

**VISUALIZING AND INTERACTING WITH CONSTRUCTION PROJECT
PERFORMANCE INFORMATION**

Submitted in partial fulfillment of the requirements for

the degree of

Doctor of Philosophy

in

Civil and Environmental Engineering

Guzide Atasoy

B.S., Civil Engineering, Middle East Technical University

M.S., Civil Engineering, Middle East Technical University

Carnegie Mellon University
Pittsburgh, PA

January, 2013

© Copyright by Guzide Atasoy 2013

All Rights Reserved

Abstract

Project engineers/managers need individually relevant and customizable information in order to monitor and assess the performance of construction projects. Current project management tools do not meet this need completely; contributing to both lack of information in certain aspects and information overload in others. The results from a contextual inquiry; including a set of interviews with different project members highlight that 60% of the information needs of the project members are not met. In order to meet such needs, relevant data needs to be processed to the desired level of detail and be displayed in the right format that fits to the characteristic of the data and the task that the project member is trying to do (e.g., compare, analyze the trend, etc.). In addition, desired interactions with the data need to be supported.

The advent of information visualization techniques could help in addressing such needs and creating new opportunities to overcome information overload by enabling effective understanding of the datasets. However, there is no comprehensive study that formalizes how various visualization techniques can be represented, reasoned and visualized to meet the unique needs of different project members in the construction industry.

As a step towards improving the visualization of and interaction with individually relevant project performance information, I have developed a framework that provides effective visual representations to project members. Identifying and addressing the individual information visualization needs of construction project team members are challenging since these needs differ in terms of the performance metrics that are of interest, the level of detail and content of the available data, and the visual representations that support the task and suitable to the data being displayed. With the goal of addressing these challenges, the research described in this thesis has had three objectives: (1) identifying the commonalities and differences in information visualization requirements

of project members when monitoring and assessing the performance of construction projects, (2) developing a framework for defining, generating and visualizing individually relevant project performance information, and (3) formalizing the selection of visualization of individually relevant information for effectively monitoring the performance of construction projects.

The research approach involved contextual design method with semi-structured interviews and direct observations of construction project members working on different projects to identify the commonalities and differences in their information visualization needs. The research tasks include consolidation of interview results, synthesizing literature on information visualization, prototyping for testing purposes, representing and navigating project performance data, developing metadata for project data, and representing and reasoning about selection of appropriate and effective visual representations. The validation studies include cross-analysis of interview results, user tests and a synthetic test to evaluate (1) the commonalities and differences in information visualization needs of project members, (2) the usability of the components of the visualization framework and, (3) the precision, usability and usefulness of the formalism in terms of retrieving effective visual representations.

The contributions of the research include; (1) identification of commonalities and differences in information visualization needs of various project members, (2) development of a generic construction project performance visualization framework, and (3) development of a formalism that can enable project members to select and visualize individually relevant project performance information in an effective way. In terms of practical implications, the proposed research can be seen as a step towards improving the current way of performance monitoring by better utilization of visualization techniques leading to possible reduction in time for visualizing related information, and contributing to effective decision-making.

Acknowledgements

Pursuing a PhD is a challenging and enriching process and I would like to thank those whose support, guidance, and friendship have made this thesis possible.

Among all I must single out my advisor, Prof. Burcu Akinci, who has always been my inspiration, and role model. Words cannot express my heartfelt gratitude to her for her vision, wisdom, guidance, tolerance, interest, and endless support both academically and personally. At times when success seemed impossible, Prof. Akinci's patience and guidance lead me to the finish line. She thought me to do successful research through informative, intellectually stimulating, motivating and insightful meetings and feedbacks.

In addition, I would like to express my sincere appreciation to the members of my PhD committee Prof. Jim Garrett, Prof. Lucio Soibelman, Prof. Anind Dey and Assoc. Prof. Mario Bergés for their encouragement, suggestions, support and invaluable feedback. I would also like to thank them for finding time in their busy schedule to serve in my committee. The partial financial support of this research from National Science Foundation under CAREER Award program is greatly acknowledged.

I would like to thank to Mehmet Akinci who offered me summer internship opportunity, facilitated reaching out to industry professionals, and participated in my user studies. I would also like to acknowledge all the professionals from the construction industry, especially Mike Gentile, Jarred Musser, and Steve Owens from Trumbull Corp. who extended their assistance in gathering the empirical data to my research.

I am thankful to the former and current Mosaic research team members, especially Semiha Ergan, Anu Pradhan, Pingbo Tang, Fernanda Leite, Asli Akcamete, Xuesong (Pine) Liu, Saurabh Taneja, Te Gao, and Engin Burak Anil for many hours of discussions, talks, and constructive feedback and their company. I have benefited greatly from having known and worked with these outstanding people. I would like extend my

thanks to Semiha Ergan who has always provided me with invaluable guidance as a mentor through repeated iterations of this thesis, and personal support, encouragement and motivation.

I must acknowledge the CEE community and the departmental staff, especially Maxine Leffard and Andrea Francioni Rooney for always welcoming me with a warm smile and being accommodating to my queries.

I would also like to thank my friends. First and foremost, I am deeply grateful to my fiancé, Onur, who has always been there for me with his love, understanding, and endless emotional and practical support even through the toughest period of my PhD. Albeit not being in Pittsburgh, my long time friends were always a phone call away when I needed to talk and cheer up. Many thanks to Gul, Gulsah, Inci, and last but not the least Emre for being my closest friends. I would also like to thank Mudasar, Emre Can, Matineh, Ayse, Batu, Duygu, and Yigit, who made the department and Pittsburgh a better place to live.

Finally and most importantly I would like to thank to my beloved parents Basri and Nazan Atasoy, my sister Latife and my brother-in-law Mutlu Karakas. I feel truly blessed to have them as my parents and be able to make them proud. They are the source of joy and happiness and primary reason for my achievements, including this work. Without their love, care, encouragement, and financial and emotional support, I would not be who I am today. That is why I am dedicating this dissertation to Onur and my family.

Table of Contents

Abstract.....	iii
Acknowledgements	v
Table of Contents	vii
List of Figures.....	xii
List of Tables	xv
List of Abbreviations and Acronyms	xvii
Glossary	xviii
1 Introduction.....	1
1.1 Problem Statement	2
1.2 Motivating Case Study.....	4
1.3 Underlying Problems and Challenges.....	8
1.4 Overview of the Research Approach	10
1.5 Objective, Scope and Assumptions of the Research.....	12
1.6 Research Questions	14
1.7 Research Method	19
1.8 Readers' Guide to the Dissertation	21

2 Information Visualization Requirements of Project Members for Construction

Project Performance Monitoring 23

2.1	Motivating studies depicting the need for identifying information visualization requirements of project members.....	26
2.2	Research Objectives and Scope	30
2.3	Background Research	31
2.3.1	Performance Monitoring Frameworks and Approaches for Construction Projects.....	31
2.3.2	Differences in Performance Metric Needs of Project Members.....	36
2.3.3	Visualization of Project Performance in Construction Management Domain	38
2.4	Research Method	41
2.4.1	Contextual Design.....	42
2.4.2	Pilot Study and Recruitment	43
2.4.3	Contextual Inquiry	44
2.4.4	Work Modeling.....	45
2.4.5	Consolidation	49
2.5	Findings.....	51
	Overview.....	51
2.5.1.....		51
2.5.2	Commonalities	55
2.5.3	Differences in Metric Queries.....	58
2.5.4	Major Differences Based on Demographics	72

2.6	Analysis Results and Validation	77
2.7	Conclusions	86
3	Construction Project Performance Visualization: A Framework towards the Development of a Performance Visualization Environment	90
3.1	Introduction.....	90
3.2	Case Studies and Needs	93
3.2.1	Availability of the Information in the Required Form.....	93
3.2.2	Software tools designed for Construction Projects and Information Visualization	94
3.3	Research Objective and Scope.....	96
3.4	Background	97
3.4.1	Construction Project Performance Monitoring Frameworks developed for the AEC Domain.....	97
3.4.2	Construction Project Performance Visualization Frameworks.....	99
3.4.3	Generic Information Visualization Frameworks	100
3.4.4	Taxonomies and Findings on Visual Representations, Data types, and Interaction Tasks	102
3.5	Research Method	104
3.6	Findings and Discussions.....	106
3.6.1	Overview of the Construction Project Performance Visualization Environment.....	106
3.6.2	Construction Project Performance Visualization Framework	109
3.7	Validation.....	124

