more messages than their daily average today?
1	Did Quest report more messages than their daily average today?
1	Have any messages failed validation today?
1	Were there any structural errors today?
1	Were there any content errors today?
1	Were there any duplicate messages reported today?
1	Is the MSS system currently running?
1	What is the current status of MSS Server?
1	How is the MSS message exchange status classified as?
1	How is the MSS message validation status classified as?
1	Which Lab reported more messages than other labs?
1	Did Mayo Report more messages than other labs?
1	Which lab reported the least messages today?
1	Did Quest Report more messages than other labs?
1	Which of the COMM points are not active today?
1	Which comm point received more messages today?
1	Which working comm point received the least number of messages or no messages
1	Were there any comm points restarted today?
1	How many comm points are not working now
1	Which comm point is highly used to receive messages
1	Did any MSS comm points failed during the last weekend?
1	How many messages failed validation today?
2	How many labs sent more messages today than their daily average?
2	What percentage of messages received today failed validation?
2	What is an appropriate action regarding the number of content errors received today?
2	What is an appropriate action regarding the number of duplicate errors received today?
2	Is the number of failed messages today more than the daily average?
2	What percentage of the error messages received today had structural errors?
2	Did MSS receive more structural errors today than the daily average for structural errors?
2	What percentage of the error messages received today had content errors?
2	Did MSS receive more content errors today than the daily average for content errors?
2	What percentage of the error messages received today were because of duplicate

	messages?
2	Did MSS receive more duplicate message errors today than the daily average for duplicate messages?
2	What is an appropriate action regarding the number of structural errors received today?
2	What is the best explanation for current MSS system health to be in Red?
2	What is the best explanation for current exchange health to be in Red?
2	What is the best explanation for current validation health to be in Red?
2	Which labs have abnormal reporting today? (Normal = +/- 20% than average)
2	Was there any event at the source today that could have affected message exchange?
2	Was there any event at the source today related to message validation?
2	What is your interpretation for the overall message exchange to be 20% less than daily average?
3	Shutting down of which COMM pt or points will have a maximum negative impact on message exchange?
3	Shutting down of which COMM pt or points will have the least negative impact on message exchange?
3	Based on historical data when can you not expect any messages to be reported?
3	Will a failure of FTP COMM point on Sunday have an impact on message exchange?
2	What is the correct distribution of errors today?
3	Which of the statement can be true regarding the less content errors than daily average?

APPENDIX D – Dashboard Prototypes

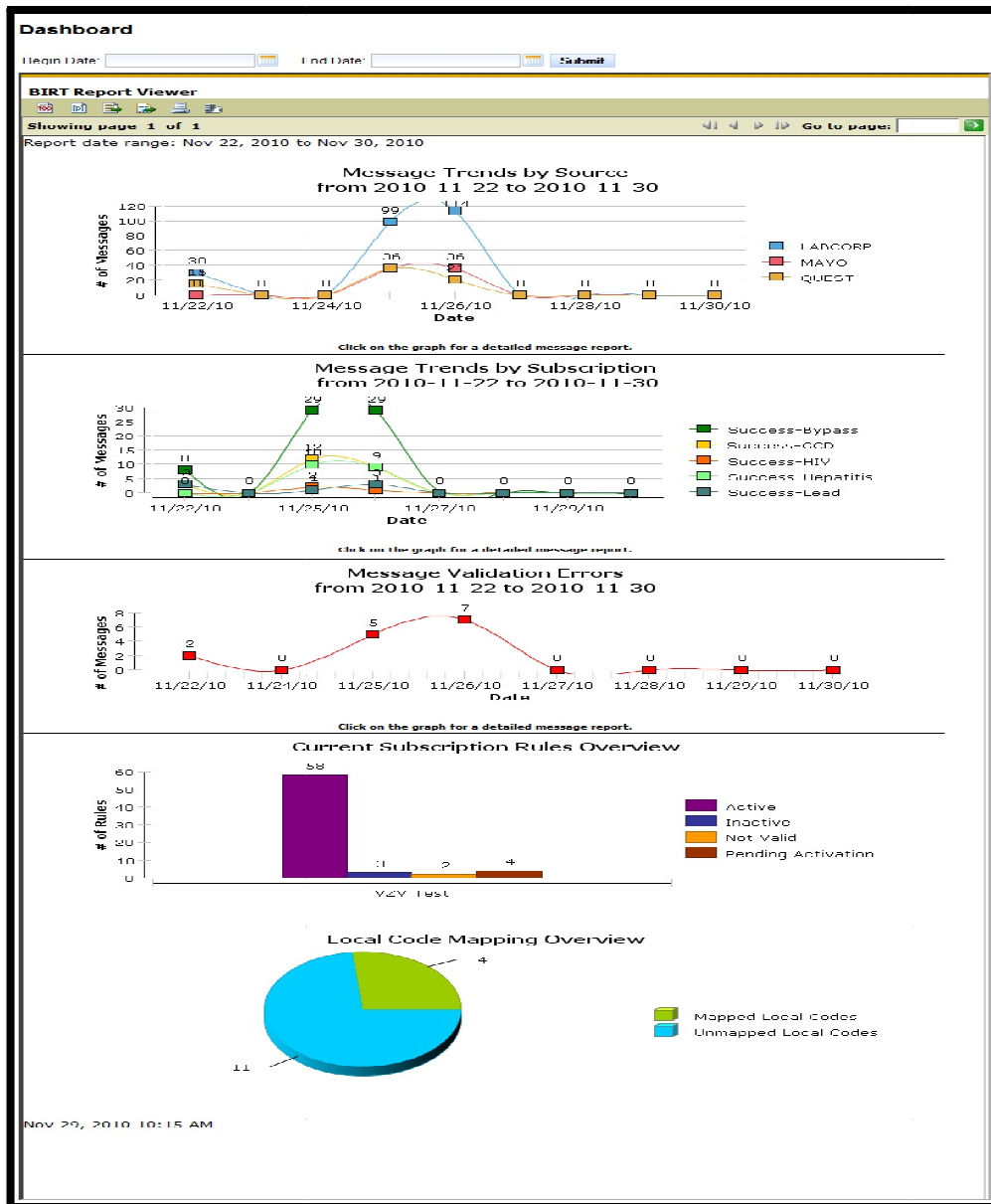
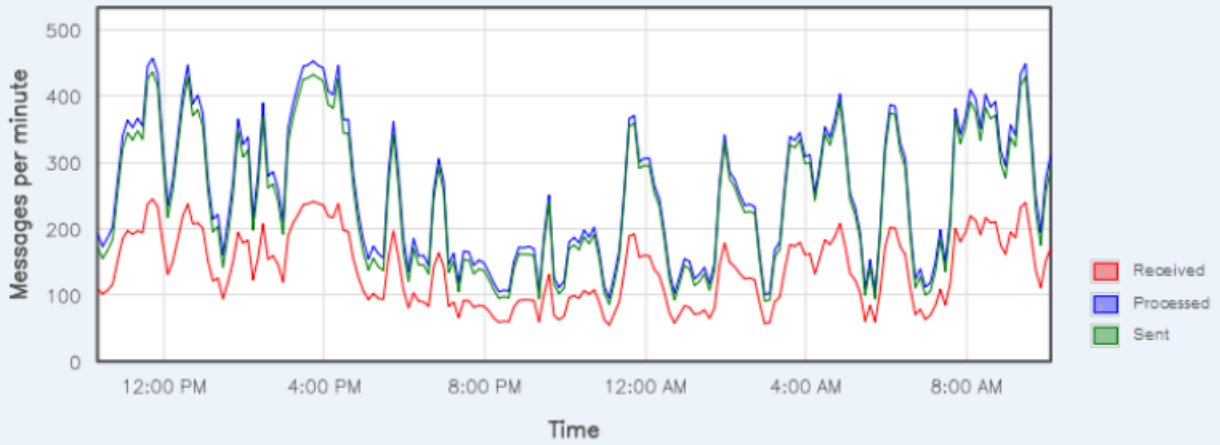