3.8	Conclusion	131
4	A Formalism for Effective Visualization of Construction Project Performance through Individually Relevant Information.....	133
4.1	Introduction.....	133
4.2	Objective and Scope	136
4.3	Background	137
4.3.1	Research Studies on performance visualization studies in CM domain	137
4.3.2	Research Studies on Information Visualization.....	138
4.3.3	Formalisms Towards Generation of Visual Representations.....	141
4.4	Elicitation and Representation of Heuristics.....	144
4.4.1	Elicitation of Heuristics	144
4.4.2	Representation of Heuristics	150
4.5	Reasoning with the Heuristics	155
4.6	Validation.....	165
4.6.1	Testbed for the synthetic tests with the formalism and its performance.....	166
4.6.2	Testbed for the Survey and its performance	175
5	Conclusion and Future Works.....	184
5.1	Summary of Contributions.....	184
5.1.1	Identification of the Commonalities and Differences in Project Performance Visualization Needs of Project Members	184
5.1.2	Development of a Construction Project Performance Visualization Framework	187

5.1.3	Representation of and reasoning with information visualization related heuristics	188
5.2	Practical Implications.....	192
5.2.1	Supporting project members for project awareness.....	192
5.2.2	Assisting software developers / other domains.....	193
5.3	Future Research	193
5.3.1	Extending the Elicitation of Information Requirements.....	193
5.3.2	Extending the Framework.....	194
5.3.3	Extending the Formalism.....	197
6	Bibliography	200
7	Appendices.....	225
7.1	Appendix A: Performance Metric Categories and Related Performance Metrics Identified from Prior Research Studies.....	226
7.2	Appendix B: Consolidated Interview Results.....	238
7.3	Appendix C: Results of the PCA analysis of m queries	261
7.4	Appendix D: Heuristics and Visual Representations.....	263
7.5	Appendix E: Survey	280
7.6	Appendix F: Synthetic Tests.....	286

List of Figures

Figure 1.1: Scatter Graph versus Column Chart for comparing the costs of activities	7
Figure 1.2: The vision on enabling project members to effectively visualize project performance information	11
Figure 2.1: The differences in the way project members want to track the Weekly Cost Reports and the associated purposes of usage divided according to the responsibilities of the project members.....	29
Figure 2.2: Flow Model for a project engineer (interviewee #7).....	46
Figure 2.3: Artifact Model for a project engineer (interviewee #7)	47
Figure 2.4: Example dashboard: Dashboard of a superintendent of a highway project ...	48
Figure 2.5: Example Dashboard: Dashboard of an owner representative of a building project	48
Figure 2.6: Affinity diagram of the consolidated interview results	50
Figure 2.7: The commonalities identified based on the interviews	56
Figure 2.8: Distribution of Performance metric categories amongst all metric queries ...	59
Figure 2.9: Visual Representations selected by project members	63
Figure 2.10: The distribution of the visual representations amongst the performance metric categories	64
Figure 2.11: Distribution of intents.....	66
Figure 2.12: Distribution of performance metrics based on all metric queries	68
Figure 2.13: Notifications status based on controllability of the performance metric queries	72
Figure 2.14: Performance Metric Categories based on project member responsibility....	74
Figure 2.15: Trade-off between exchange of information and productivity performance metric categories based on responsibility of the project members (the numbers are not normalized for the number of interviewee)	74
Figure 2.16: Distribution of the identified metric queries based on their performance metric categories and the types of projects interviewees are experienced with.	75
Figure 2.17: Distribution and Comparison of metric queries based on CoIs and the types of projects interviewees are experienced with	76

Figure 2.18: Distribution and Comparison of metric queries based on CoIs and the responsibilities of the interviewees	76
Figure 2.19: The color-coded representation of coefficients of original variables in forming PCs.	79
Figure 2.20: The dataset plotted against the first two PCs.	80
Figure 2.21: The Overlap and Goodall similarity analyses using the equations in Boriah et al. 2008.....	83
Figure 2.22: Similarity matrix for Overall measure.....	84
Figure 2.23: Similarity matrix for Goodall measure.....	84
Figure 3.1: Research Method.....	105
Figure 3.2: A Conceptual Model for Project Performance Visualization Environment.	108
Figure 3.3: An affinity diagram depicting the requirements of project members for monitoring the performance of construction projects.....	110
Figure 3.4: UML Class diagram representation of <i>ProPerVisualWork</i>	111
Figure 3.4: Primary <i>ProPerVisualWork</i> Components	114
Figure 3.5: Secondary Components.....	117
Figure 3.6: Data Characteristic defined in prior research studies and the data characteristics modeled in <i>ProPerVisualWork</i>	118
Figure 3.7: The sparse distribution of the contextual inquiry participants	125
Figure 3.8: Venn Diagram of Visual Representations	130
Figure 4.1: Examples of the <i>heuristics</i> on visual representations, where problems are identified with ellipses and boxes, and ND refers to not-defined.....	139
Figure 4.2: Data Variability	147
Figure 4.3: The distribution of publications describing the <i>heuristics</i> included in this study based on their year of publication	150
Figure 4.4: A Summary of steps of the reasoning mechanism	156
Figure 4.5: Composing and Filtering LoDs	157
Figure 4.6: Project data metadata.....	159
Figure 4.7: Identification of Data Characteristics.....	160
Figure 4.8: The Distribution of Heuristics based on how they use the components of the framework	161

Figure 4.9: Looking up heuristics	162
Figure 4.10: Usage of Heuristics	171
Figure 4.11: The number of visual representations used in the test cases	173
Figure 4.12: Box plot distributions	179
Figure 5.1 Visual Representations Metadata	197

List of Tables

Table 2.1: Differences in five project members' interests in monitoring the weekly cost performance of activities.....	27
Table 2.2: Performance metric categories and subcategories identified in literature	36
Table 2.3: Differences in information needs of project members identified from construction monitoring related research efforts	38
Table 2.4: Demographics of the project members that are interviewed	44
Table 2.5: 132 Performance Metric Queries combined under short descriptions ignoring specific LoDs	52
Table 2.5 Cont'd: 132 Performance Metric Queries.....	53
Table 2.6: Summary of performance metric categories per job title of interviewee	55
Table 2.7: The CoIs selected by different responsibilities of interviewees	60
Table 2.8: Enumeration and Types of CoIs and LoDs.....	62
Table 2.9: Numerical values assigned to metric categories for the PCA analysis.....	78
Table 2.10: Latent and cumulative latent values for the PCA analysis	79
Table 2.11: The distribution of variables according to the magnitude of their coefficients in the representation of the PCs	81
Table 2.13: The statistics of similarity matrices	85
Table 3.1: Construction Project Visualization Frameworks	100
Table 3.2: Existing Information Visualization Frameworks.....	102
Table 3.3: Framework Components.....	113
Table 3.4: Demographics	124
Table 3.5: Literature review of project data / information visualization frameworks	126
Table 3.6: Coverage of the framework based on previous research studies	126
Table 3.7: Usability of framework components	128
Table 3.8: Component Usefulness	131
Table 4.1: Research studies used in <i>heuristics</i> identification	149

Table 4.2: Demonstrative set of <i>Heuristics</i>	153
Table 4.3: The number of enumerations of heuristics table major variables and the number of heuristics that capture a value for these variables	154
Table 4.4: Metric queries with composite intents.....	166
Table 4.5: Test Cases	168
Table 4.6: Facets of Visual Representations for the initial four metric queries	169
Table 4.7: Precision of formalism.....	169
Table 4.8: Frequency of highest and lowest number of visual representations out 163.	173
Table 4.9: Survey questions.....	176
Table 4.10: The selection of charts within facets of 1s and 2s as the 2 nd chart option ...	177
Table 4.11: Survey results.....	178
Table 4.12: The statistical results for the data survey data	178
Table 4.13: Mean and Max. scales assigned to Charts	181

List of Abbreviations and Acronyms

AEC Architecture Engineering Construction

BIM Building Information Model

CAD Computer Aided Design

CM Construction Management

CoI Concept of Interest

D_dim Data Dimension

D_var Data Variability

D_str Data Structure

HCI Human Computer Interaction

LoD Level of Detail

PCA Principal Component Analysis

ProPerVisualWork Project Performance Visualization Framework

Glossary

Concept of Interest: Facets of major subject of interest of each metric query. For example, if the metric query of the user is “actual versus budgeted cost of activities performed last week”, its concept of interest is “activity”. Similarly, the concepts of interests can be project, foreman, component etc.

Heuristics: Findings from previous research studies about the relationship and characteristics of visualization techniques, their categorizations, and their performance comparisons

Level of Detail: The granularity of the requested performance metric in terms of time (e.g., monthly, weekly), space (e.g., building, floor), or Concept of Interest (e.g., foremen, component)

Metric Query: Each performance metric that is required by project members for monitoring an aspect of a single or multiple project(s). A metric query is majorly composed of the performance metric category, concept of interest and level of detail of the requested information, and the interaction and visual representation needs of the users.

Performance metric: “A measurable quantity that indicates some aspect of performance” (Deru and Torcelini 2005, pg 5).

Visual representation: “A collection of graphical symbols that visually encode causal, functional, structural, and semantic properties and relationships of a represented world” (Sedig and Sumner 2006, pg 2). This term captures tables, texts, graphs, maps, diagrams, icons and networks.

CHAPTER 1

1 Introduction

While managing and monitoring the performance of construction projects, project team members need customizable and individually relevant project information. However, currently, such needs of project members are not supported. Moreover, the way that they desire to visualize the data that they are interested in is not always ideal for the type of decision that they would like to make or the type of data that they would like to visualize.

The goal of this PhD research is to develop an approach to enable construction project members to visualize project performance information in an interactive, and individually configurable way. The scope of the proposed work includes identifying the commonalities and differences in information visualization needs of project members, identifying a generic framework to capture information items required to enable construction project performance visualization, and formalizing an approach to enable project members to effectively visualize and interact with such information.

The following sections present the importance of the problem, the challenges for effectively visualizing project performance information, the research vision, objectives and scope, and overview of research questions and methods.

1.1 Problem Statement

While managing construction projects, project team members need to easily access, customize, and visualize project performance data that are relevant to them (Cadiz et al. 2002; Balfanz et al. 2003; Russell et al. 2009). Relevant information refers to different context and detail of project information that are required by project members based on their responsibilities and interests in the projects. However, such needs are not always fully supported. Various commercially available software tools available for construction industry support different tasks associated with construction management, such as construction monitoring, collaboration and communication, cost control, schedule control, and information management. However, these tools serve a general audience of construction engineers and managers, and lack necessary means and methods to generate customized views per individuals. Hence, the tools, currently used at job sites for project performance monitoring and visualization, do not always support the project members' expectations and are not customizable to visualize performance data per individual preferences of various parties.

Monitoring and assessing the performance of a project is claimed to be one of the most important tasks of construction project management (Memon et al. 2007). Through an initial set of fourteen interviews with different project members, I identified that currently 78% of the project performance visualization requirements of project members were not available or not in a format required by them. Hence, my PhD research has been driven by the vision that a project visualization environment is required to manage and monitor the construction projects so as to effectively visualize the project performance data and to

make sense of it. By analyzing the interview results (i.e. metric queries) that I performed with different project members, I observed that their information visualization needs change based on various factors, such as their responsibilities, decisions to be made, experience levels, and characteristics of the project. These information visualization needs show commonalities and differences in terms of the monitored performance metrics, the level of detail of the metric queries, and the frequency and ways of visualizing these metric queries. Given these commonalities and differences, there is a need to understand the benefits and limitations of visualization techniques so that appropriate ones can be suggested per individual needs. Hence, I explored the research efforts done in Human Computer Interaction (HCI) domain on information visualization.

Several researchers studied and pointed out the potential advantages of information visualization within the Human Computer Interaction (HCI) domain. These studies target exploring types and relationships of data, visual representations and interaction intents, as well as effectiveness of visual representations. In these studies, several researchers identified, categorized and mapped visual representations (e.g., Keim 2002; Hoffman and Grinstein 2002) to data types (e.g. Schneiderman 1996; Myatt and Johnson 2009), and users' interaction needs (e.g., Amar et al. 2005; Zhou and Feiner 1998), and they evaluated the effect of visual representations in augmenting human cognition (Card 2008). These studies differ in terms of their approaches and focuses. For instance, the developed approaches include defining taxonomies (e.g., Schneiderman 1996), categorizing visual representations (e.g., Lohse et al. 1994), comparing performance of visual representations (e.g., Kumar and Benbasat 2004), and mapping interaction tasks to

specific visualization techniques (e.g., Chengzhi et al. 2003). Besides, some of the studies have focused on visualization of specific datasets, or interaction types. For instance, Aigner et al. (2007) explored time-oriented datasets and Chin et al. (2009) explored dynamic data-streams. While such wide variety of studies and findings exist, they have not been synthesized and formally represented to come up with a unifying framework. Thus, there is a need for synthesizing and formalizing the knowledge on information visualization. However, such studies might not be easily applicable to each domain due to variations in data types and context, and the purposes of users for visualizing and interacting with data. Hence, domain-specific customization of visualization of information is required (Heer et al. 2005).

Applying the findings identified within the HCI domain to the AEC domain requires identifying the information visualization needs of construction project members, capturing and representing the knowledge on information visualization, and associating and mapping the findings of the HCI domain to the needs of the construction industry. As a step towards this, there is a need to form the knowledge on commonalities and differences in information visualization needs of project members and to formalize an approach for enabling effective and efficient means of visualizing project performance information for construction management.

1.2 Motivating Case Study

Within the scope this research study, I have performed semi-structured interviews with fourteen project members in order to understand how they would like to monitor the

performance of construction projects. During the interviews, I captured the context of the project information they monitor, in addition to how they would like to visualize and interact with such information. As a result, I observed a total of 132 performance metric queries from the interviews. Through the detailed analysis of the interview results, I observed that the information visualization needs of project team members change based on various factors, such as their responsibilities, decisions to be made, experience levels, and characteristics of the project. These information visualization needs have commonalities and differences in terms of the performance metrics monitored, the level of detail of the metric queries, and the frequency and ways of visualizing these metric queries. The commonalities and differences in metric query visualization needs of project members are discussed in detail in Chapter 2. Currently, there are no commercially or openly available tools that support visualization of performance information according to individual needs of the project members. Moreover, there are no research papers that studied the extent that these information visualization needs differ. Hence, this research study is performed to address the need of capturing and representing the project performance visualization needs of project members.

In addition to understanding the project performance visualization needs of project members, project members should be provided with effective visual representations to support their decision-making process. For instance, during the interviews, I identified that a project engineer wants to see the activities' actual and budget costs compared on a scatter graph with the intent of identifying the activities with the largest cost overruns. A cluttered view generated in the scatter graph was caused due to plotting large number of

activities and this made it more difficult to perform pair-wise comparisons as shown in Figure 1.1a. The problem is mainly caused by the fact that scatter graphs are best suited for continuous data sets (Schonlau 2003); however, the cost codes are ordinal (ranked categorical data) numbers. Helfman and Goldberg (2007) stated that column charts are better suited to perform comparisons than scatter graphs. Moreover, column charts can be used to plot categorical data as shown in Figure 1.1b. In order to address the intent of the user in observing the highest cost-overflow activities, these cost-overruns can be highlighted with color coding; such as using red colored columns for the top five highest cost-overruns.

With the use of two columns for each activity's actual and budget costs, the chart might still look cluttered and use of stacked column chart could be suggested. However, Cleveland and McGill (1984) stated that in order to measure the relevant magnitude, position is a better indicator than length. In other words, instead of using a stacked graph and let user compare the lengths of columns on top of each other, it would be faster to observe the differences in a standard column chart by looking at the top position of the columns. In order to eliminate the problem of reading values, dynamic labeling of the values can be provided. Similarly, Kumar and Benbasat (2004) performed user tests and suggested the usage of dynamic labeling in overcoming ambiguity. Hence, as compared to a scatter graph or stacked column chart, column charts with dynamic labeling are better suited for this comparison task. In this way, the suggested charts would eliminate the problem of clutter; which provides faster observation of data points. In addition to assessment of visual representations in terms of the speed in observing data points,

various research studies explored the effectiveness of visual representations in terms of accuracy.