Figure 54: Traditional (old) MSS Dashboard integrated with BIRT Reporting

System Performance

| Last Hour | Last 24 Hours | Last 7 Days | Last 30 Days | Custom |

Total Engine Throughput (Last 24 Hours)



Received Messages	199,580
Sent Messages	347,875
Processed Messages	368,892

[Earlier Start Time](#) | [Later Start Time](#)

[Earlier End Time](#) | [Later End Time](#)

Memory Usage (Last 24 Hours)

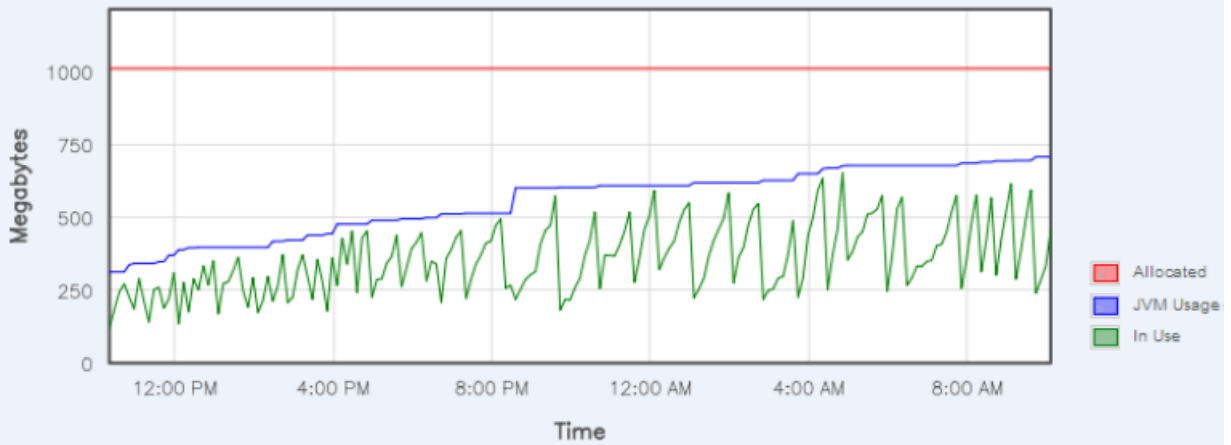


Figure 55: Supplement Interfaces- Orion Rhapsody Dashboard

Communication Point List

Filter For Showing All Communication Points

Filtered Communication Points							
Name	Action	Connections	Unprocessed	Waiting to be Sent	Received	Sent	Idle Time
📁 LatencyResponseDemo							
▶ LatencyFeeder (Demo:Message Feeder)	Stop	1	<input type="text" value="0"/>	<input type="text" value="0"/>	14,095	0	6 seconds
▶ Sink (Sink)	Stop	1		<input type="text" value="0"/>		14,092	6 seconds
▶ TCP Client (TCP Client)	Stop	20	<input type="text" value="0"/>	<input type="text" value="0"/>	14,092	14,095	6 seconds
▶ TCP Server (TCP Server)	Stop	20	<input type="text" value="0"/>	<input type="text" value="0"/>	14,095	14,092	6 seconds
📁 Map and Acknowledge							
📁 Cleanup							
▣ Cleanup EDI Files (Directory)	Start	0	<input type="text" value="0"/>		84		(Not Running)
▶ Cleanup XML files (Directory)	Stop	1	<input type="text" value="0"/>		28		39 minutes
▶ Sink (Sink)	Stop	1		<input type="text" value="0"/>		112	39 minutes
▶ EDI Messages (Directory)	Stop	1		<input type="text" value="0"/>		141	4 minutes
▶ Sink (Sink)	Stop	1		<input type="text" value="0"/>		0	1 day 4 hours
▶ Standard Feeder (Demo:Message Feeder)	Stop	1	<input type="text" value="0"/>	<input type="text" value="0"/>	14,162	14,162	5 seconds
▶ XML Messages (Directory)	Stop	1		<input type="text" value="0"/>		14,162	5 seconds
📁 VMS3							
▶ Main to Sub (Error Message Redirector)	Stop	1	<input type="text" value="0"/>	<input type="text" value="0"/>	178,769	178,769	2 seconds
▶ VMS3-Sink (Sink)	Stop	1		<input type="text" value="0"/>		176,897	1 second
▶ VMS3-TCPserver (TCP Server)	Stop	0	<input type="text" value="0"/>	<input type="text" value="0"/>	0	0	1 day 4 hours
▶ VMS3-Timer (Timer)	Stop	1	<input type="text" value="0"/>		10,312		2 seconds

[Start All](#) | [Stop All](#) | [Reset All Counters](#) | [Live tracked messages](#)

Figure 56: Supplement Interfaces- Comm Point Status

Route List

Filter For Showing ▼

Filtered Routes

Route Name	Action	Waiting	Currently Processing	Total Processed	Idle Time
LatencyResponseDemo					
LatencyGenRoute	Stop	<input type="text" value="0"/>	<input type="text" value="0"/>	62,265	1 second
LatencyServerRoute	Stop	<input type="text" value="0"/>	<input type="text" value="7"/>	31,134	1 second
Map and Acknowledge					
Batch Zip Route	Stop	<input type="text" value="0"/>	<input type="text" value="71"/>	311	0 seconds
Cleanup					
Delete Files Route	Stop	<input type="text" value="0"/>	<input type="text" value="0"/>	297	21 minutes
Message Generation	Stop	<input type="text" value="0"/>	<input type="text" value="0"/>	31,171	0 seconds
TypeSorter	Stop	<input type="text" value="0"/>	<input type="text" value="0"/>	62,342	0 seconds
VMS3					
VolumeMsgSender-Main	Stop	<input type="text" value="0"/>	<input type="text" value="0"/>	25,052	3 seconds
VolumeMsgSender1-MakeMsg	Stop	<input type="text" value="194,121"/>	<input type="text" value="7"/>	288,872	0 seconds
VolumeMsgSender2-Validate	Stop	<input type="text" value="0"/>	<input type="text" value="6"/>	288,871	0 seconds

[Start All](#) | [Stop All](#) | [Reset All Counters](#)