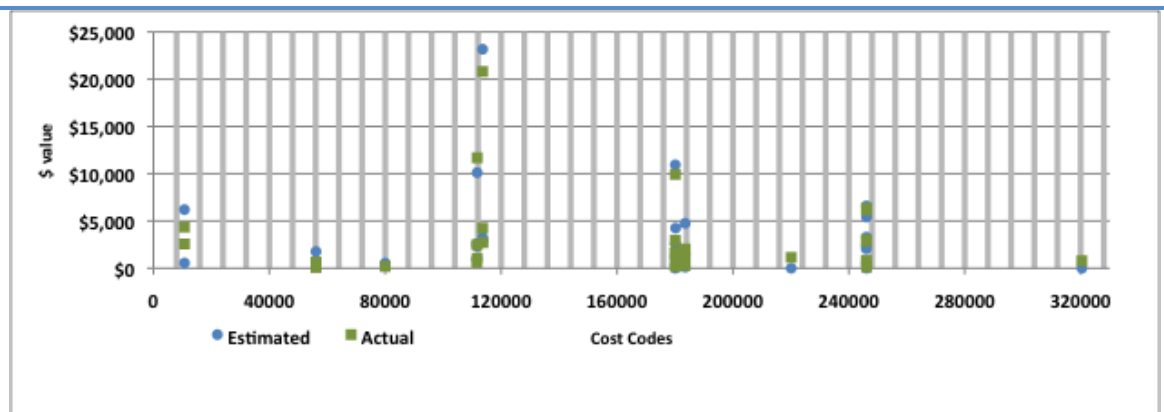


Figure 1.1a: Activity cost presented in a scatter plot

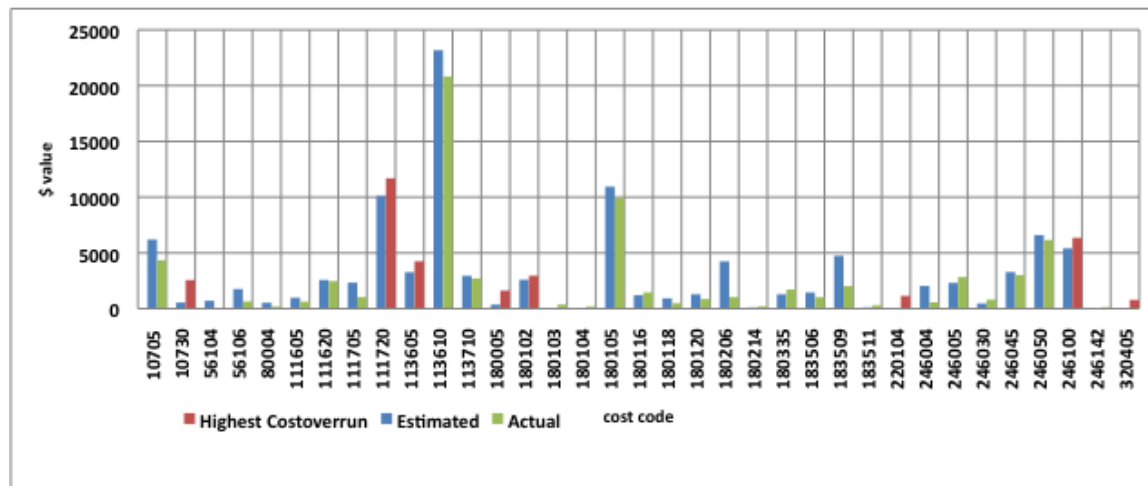


Figure 1.1b: Activity cost presented in a column chart

Figure 1.1: Scatter Graph versus Column Chart for comparing the costs of activities

In conclusion, through the initial case studies, ineffective visualization of project performance information is observed. This problem is mainly caused by limited knowledge on the different project performance monitoring needs of project members, and the effective ways to visualize this information. In the following subsection (Section

1.3), the underlying problems and challenges associated with project performance visualization are discussed.

1.3 Underlying Problems and Challenges

The motivating case demonstrated that currently project members do not have access to project performance information in the way they desire and this results in wasted time or losing valuable insights. Hence, detailed exploration of the information visualization needs of project members is required to understand the main features of project performance visualization and the extent to which they vary. In doing so one can identify and describe the patterns of the commonalities and differences in project members' performance visualization needs to suggest more effective ways.

Information visualization can lead to new insights, more efficient decision making and improved problem solving capabilities (Gershon et al. 1998; Dull and Tegarden 1999). In the motivating case, the situation within which the visual representation requested by an interviewee was not necessarily the best based on his purpose of visualization and the type of data that needed to be visualized is discussed. Indeed, a scatter plot was requested to visualize delta cost of activities by a project engineer. The dataset is composed of activity codes that are ordinal in nature. However, a scatter plot is better suited for continuous data while column chart can be used for ordinal data. Thus, it is necessary to identify the factors that impact the selection of an effective visual representation for a given context. Although various research efforts studied information visualization, there is no single research that has identified, captured, represented, and

reasoned about the findings from prior research studies in a formalized and integrated way. Moreover, relationships between time, resources, project requirements, and cost are not fully communicated in today's traditional graphical representation in the AEC domain and a fundamental model of project control data representation and improved information displays are needed (Liston et al. 2001; Songer and Hays 2003). The initial motivating case study presents the differing needs of project members, and the varying effectiveness of visual representations. However, the extent of project performance visualization needs of project members, and the ways to enable representation and facilitation of such needs has not been formalized for the construction domain. Hence, one of the goals of my PhD research has been to develop a general framework that captures and addresses differing information visualization needs of project members.

Although there are hundreds of visual representation options, not all of them are effective in a given situation. Data characteristics and interaction intents, and the context and LoD of the information might impact the selection and generation of different visual representations. Furthermore, Horn (2001) state that in the last 50 years, hundreds of visual representations appeared and this domain has been rapidly evolving over the last decade. Hence, in order to support the use of possible future representations to be developed, the findings of previous research studies should be represented in a easily traceable and maintainable format.

In order to enable selection and visualization of individually relevant information, a formalism needs to be developed. Indeed, in visualization summit in 2007, it is stated that

“Yet, no methodology, on describing how to provide visualization is generated” (Burkhard et al. 2007). Hence, there is a need to formalize a project performance monitoring framework and corresponding representation and reasoning mechanisms to facilitate the project members to effectively and efficiently visualize their performance metrics.

1.4 Overview of the Research Approach

Considering the limitations of the current practices and needs for overcoming the discussed challenges, I have developed an approach that enables project members to visualize and interact with individually relevant information. Figure 1.2 is an IDEF0 diagram providing an overview of the proposed approach which leverages project information with performance metric requirements of project members and augments it with visual representations and techniques according to the characteristics of the existing data, and the intent of users.

In order to address different information visualization and interaction needs of project members; first it is necessary to understand how different such needs are. The first research question of this thesis addresses this need through user studies. In addition, a framework and formalism that represents and reasons within such differences and requirements, needs to be developed. The second and third research questions address this need through the integration of the user studies with literature review.

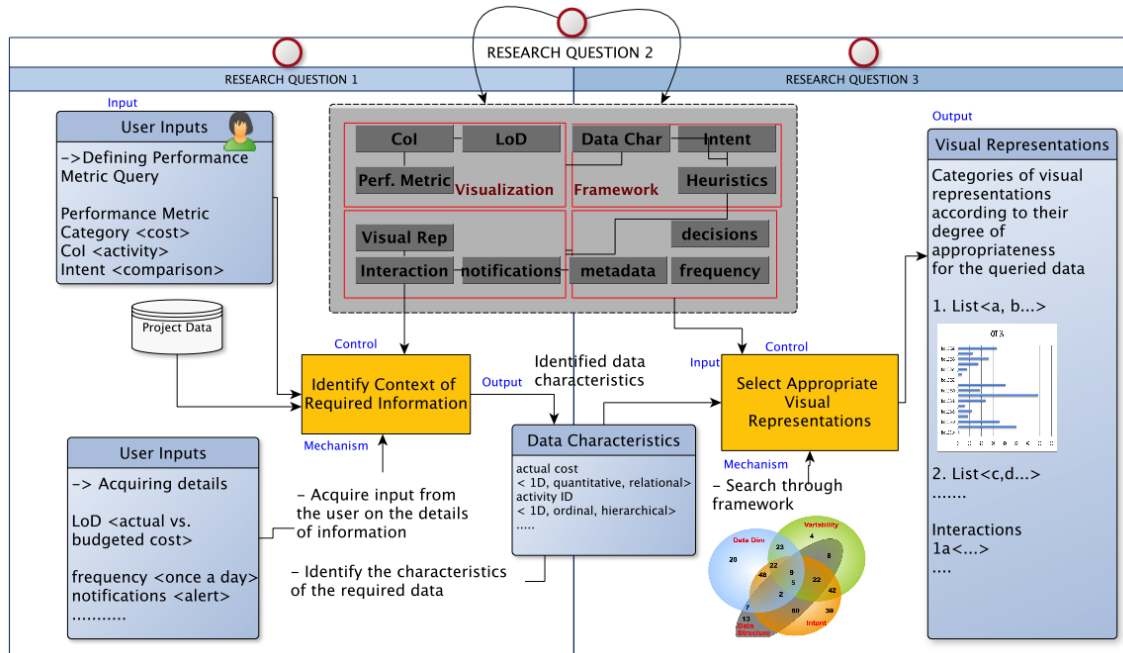


Figure 1.2: The vision on enabling project members to effectively visualize project performance information

The envisioned approach first acquires input from the user regarding the context of the required information to generate performance metrics. Here, the context of the information is defined as the performance metric category, concept of interest and the intent of the user as shown in Figure 1.2. The vision diagram is populated with an example metric query of “comparison of actual vs. budgeted cost of activities”. Given this metric query the initial input acquired from the user are cost, activity and comparison. Then, navigating through project data and user’s query, details of the metric query are identified. Given the same example, the level of detail of the information is actual vs. budgeted cost of activities that are performed last week, to be visualized once a day with alerts on the highest cost variations. Indeed, these detailed information items are identified and represented according the developed framework. As discussed earlier, the

framework is composed of components that are used to define and enable performance visualization needs of project members, and these components include the performance metric category (cost), intent (comparison), CoI (activity), frequency (once a day) and so on.