Figure 57: Supplement Interfaces- Rhapsody Route Monitor

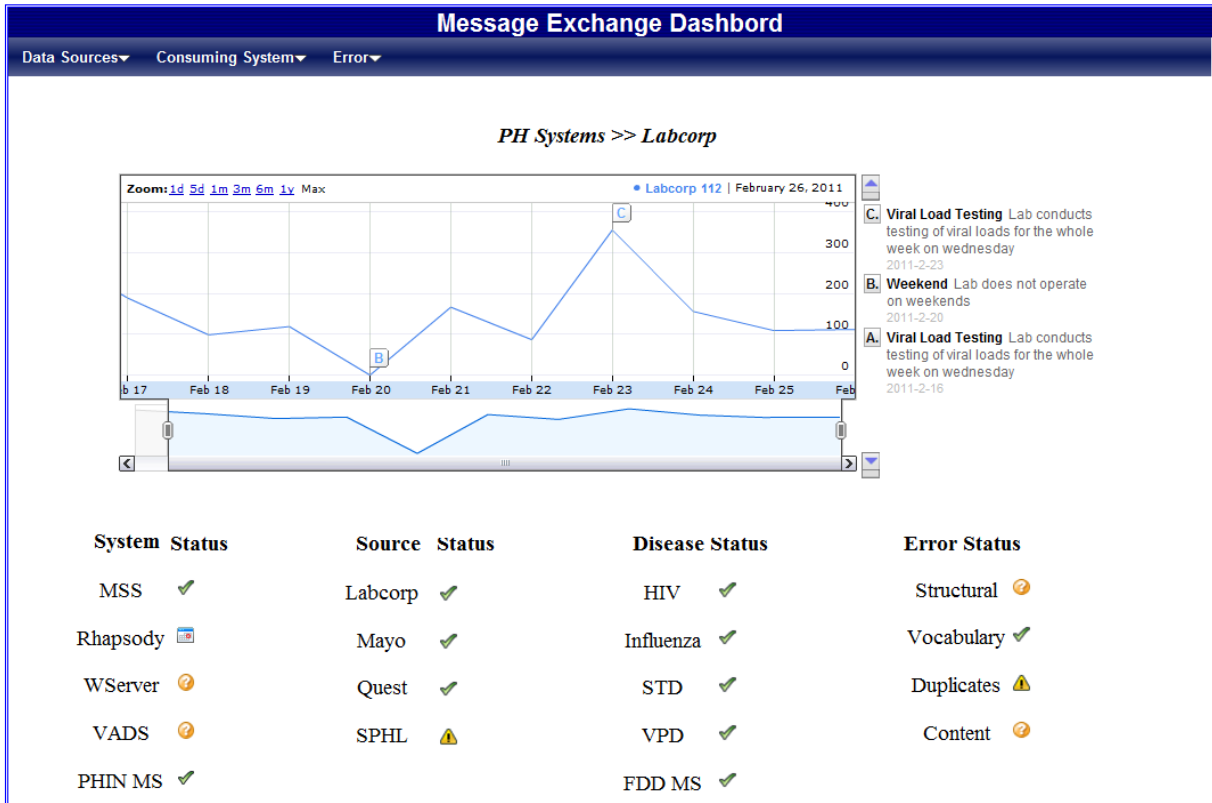


Figure 58 New Prototype Main Page with Labcorp Trend

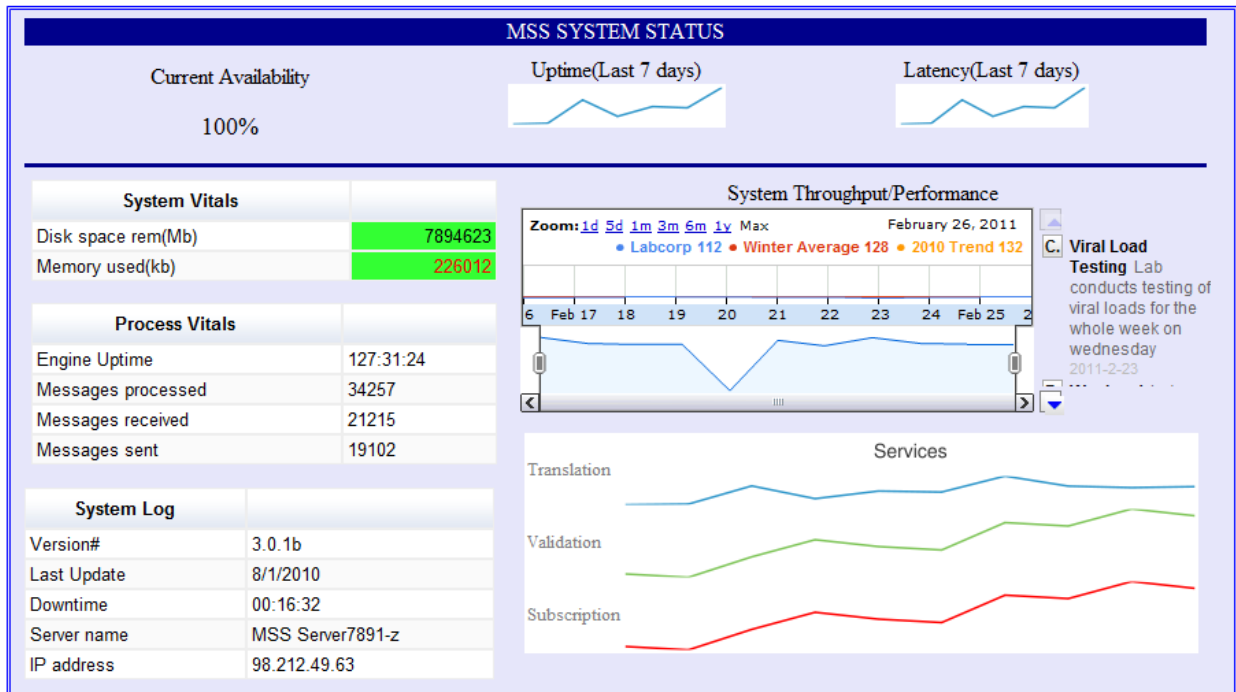


Figure 59: MSS System Dashboard Window

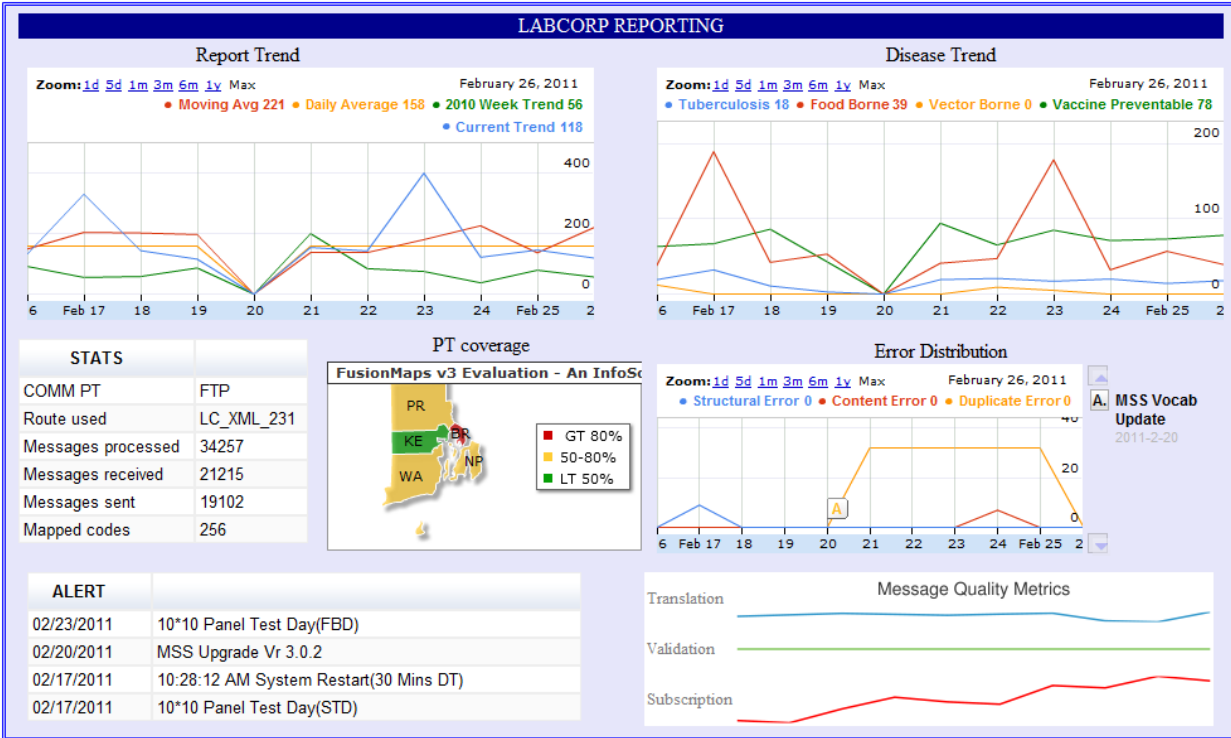


Figure 60: Labcorp Reporting Window

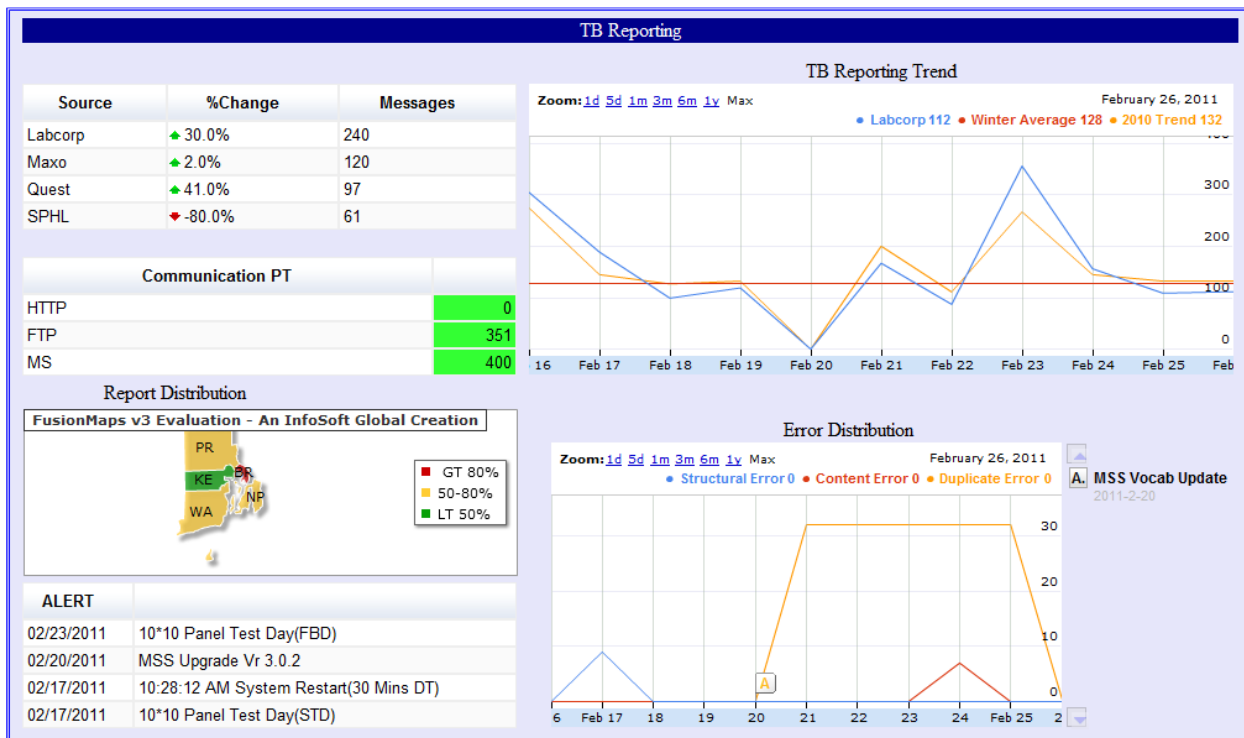


Figure 61: Labcorp TB Reporting Window

BIBLIOGRAPHY

- Arzt, N. H. (2007). *The evolution of public health information exchange*.
- Baddeley, A. (1972). Selective Attention and Performance in Dangerous Situation. *British Journal of Psychology* , 537-546.
- Ball, M. J. (2002). Better Health Through Informatics: Managing Information to Deliver Value. In P. O. Carroll, *Public Health Informatics and Information systems* (pp. 39-51). Springer.
- Beringer, D. B., & Hancock, P. A. (1989). Exploring situational awareness: A review and the effects of stress on rectilinear normalisation. *Proceedings of the Fifth International Symposium on Aviation Psychology* , Volume 2, pp. 646-651.
- Boehm, B. (1986). A Spiral Model of Software Development and Enhancement. *ACM SIGSOFT Software Engineering Notes*, (pp. 11(4):14-24).
- Boff, K., Kaufman, L., & Thomas, J. (1986). *Handbook of Perception and human Performance*. New York: Wiley & Sons.
- Burns, C. (2007). Supporting situation awareness through ecological interface design. *Proceedings of the human factors and ergonomics society 51st annual meeting*, (pp. 205-209).
- Casey, S. (1993). *Set phasers on stun*. Santa Barbara: Aegean.
- Christofferson, K. (1998). A longitudinal study of the impact of ecological interface design on deep knowledge. *International Journal of Human Computer Studies* , 729-762.
- Clark, R., Lyons, C., & Hoover, L. (2004). Graphics for learning: Proven guidelines for planning, designing, and evaluating visuals in training materials. *Performance Improvement* , 43(10).

- Dey, A. K. (2000). *Providing Architectural Support for Building Context-Aware Applications*. Atlanta: Georgia Institute of Technology.
- Dey, A. K. (2001). Understanding and Using Context. *Personal and ubiquitous Computing* , 4-7.