Once the relevant data is determined using the developed framework, its data characteristics are identified. Data characteristics capture information about the dimensionality of the data (e.g., 2D, multiD etc.), structure of the data (temporal, hierarchical and etc.) and variability of the data (e.g., categorical, quantitative and etc.). These data characteristics in addition to the user intent (comparison, trend and etc.) are used for the identification of effective visual representations that are captured within the framework. Indeed, the framework includes *heuristics* that synthesize and represent the information visualization related studies performed to help in identifying which visual representations should be used for which datasets and user intents. Searching through the framework to identify the effective visual representations of results is a list of visual representations that are categorized based on their degree of appropriateness for the given dataset.

1.5 Objective, Scope and Assumptions of the Research

The main goal of this research is to support the generation of effective visualization of individually relevant project performance information to project team members. The specific objectives are to:

- Identify the major commonalities and differences in information visualization requirements of construction project team members when monitoring the progress of construction projects.
- Develop a project performance visualization framework and identify its major components to support the generation of customized information visualization needs of project team members.
- Develop a formalism that utilizes the visualization framework and enables project members to effectively visualize and interact with project productivity information.

This research is focused on the performance measurements that are relevant to the construction phase of projects. The scope of the proposed research is also limited to quantitative performance metrics. Hence, the qualitative metrics that reflect the perspectives of people on the status of work or other project members are out of scope.

The target user group of this research is the project members, who are working on a single project at a given time, such as project managers and engineers. As this is an initial effort in this area, it can be seen as a step towards addressing multiple projects related performance measurement in the future.

The scope of the physical environment is the office space. In other words, desktop computers are the target technology using which the project team members will generate and visualize performance metrics to monitor the construction projects.

The scope of the research is addressing the information visualization and interaction needs of project members, and does not focus on addressing how comprehensive and/or generic the project data needs to be represented and stored since such issues are large-scale research problems within themselves and are beyond the scope of this research, which target enabling effective visualization of data to support construction project team members' decisions. Hence, it is assumed that the project data required to generate performance metrics is available, and a project data metadata that states the required information will be developed.

1.6 Research Questions

This research addresses three research questions in order to facilitate project members to visualize and interact with project performance data in an effective way according to their desires:

Research Question 1: *What commonalities and differences exist in information visualization requirements of different project members when monitoring and assessing construction project performance?*

As discussed in the introduction section, several differences exist in the way project members would prefer to monitor the performance of construction projects. Without customizing information and supporting different ways of interacting with such customized representations, information overload is probable. In order to eliminate the consequences of information overload; project members need to generate and visualize the project performance information in an effective way. However, the ways in which the

information visualization requirements of project team members for project performance monitoring vary are not known.

To enable project members to visualize the relevant information, main commonalities and differences in the ways they monitor the project performance should be investigated. Previous research efforts focused on identification of the performance metric categories with respect to different project members (e.g., Sanvido et al. 1992), the differences in the LoDs of the information, such as project or activity level information (e.g., Navon and Sacks 2007), and the differences in the users' interests for visual representations. (e.g., Songer et al. 2004). Although such categories of differences are identified, they are not sufficient to capture the differences identified through interviews. For example, types of intents, and the context of the information should also be considered.

This research question targets further exploration and identification of such commonalities and differences in the information visualization needs of project members through interviews and case studies. The testing and validation of the commonalities and differences are performed with the use of clustering methods such as Principal Component Analysis and similarity analysis.

Research Question 2: *What is a generic framework that is required to define, and represent the information visualization requirements of project members to monitor the performance of construction projects?*

Several research efforts stated that construction industry is in need of a systematic approach and representation for project performance information (Songer and Hays 2003; Navon and Sacks 2007; Golparvar-Fard et al. 2009). Several studies have intended to address this need by using different approaches, such as using augmented photographs (e.g., Golparvar-Fard et al. 2009), integrating data acquisition and simulation technologies (e.g., Azimi et al. 2010 for steel fabrication), and identifying some information items that need to be captured (Navon and Sacks 2007). However, these studies are still limited in terms of enabling navigation through different LoDs and context of information, and different visual means to effectively represent them. Hence, there is a need to develop a project data metadata that formalizes the characteristics and relationships of different LoDs of information.

Within the HCI domain, several studies have defined frameworks for information visualization (e.g., Aigner et al. 2007; Shrinivasan and van Wijk 2008; Icke and Sklar 2009). However, these frameworks are not comprehensive and integrated into a unified framework. In numerous cases, they also lack formalized definitions, detailed enumerations or mapping of components that make up the framework. Finally, they cannot be directly applied to construction domain; since they do not cover all of the natural idioms used in this domain (e.g., Gantt Chart, Building Information Models), and characteristics of data (e.g., point cloud data, different abstraction and details of information) used during construction projects. In order to provide appropriate visual representations to the users, not only the characteristics of the data, and the intent of the user, but also the effect of visualization on human cognition (e.g., color, shape,

dimension selections) should be explored. In the HCI domain, several research efforts studied these topics. For instance, while Lohse et al. (1994) identified 11 categories of visual representations, Shneiderman (1996) identified 6 types of interaction tasks and both Chengzhi et al. (2003) and Keim (2002) defined taxonomies for the association of visual representations with users' intents and interaction tasks. More detailed discussions on the instances of components can be found in Chapter 3.

Although there are numerous visual representations that have been explored in previous research studies, these studies generally have specific focuses, such as certain data types, and a set of interaction tasks. Hence, a study, which integrates prior studies and formalizes a framework for visualization of information that can be applied to construction industry, is missing. Moreover, the number of visual representations that are being studied and developed is increasing over the last few years, and hence there is a need for development of an extensible framework to incorporate visual representations to be developed in the future. With the purpose of addressing this need, a grammar / metadata of visual representations is required.

In summary, the aim of the second research question is to identify a general project performance visualization framework that supports project members to visualize the performance monitoring metrics that they are interested in. The validation of the framework in terms of its coverage and usefulness is performed through a combination of literature review, synthetic tests of generating visual representations for the metric

queries identified through interviews, and a qualitative survey to learn the perspectives of project members for the components of the framework.

Research Question 3: *What representation and reasoning mechanisms support selection and visualization of individually relevant information for effectively monitoring the productivity performance of construction projects?*

This research question aims at investigating what mechanisms can enable customized generation and visualization of performance metrics by reasoning about the framework to be generated. While it is widely recognized that the selection of visual representations is dependent on the characteristics of data and intent of the user, there are no studies that formalize the selection and configuration of visual representations within the construction domain. Within the HCI domain, a limited number of formalisms exist (e.g., Elmqvist and Fekete 2010; Helfman and Goldberg 2007); however, the implemented approaches have limitations, such as considering only data types (and ignoring the effect of intents), focusing only on a few visual representations, and not always considering the effectiveness of the visual representations. Although the number of formalisms is limited, numerous research efforts explored the effectiveness and/or appropriateness of specific visual representations, and these studies can be synthesized to generate a formalism to identify effective visual representations. However, findings of some of these studies are semi-structured, ill-defined and sometimes contradicting. Consequently, in order to use the existing knowledge to find the applicable and effective visual representations, these heuristics need to be synthesized, and formally represented. This research question

targets addressing this need. The developed formalism is validated in terms of its precision to retrieve applicable visual representations, usability in construction management domain, and usefulness to suggest effective visual representations through synthetic tests and user survey.

1.7 Research Method

The research method involved identification of commonalities and differences in performance visualization needs of project members by following contextual design guidelines and performing semi-structured interviews with a set of project members from different construction projects. I have also performed direct observation and exploration of existing performance monitoring processes during two-months internship. During the internship, I have started to identify the different needs and challenges of project performance tracking on a highway construction jobsite. Following this initial investigation, I performed two additional case studies. First case study is the interviews that I performed with fourteen project members to explore their information visualization needs, and as a second case study I tested the usability of information visualization toolkits.