- Dey, A. K., & Abowd, G. D. (1999). Towards a Better Understanding of Context and Context Awareness. *Proceedings of the 1st international symposium on Handheld and Ubiquitous Computing*. HUC.
- Dinadis, N. &. (1995). Designing functional visualizations for aircraft system status displays. *Journal of Aviation Psychology* , 241–269.
- Drews, F. A., & Westenskow, D. R. (2006). The Right Picture Is Worth a Thousand Numbers: Data Displays in Anesthesia. *Human Factors: The Journal of the Human Factors and Ergonomics Society* , 48-59.
- Eckerson, W. (2006). *Performance Dashboards*. JohnWiley & Sons, Inc.
- Edmondson, W., & Meech, J. (1994). Putting Task Analysis into Context. *SIGCHI Bulletin* , (26) 59-63.
- Endsley M.R., Kaber D.B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, (pp. 461-492).
- Endsley, M. (1990). Predictive utility of an objective measure of situation awareness. *In: the Human Factors Society 34th annual meeting.*, 41–45.
- Endsley, M. (1990). Predictive utility of an objective measure of situation awareness. *Human Factors Society* , 41–45.
- Endsley, M. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37:32–64.

- Endsley, M. (2000). Theoretical Underpinings of Situation Awareness: A Critical Review. In G. D. Endsley MR, *Situation Awareness Analysis and Measurement*.
- Endsley, M. (2000). Validity and Use of SAGAT. In G. D. Endsley M, *Situation Awareness Analysis and Measurement*.
- Endsley, M. R. (1988). Situation awareness global assessment technique (SAGAT). *National Aerospace and Electronic Conference (NAECON)*. Dayton, OH.
- Endsley, M. R., Bolte, B., & Jones, D. G. (2003). *Designing for situation awareness: An approach to user-centered design*. Taylor & Francis.
- Endsley, M., & Kaber, D. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, (pp. 462-92.).
- Endsley, M., & Kiris, E. (1994). Information Presentation ofr Expert Systems in Future Fighter Aircraft. *Internation Journal of Aviation Psychology* , 333-348.
- Endsley, M., Mogford, R., & Allendoerfer, K. (1997). Effect of free flight conditions on controller performance, workload, and situation awareness: a preliminary investigation of changes in locus of control using existing technology. *Federal Aviation* .
- Endsley, M.R, (2000). Validity and Use of SAGAT. In G. D. Endsley M, *Situation Awareness Analysis and Measurement*.
- Endsley, M.R. (1988). Design and evaluation for situation awareness enhancement. In: the Human Factors Society 32nd annual meeting. *Human Factors Society*, (pp. p. 97-101.). Santa Monica, CA.
- Eshraghi, A. (Information Management Special Report). *Dashboards: The New Face of BI*. 2008.

- Farcot, R. (2010). Digital Dashboard Technology—Visualize The Possibilities. *INNOVATION* , 10 (15-20).
- Few, S. (2006a). *Common Pitfalls in Dashboard Design*. Proclarity.
- Few, S. (2006b). *The effective visual communication of data*. Sebastopol: O'Reilly.
- Few, S. (2008a). *Dashboard Design for Real Time Situational Awareness*. Inova Solutions.
- Few, S. (2008b). *Dual-scaled axes in graphs*. Retrieved April 25, 2010, from perceptualedge.com:
http://www.perceptualedge.com/articles/visual_business_intelligence/dual-scaled_axes.pdf
- Few, S. (2009). Dashboard design Beyond meters gauges and traffic lights. *Business Intelligence Journal* , 18-24 10(1).
- Flach, J. (1995). *Global Perspectives on the ecology of human machine systems*. New Jersey: Hillsdale.
- Fracker, M. (2001). Measures of situation awareness: review and future directions. *Armstrong Laboratory* .
- George, M. (1956). “The Magical Number Seven, Plus or Minus Two”. Some limits on our capacity for processing information. *The Psychological Review* , 81-97 (63).
- Gibson, J. (1988). *The senses considered as perceptual systems*. Boston: Houghton-Millin.
- Goldberg, L.R. (1970). Man versus model of man: A rationale, plus some evidence for a method of improving clinical judgement. . *Psychological bulletin* , 422-432.
- Graeber, R. (1996). Integrating human factors and safety into airplane design and operations. In B. H. Lowe, *Applied aviation Psychology* (pp. 27-38). Aldershot, UK: Avebury Aviation.
- Hancock, P. (1988). *Human Mental Workload*. Amsterdam: North-Holland.