In order to implement the approach for facilitating the effective visualization of individually relevant project performance information, the following tasks are performed as indicated within each bullet:

1. Identifying the commonalities and differences in information visualization requirements of project members for monitoring and assessing the performance of construction projects:

- a. Direct observations of two project engineers and a project manager for monitoring the performance of a construction project.
 - b. Semi-structured interviews with fourteen different project members working on seven different projects as a part of a contextual design process.
 - c. Exploration of the existing performance monitoring metrics identified in prior research studies.
 - d. Generating individual and consolidated work models to understand and develop a common language regarding the interview data.
 - e. Identifying the commonalities and differences in information visualization needs of project members.
2. Formalizing a generic framework for construction project performance visualization:
- a. Performing a detailed literature survey to identify the research studies that are related with information visualization.
 - b. Comparative analysis of findings with literature review and interviews.
 - c. Developing a project performance visualization framework.
 - d. Identifying the enumerations of the framework components for performance visualization through integration of interviews and literature.
 - e. Validating the coverage in terms of the demographics of the interviewees and literature, and usefulness of the framework in terms of the value of each framework component.

3. Representation of and reasoning with information visualization related heuristics
 - a. Performing a case study to evaluate information visualization toolkits.
 - b. Developing a representation schema to represent the information visualization related studies performed in prior research.
 - c. Developing an approach to formalize the selection of effective visual representations.
 - d. Developing a project database using data from internships to validate the developed representation schema and reasoning mechanisms.
 - e. Performing user tests and synthetic tests to verify and validate research questions 2 and 3.

1.8 Readers' Guide to the Dissertation

The introduction chapter of this dissertation provides a description of the problem, current practice and existing challenges, the research vision, objectives, scope, assumptions, and questions, and an overview of the research methods.

Chapter 2 describes the need, motivating case studies, research tasks, and findings in relation to identification of the commonalities and differences in information visualization needs of construction project members. It describes in detail the case studies performed as a part of contextual design for information identification, consolidation and results based on this investigation.

Chapter 3 describes the lessons learned from motivating case studies on visualization and interaction requirements and functionality of visualization environments serving that purpose, research tasks, and findings in relation to the project performance visualization framework and its components.

Chapter 4 describes the need, motivating cases, research tasks, findings and validation results in relation to consolidation, representation and reasoning of information visualization related *heuristics*.

Chapter 5 discusses of the results of the research by describing the findings, contributions, and expected practical implications. It also describes possible future work that can be performed to complement this research.

CHAPTER 2

2 Information Visualization Requirements of Project Members for Construction Project Performance Monitoring

In the Architectural, Engineering and Construction (AEC) industry, researchers agree on the fact that for successful completion of a project, accurate and relevant information should be available at the right place, at the right time, and with the right format (Abou-Zeid et al. 1995; de la Garza and Howitt 1998; Ganah et al. 2010; Hiremath and Skibniewski 2004; Soibelman and Caldas 2000). This is in particular true for monitoring and assessing the performance of a project, which is claimed to be one of the most important tasks of construction project management (Memon et al. 2007). When managing construction projects, project team members prefer easily accessible, customizable and individually relevant just-in-time information (Cadiz et al. 2002; Balfanz et al. 2003; Russell et al. 2009); however, such needs are not fully supported. For instance, Lee and Rojas (2009) emphasized the ineffectiveness of several traditional visual representations (e.g., CPM, PERT) for monitoring the performance of projects as opposed to their use and effectiveness for planning purposes. Specific to superintendents, Reinhardt et al. (2005) emphasized the ineffective and inefficient navigation functionality of mobile computing devices for contextual and visual representation of project

information. Furthermore, Navon and Sacks (2007) identified the need for exploring different project members' information needs with varying levels of detail for project control, as a research issue to be pursued.

Construction project team members need individually relevant information (Sanvido 1992; Songer and Hays 2003; Hiremath and Skibniewski 2004) in order to gain awareness about project performance. Information required to monitor the performance of projects varies amongst the project team members based on their responsibilities, objectives and experience levels. In order to support project members in accessing critical information, commercially available project management systems have started to implement dashboard views within which the most critical information is summarized on a single screen. Even though such dashboards present the prominent information in a graphical way, only a limited number of visual representations and interactions with the data are offered to the users, and these do not always align well with the different needs of the project team members (Song et al. 2005). In other words, the users cannot flexibly explore the project information from different perspectives, and the visual means for presenting and interacting with information do not always match with the intent of the user and the characteristics of the data.

In the absence of flexibility in visualization of and interaction with data, information overload is probable. Indeed, numerous existing studies highlight the existence of information overload in construction projects (e.g. Leung et al. 2003; Haksever 2000; Thorpe and Mead 2001; Thorpe and Murray 1996; Zantout and Marir 1999). Such

information overload decreases the quality of the decisions (Speier et al. 1999; Abdel-Khalik 1973; Chewning and Harrell 1990; Shields 1980; Snowball 1980; Mehrotra et al. 2004), and causes waste of time and money (Keim et al. 2008). This would contribute to project members wasting unnecessary time in accessing and visualizing the project information. McCullouch (1997) stated that on the average, 30-50% of field supervisory personnel's time is spent on recording and analyzing field data (McCullouch 1997). Moreover, during my internship at a construction job site, I observed that one of the project engineers spent 10 % of his time on a daily basis to aggregate and visualize the timecard information in a spreadsheet view, a view that he preferred to visualize and analyze such daily production data. In order to overcome information overload, project members should be provided with project performance information that is unique to their needs.

The information required to monitor the performance of projects varies amongst the project team members based on their responsibilities, objectives and experience levels. A main challenge associated with not being able to enable the visualization of individually relevant information stem from the lack of foundational knowledge describing the commonalities and differences in information needs of project members. Hence, the objective of this study is to investigate the performance monitoring and visualization needs of project members in order to foster the performance monitoring process of construction projects.

2.1 Motivating studies depicting the need for identifying information visualization requirements of project members

I performed a set of initial interviews with different project members working on a range of construction projects to understand whether they have access to individually relevant and customizable project information as they desire. In this section, I will focus on five of these interviewees to illustrate the motivation for the proposed research. These five project members, with whom I interviewed, are working within the same company and on highway projects. One of the performance indicators that they are all interested in is weekly cost performance, which they monitor using the reports generated from an internal software system. However, the information items that they utilize from these reports, how they would like to visualize them and why they would like to visualize them in a particular way differ amongst different project team members. These differences are listed in Table 2.1, and grouped under their intent in visualizing and interacting with the data, the usage of purpose of the information, the criterion to evaluate the information, the LoD of the information, the requested additional contextual information, and the visual representations to display the information. For instance, the first row in Table 2.1 depicts the information visualization needs of an Area Manager, who manages more than one project at a given time. With the intent of identifying the projects with the highest cost overrun, his purpose was to check if the projects are on track. The LoD that he was concerned about was at the project level, and he set the criteria to see the projects if the total cost difference is over -\$5,000 (in other words, when there is a cost overrun of more than \$5,000). At this level, he is only interested in getting notified about the project(s) that has a cost overrun above the given threshold; so, the visual representation that he desires is a flag with the project's name. Similarly, the information visualization needs of

a project manager are described in the second row of Table 2.1. With the intent of projecting the cost of activities performed in a given week, his purpose is to compare whether the observed performance matches the estimated performance. He is interested in activity and resource level cost, resulting in him asking for material costs and projected activity cost for all of the activities. Finally, without any criteria to eliminate or highlight activities, he prefers to visualize a list to display the information.

Table 2.1: Differences in five project members' interests in monitoring the weekly cost performance of activities

Who	Interaction Intent	Decisions to be made	Level of Detail (LoD)	Criteria for filtering/ highlighting information	Additional Information	Visual Representation
Area Manager	Identify highest cost overrun projects	Check if the project is on track	Project level information	See project if total cost difference > -\$5000	Not requested	Flag
Project Manager	Project cost of activities performed that week	Compare whether the observed performance matches the estimated performance	Activity and resource level information	Not requested	- Material costs - Projected activity cost based on actual & estimated productivity	List
Project Eng 1	Compare all activities performed that week	Find out major discrepancies	Activity level information	Total Actual versus estimated cost of activities	Not requested	Scatter Graph
Project Eng 2	- Identify the foreman of overthreshold activities - Identify their past pattern	Report to higher level members Report to foremen Track critical activities	Activity and foremen level information	- Cost overrun10% - Top 5 activities with the highest cost overrun	Foremen responsible from the activities performed that week	List
Project Eng 3	Identify activities with higher cost-overrun	Identify the causes to fix them	Activity, resource level information	- Cost overrun10% - "Activity % complete" less than 75% in terms of quantity	Quantities completed to-date	List

These major differences are also graphically depicted in Figure 2.1 as a sequence of actions including processing of data, selection of visual representations, and decision-making based on visual representations analyzed. As shown in Figure 2.1, when a cost report is generated, it is not in a format that is requested by the project members and hence further processing of the data contained in the report is needed. The initial three layers in Figure 2.1 are interaction intent, criteria, and additional information. These three layers indicate the differences and similarities in the expectations of project members about how to monitor the cost performance. The final two layers are visual representations and decisions of the project members. These two layers reflect the similarities and differences in the use of the identified information to visualize and assess the required information.

Table 2.1 and Figure 2.1 highlight that it is possible to declare the information visualization needs of the project members in a common way by using concepts related with the context of the information, such as the Level of Detail, and the visualization and interaction needs of the project members, such as visual representations, and criteria for filtering information. However, the details in which each of these concepts are used to depict each project team member's visualization needs, differ.

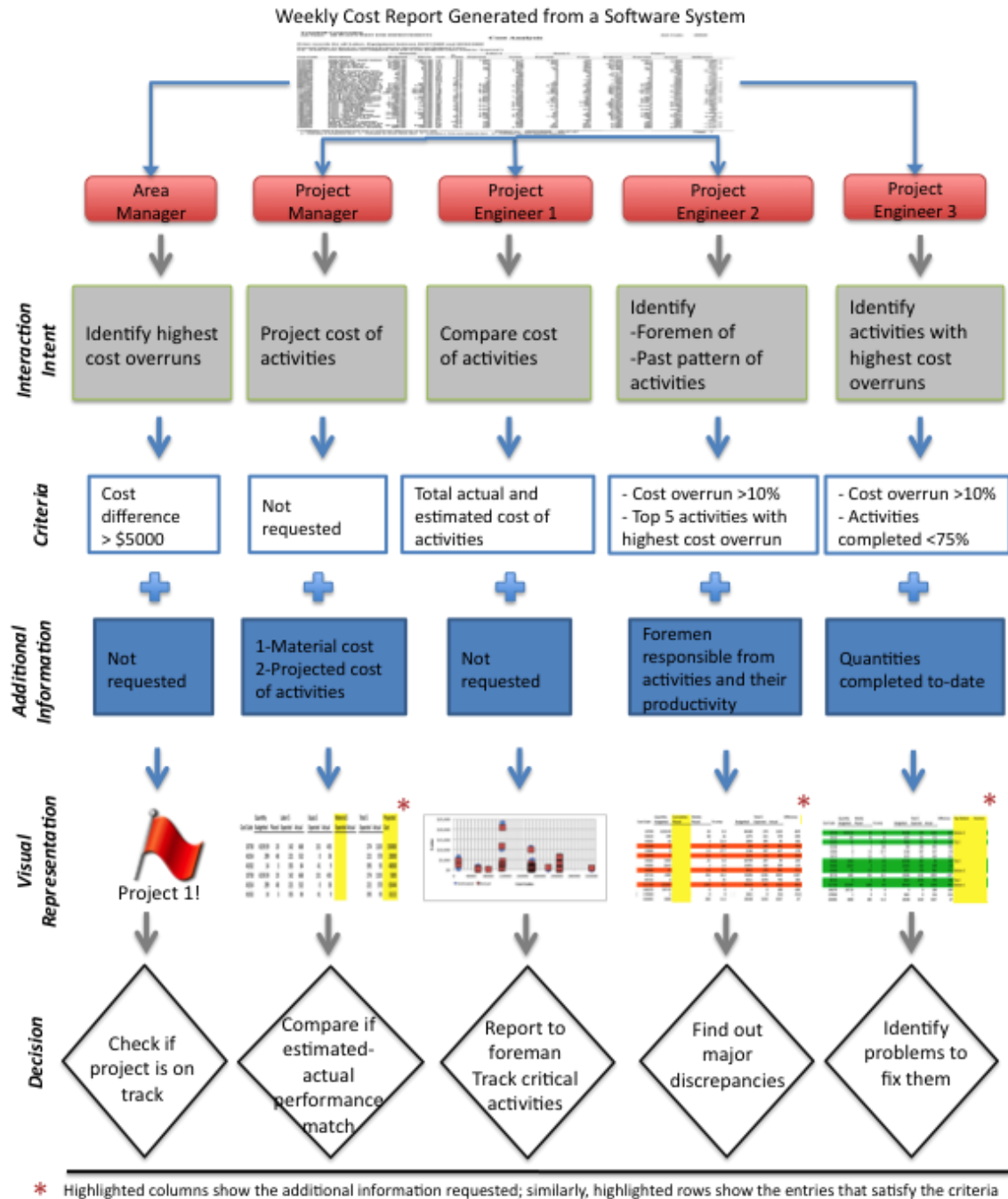


Figure 2.1: The differences in the way project members want to track the Weekly Cost Reports and the associated purposes of usage divided according to the responsibilities of the project members

As stated, although the presented project members are all interested in monitoring the comparison of the actual and budgeted cost of activities, differences exist in selected visual representations, criteria on assigning notifications or filtering information, and the usage purpose and interaction intent of different project members. By taking into account these differences in needs presented for a set of a project members and performance metrics and by looking into 14 interviews with project members involved in 7 different projects, I observed that for the 60% of the performance metrics; project data needs to be processed to visualize it in the required form. These information visualization needs that are not met through the existing practices of project members highlight the need to foster the understanding of project members' unique performance monitoring and visualization needs.

2.2 Research Objectives and Scope

The objectives of the research presented in this chapter are: (a) To investigate the information requirements of construction project team members when monitoring the progress of construction projects, and (b) To identify important commonalities and differences in these information visualization needs.

I specifically focused on the performance measurements that are relevant to the construction phase of projects and limited the performance metrics to quantitative ones. Since qualitative metrics that reflect the perspectives of people on the status of work or other project members are subjective and cannot be generalized, I left them out of scope of this research. In addition, the target user groups of this research project have been identified as project members, who are working on a single project at a given time, such

as project managers and engineers. As this is an initial effort in this area, it can be seen as a step towards addressing multiple projects related performance measurement in the future. The scope of the physical environment that project personnel will interact with project performance metrics has been defined as their office spaces and their personal computers that they will use to generate and visualize performance metrics to monitor the construction projects.

2.3 Background Research

The research presented in this chapter builds on previous research studies done in the areas of project performance measurement metrics, differences in performance monitoring needs of project members and construction project performance visualization.

2.3.1 Performance Monitoring Frameworks and Approaches for Construction Projects

I have reviewed over 70 papers published over the last 20 years in relation to performance metrics for construction projects. These research studies can be grouped under two main categories: (a) project performance being evaluated at project completion, and (b) project performance being measured throughout the project. The first group of research studies focused on defining performance measurement with the respect to project success at the completion, and considered the satisfaction level of the clients (e.g., de Wit 1998; Chan and Chan 2004; Yang and Peng 2008). The second group of studies considered the need for monitoring a project's performance during the construction phase with a focus on either at Project-level (e.g., Alarcon and Ashley 1996;

Korde et al. 2005) or at Company-level (e.g., Takim and Akintoye 2002; Cheah et al. 2004) performance. While the project-level performance measurement mainly considers processes and operations performed on site, company-level performance measurement targets strategic management and financial aspects.

Several of the studies, which highlighted the need for monitoring project performance during construction, suggested the use of or development of performance measurement framework for construction (e.g., Alarcon and Ashley 1996; Cheung et al. 2004). Three major frameworks developed prior studies are Key Performance Indicators, Balance Scorecards, and European Model of Quality Management Excellence (Yang et al. 2010). Key Performance Indicators (KPIs) are the indicators that enable measurement of a project's and organization's performance throughout the construction industry (DETR 2000). This approach has been proposed and evaluated by researchers to identify the key characteristics of indicators and assess their application in the industry (e.g., Cox et al. 2003; Beatham et al. 2004; Bassioni et al. 2004; Chan and Chan 2004). Balanced Scorecard (BSC) is a performance measurement system that incorporates four main measurement categories; (1) financial performance, (2) internal business process, (3) customer satisfaction, and (3) learning and growth. While this method is used across different domains, numerous researchers studied identification of potential sub-measures for each of the four categories and their application in construction industry (e.g., Kagioglu et al. 2001; Stewart and Mohamed 2001; Mohamed 2003; Yu et al. 2007). European Foundation of Quality Management (EFQM) Excellence Model is designed to define and assess continuous improvement of organizations based on a set of predefined

criteria (Beatham et al. 2004). Major research performed in this area is towards studying the implementation of EFQM within the construction industry (e.g., Robinson et al 2001, Watson 2002; Robinson et al. 2005). However, Bassioni et al. (2004) and Lin and Shen (2007) highlighted the fact having predefined fixed criteria can be non-comprehensive and limit the flexibility of application to different processes.