- Hansoti, B. (2010). *Business Intelligence Dashboard in Decision Making*. Purdue University, PhD Thesis.
- HHS. (2010, Feb). Health Information Technology: Initial Set of Standards, Implementation Specifications, and. *Extracted from 45 CFR Part 170 RIN 0991-AB58* . Washington, DC, USA: HHS.
- Hoffer, J. A. (2002). *Modern systems analysis and design*. NJ: Prentice-Hall.
- Hurwitz, J., Halper, F. (2005). *Dashboards — Enabling Insight and Action*. Waltham: Hurwitz & Associates.
- Inc., J. S. (2009). A guide towards creating a dashboard people love to use. Juice Innovations.
- Itoh, J. S. (1995). An ecological interface for supervisory control of BWR nuclear power plants. *Control Engineering Practice* , 231-239.
- Jamieson, G. (2003). *EID Rationale Design Project*. Toronto: CIL- University of Toronto.
- Jamieson, G. A. (2001). *A prototype ecological interface for a simulated petrochemical process*. Toronto.
- Jernigan, D. B. (2002). Data Standards in Public Health Informatics. In P. O. Carroll, *Introduction to Public Health Informatics and Information System* (pp. 213-238). Springer.
- Jones, D. (2000). Overcoming representational errors in complex environments. *Human Factors* , 367-378.
- Klein, G. (2000). Analysis of situation awareness from critical incident reports. In G. D. Endsley MR, *Situation Awareness Analysis and Measurement* (pp. 51-71.). NJ: Lawrence Erlbaum Associates;.
- Kuhn, T. (1970). *The Structure of Scientific Revolution*. Chicago: University of Chicago Press.

- Kunapareddy, N., Mirhaji, P., Zhang, J., Srinivasan, A., & Michea, Y. (2005). Information Visualization for Quality Control in Health Data Exchange Platforms. *AMIA Annu Symp Proc.*, (p. 1013).
- Kushniruk, A., & Patel, V. (2004). Cognitive and usability engineering methods for the evaluation of clinical information systems. *Journal of Biomedical Informatics* , 56–76.
- Langmuir, A. D. (1976). William Farr: Founder of Modern Concepts of Surveillance. *Int Journal of Epi* , 8-13.
- Lee, J. D. (2006). Application of Ecological Interface Design to Driver Support Systems. *IEA*.
- Liang, L. Y. (2001). *Dashboards and scorecards: Executive Information Technology for the public sector*. Government Finance Review.
- Lumpkin, J. R. (2002). History and Significance of Information Systems and Public Health. In P. O. Carroll, *Public Health Informatics and Information Systems* (pp. 16-38). Springer.
- MacLean, A., Young, R. M. (1989). Design rationale: the argument behind the artifact. , *Proceedings of the SIGCHI conference on Human factors in computing systems: Wings for the mind*, (pp. 247-252).
- Magnuson, J. A. (2005). Electronic laboratory reporting in the usa: 2004 national snapshot survey. *The Journal on Information Technology in Healthcare* , 3:31–45.
- Malik, S. (2005). *Enterprise dashboards : design and best practices for*. Wiley.
- Marcus, A. (2006). Dashboards in your future. interactions. *ACM* , Volume 13 , Issue 1.
- Maskery, H., & J. M. (1992). CONTEXT: In the eyes of users and in computer systems. *SIGCHI* , (24-2) 12-21.
- Mclead. (2004). *Interface Engines for Health Care: 21st Century Trends*. Mclead Inc.

- Miller, G. (1956). The Magical Number Seven Plus or Minus Two: Some limits on our capacity for processing information. *Psychological Review* , 81-97.
- Mintzberg, H. (1973). *The nature of Managerial Work*. New York: Harper and Row.
- Mirhaji, P., Michea, Y., Casscells, S., & Zhang, J. (2005). Situational awareness in public health preparedness settings. *SPIE; Technologies for Homeland Security and Homeland Defense IV*, (pp. 81-91).
- Mirhaji, P., Srinivasan, A., Casscells, S. W., & Arafat, R. (2004). Semantic Approach to Public Health Situation Awareness - Design and Methodology. *Morbidity and Mortality Weekly Report (MMWR)* , 53(supp) 252.
- Moray, N. (2000). Are observers complacent when monitoring automated systems? *IES 2000/HFES 2000 Congress* (pp. 592-595). Santa Monica: Human Factors and Ergonomic Society.
- Morrissey, R. (2007). Tailoring performance dashboard content. *Business Intelligence Journal*, 1 , 12(4) 7-16.
- Nagel, D. (1988). Cockpit-crew systems design and integration. In E. W. Nagel, *Human Factors in Aviation* (pp. 263-303). San Diego: Academic Press.
- Nagel, D. C. (1988). Human error in aviation operators. *Human factors in aviation* , 263-303.
- National Transportation Safety Board. (1991). *Aircraft Accident Reports: US Air-SkyWest Air at LAX*. Washington D.C.: NTSB.
- National Transportation Safety Board. (1973). *Aircraft Accident Report : Eastern Airlines401*. Washington D.C.
- O' Carroll, P. (2002). *Public Health Informatics and Information Systems. 2002*:. Springer Verlag.

- Parnas, D.L. (1986). A rational design process: How and why to fake it,. *IEEE Transactions on Software Engineering*, v.12 n.2, , 251-257.
- Parsa, M., Richesson, R., & Srinivasan, A. (2004). Public Health Surveillance a Semantic Approach. *Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications*, (pp. 339-350).
- Parsa, M., Richesson, R., Smith, J. W., Zhang, J., & Srinivasan, A. (2004). Knowledge based Public Health Situation Awareness. *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense III*, (pp. 198-209).
- Parsa, M., Zhang, J., Smith, J. W., Majid, M., Casscells, S. W., & Lillibridge, S. R. (2003). Informatics Critical to Public Health Surveillance. *Sensors, and Command, Control, Communications, and Intelligence (C3I)Technologies for Homeland Defense and Law Enforcement II*, (pp. 151-163).
- Pritchett, A., Hansman, R., & Johnson, E. (1995.). Use of testable responses for performance-based measurement of situation awareness. In E. M. Garand DJ, *Experimental Analysis and Measurement of Situation Awareness*. (pp. 75-81.). Daytona Beach, FL;
- Rasmussen, J. (1983). Skills, Rules, and Knowledge signal, sign and symbol, and other distinctions in human performance. *IEEE Trans on System man and Cybernetics SMC* , 257-266.
- Remo, R. (2009). *Public Health & Electronic Health Information Exchange: A Guide To Local Agency Leadership*. The Institute for Public Health Informatics and Research.
- Royce, W. W. (1987). Managing the Development of Large Software Systems: Concepts and Techniques. *ICSE* , 328-339.

- Salvendy, G. (1987). *Handbook of Human Factors*. New York : Wiley & Sons.
- Sanders, M., & McCormick, E. (1992). *Human Factors in Engineering and Design*. New York: McGraw Hill.
- Sarter, N., & Woods, D. (1995). How in the World did i ever get into that mode? Mode error and awareness in supervisory control. *Human Factors* , 37:5-19.
- Selcon, S. (1990). Decision Support in Cockpit- Probably a good thing? *Proceeding of Human factors Society 34 Annual Meeting* (pp. 21-25). Santa Monia: Human Dactors Ergonomic Society.
- Sexton, A. G. (1988). Cockpit-crew systems design and integration. *Human factors in aviation* , 495-526.
- Sexton, G. (1988). Cockpit crew systems design and integration. In E. W. Nagel, *Human Factors in Aviation* (pp. 495-526). San Diego: Academic Press.
- Smith, W.M, Rydzak, C, Szeto, H., Buckeridge, D.L., Bravata, D., & McDonald, D. (2004). Systematic review: Surveillance systems for early detection of bioterrorism related diseases. *Annals of Internal Medicine* , 910-922.
- Srinivasan, A. (2009). NEDSS Messaging Solution- A National Initiative to accelerate Standard Message Exchange. *CSTE Annual Meeting*. Buffalo, NY.
- Srinivasan, A. C. (2008). Lessons in adopting the NEDSS Message Subscription Service. *Public Health Information Network Proceedings*. Atlanta.
- Srinivasan, A., & Abellera, J. (2010). State of NEDSS- Past, Present and Future. *Council of State and Territorial Epidemeologists Annual Conference*. Portland, OR.
- Srinivasan, A., Abellara, J., Danos, C., & McNabb, S. (2007). Improving Awareness in Message Exchanging Platform - A Knowledge Driven Approach. *AMIA*.

- Srinivasan, A., Abellera, J., Danos, S., & McNabb, S. (2008). Improving Awareness in Message Exchanging Platform- A Knowledge Driven Approach. *American Medical Informatics Conference Annual Meeting*. Washington DC.
- Srinivasan, A., Danos, S., McNabb, S., & Rhodes, B. (2008). Unified Messaging Solution for Biosurveillance and Disease Surveillance. *Association of Public Health Labs Annual Meeting*.
- Suchman, L. (1987). *Plans and Situated Action*. Boston: Cambridge Univ Press.
- Thorvald, P. (2009). Incorporating Cognitive Aspects in Digital Human Modeling. *HCI* (pp. 323-332). Springer.
- Toner, E. S. (June 2009). *Creating Situational Awareness: A systems approach*. Institute of Medicine Forum on Medical and Public Health Preparedness for Catastrophic.
- Tufte, E. R. (1983). *The Visual Display of Quantitative Information*. Cheshire, CT.
- Tufte, E. R. (2005). *Beautiful Evidence*. Cheshire CT: Press.
- Vicente, K. (2002). Ecological interface design: Progress and challenges. *Human Factors* , 62-78.
- Vicente, K. R. (1992). Ecological Interface Design- Theoretical Foundations. *IEEE Transactions on System Man and Cybernetics*, (pp. 589-606).
- Vidulich, M. (1992). Measuring Situational Awareness . *Proceedings of Human Factors Society 36th Annual Meeting*, (pp. 40-51).
- Vidulich, M. (1992). Measuring Situational Awareness. *Proceedings of Human Factors Society 36th Annual Meetin* , 40-47.
- Watson, H. G., Houdeshel, G., & Rainer, R. K. (1996). *Building executive information systems and other decision support applications*. NY: John Wiley & Sons Inc.

Wikipedia, I. (2010, December 20). *User Centered Design*. Retrieved December 20, 2010, from

Wikipedia: http://en.wikipedia.org/wiki/User-centered_design

Wright, M.C., Taekman, J. (2004). Objective measures of situation awareness in a simulated medical environment. *Qual Saf Health Care*, (pp. 65-71).