A common feature of these frameworks is that they incorporate some performance metrics against which performance is measured and evaluated. However, the identified performance metrics differ from each other as the scope of the research studies change in the following aspects:

- (a) *A specific performance metric category being studied*, such as schedule (e.g., Kog et al. 1999), safety (e.g., Abdul-Rashid et al. 2007), environment (e.g., Shen et al. 2005), subcontractor performance (e.g., Ko et al. 2007) and project manager performance (e.g., Dainty et al. 2003);
- (b) *The country in which the projects are evaluated*, such as in UK (e.g., Mahmoud and Scott 2002; Construction Consultants KPIs Working Group 2007), in Botswana (e.g., Swarnadhipathi and Boyd 2007), and in South Asia (e.g., Sohail et al. 2002);
- (c) *Techniques or technologies being applied in projects*, such as 4D (e.g., Dawood et al. 2006) and lean construction (e.g., Ballard and Howell 1994);
- (d) *Type of projects being evaluated*, such as urban micro-contracts (e.g., Sohail et al. 2002) and design-build fast-track projects (e.g., Pena-Mora and Li 2001);
- (e) *Level of decision making being studied*, such as at organizational level (e.g., Mahmoud and Scott 2002), at project level (e.g., Cheung et al. 2004; Alarcon and

Ashley 1996; Construction Institute 2005), at activity level (e.g., Li and Russell 2006; Korde and Russell 2005) or at task level (e.g., Botero and Alvarez 2005; Ballard and Howell 1994);

(f) *Metric being studied from different stakeholders' point of views*, such as from clients (e.g., Cheng et al. 2005), from project managers (e.g., Cox et al. 2003) and from consultants point of views (e.g., KPIs Working Group 2007).

Currently, 300 performance metrics have been identified to measure construction project performance. The detailed lists of performance metrics grouped under categories and subcategories are presented in the tables provided in Appendix A. Since there is no widely accepted definition of categories that I can build on in grouping this large number of metrics, I synthesized these metrics under eight overarching categories based on their purpose of usage:

1. *Time*, which incorporates project performance metrics that evaluate a project/activity/task from completion, schedule and time perspectives.
2. *Cost*, which incorporates project performance metrics that evaluate a project/activity/task from cost and budget perspectives.
3. *Quality*, which incorporates project performance metrics that evaluate a project/activity/task from quality, conformances and reworks perspectives.
4. *Communication*, which incorporate project performance metrics that evaluate the number and types of exchanged documents, such as RFIs, change orders, and submittals, for a project/activity/task.
5. *Safety*, which incorporates project performance metrics that evaluate a project/activity/task from safety requirements perspective.

6. *Productivity*, which incorporates project performance metrics that target the quantities of performed work and production levels for a given project/activity/task
7. *Business performance*, which incorporates performance metrics that are related to higher company level metrics such as return on investment and profitability.
8. *Qualitative metrics*, incorporating performance metrics reflecting the subjective views of the project participants about different aspects of the projects, and are outside the scope of this research.

Table 2.2 presents an overview of the synthesized performance metric categories and example metrics that are identified in prior work. When prior work on definitions and indexes of performance metrics are studied, it is acknowledged that the vast number of performance metrics is a powerful source to lead project members in the direction of identifying the factors to measure the performance of construction projects. Albeit the identified number and variety of performance metrics, existing research studies are limited in providing only the definitions rather than the scope of the performance metrics. In other words, the literature on performance metrics do not explicitly define the extent to which individual project members measure performance including the captured level of detail of the metric queries and how and when to visualize and interact with these metrics.

Table 2.2: Performance metric categories and subcategories identified in literature

Perf. Metric Category	Description	Examples of Performance Metric Subcategories from
Time	Project performance metrics that evaluate a project/activity/task from completion, schedule and time perspectives	Achievement of critical dates, Turn around time for submissions in period, Total Project Duration, Effectiveness, Time predictability, Rectifying defects
Cost	Project performance metrics that evaluate a project/activity/task from cost and budget perspectives.	Interim payment, Variation order, Cost claims, Final Account Forecast, Cost growth factors, Cost for construction
Communication	Project performance metrics that evaluate the number and types of exchanged documents such as RFIs, and submittals.	Communication management, Communication index, Change Cost Factor, Change Orders.
Quality	Project performance metrics that evaluate a project/activity/task from quality, conformance and reworks perspectives.	Inspections, Non-compliance records Work Rejection, Survey (Samples) Rejections, Defects, Outstanding quality issues, Quality control/ rework.
Safety	Project performance metrics that evaluate a project/activity/task from safety requirements perspective.	Statistics, Monitoring and Compliance Education training and Campaigns, Inspections and Audits, Complaints and prosecutions, Total Recordable Incident Rate (TRIR), Dart Rate (LWCIR).
Productivity	Project performance metrics that look at the quantities performed work and production levels for a given project/activity/task.	Productivity index 1, Productivity index 2, Units/MH, \$/unit, Weekly Plan Execution, Look-ahead Planning, SPI, CPI.
Business Performance	Performance metrics that are related to higher company level metrics, such as return on investment and profitability.	Profitability, Productivity Return on Capital employed, Return on value added, Interest cover, Return on investment, Ratio of value added, Repeat business, Outstanding money.
Qualitative Metrics	Performance metrics reflecting the subjective views of the project participants about different aspects of the projects, and are outside the scope of this research	Subcontractor Performance Index, Team Building, Leadership Mutuality & approachability, Honesty & integrity Communication Learning, and Understanding

2.3.2 Differences in Performance Metric Needs of Project Members

Due to different thinking styles, experiences and knowledge of project members, how they make use of project information for managing projects is not predictable (Chiu and Russell 2011). Various researchers in the construction project management domain stated that project members have unique information needs for monitoring the performance of a

project (e.g., Bryde 2003, Sanvido et al. 1992, Songer and Hays 2003). Although this need has been clear, the background research for enabling project personnel specific monitoring of performance is limited. In general, related previous research studies mainly focused on defining performance metrics. Only a limited number of research studies focused on the specific information and visualization needs of project personnel. Table 2.3 shows a synthesis of such research studies and some examples from previous research studies that demonstrate some differences in information needs of project members.

Most of the research studies on differences in performance monitoring needs focus on the performance metric categories as discussed in the previous section. For instance, Golparvar Fard et al. (2007) indicated that the level of detail and performance metric categories dependent on the job title of project members. While upper level management would be interested in cost of activities in summary levels, the daily practitioners are interested in detailed schedule of activities.

The other differences are categorized as frequency, the visual representations, and the LoD of the information as shown in Table 2.3. Although the frequency of performance monitoring is identified as a differentiating factor; its effect on the context of the visualized information is only studied by Navon and Sacks (2007). Finally, a number of research efforts studied the need for the use of different LoDs and visual representations for construction projects; however, some of the identified differences are at an abstract level and some of the differences are identified specific to a certain scope. Moreover, the results from the initial set of interviews demonstrate the existence of further differences;

such as decisions to be made, and interaction intents as detailed in the motivating case section.

Table 2.3: Differences in information needs of project members identified from construction monitoring related research efforts

Information Needs of Project Members	Studied Project Members	Examples of the Differences	Reference
Performance Metric Category	Multiple project members	e.g., Cost, schedule, quality	e.g., Sanvido 1992; Korde et al. 2005; Kog et al. 1999, Golparvard Fard et al. 2007
Frequency	Multiple project members from the main office office, project management and external	e.g., annual, monthly, daily	e.g., Navon and Sacks 2007
Visual Representations	Project manager, accountant, Subcontractors	e.g., users should select representations (e.g., tree maps, linear trees) based on their cognitive abilities e.g., artifacts used during collaboration meetings	e.g., Songer et al. 2004; Russell et al. 2009, Tory and Staub-French 2008
Level of Detail (LoD)	Superintendents, Multiple project members	e.g., space decomposition: building, floor, section, element e.g., time decomposition: yearly, monthly, daily e.g., project related objects/ facets of information items: facilities, work stations, products	e.g., Reinhardt et al. 2005; Navon and Sacks 2007; Jung and Woo 2004, Golparvard Fard et al. 2007

2.3.3 Visualization of Project Performance in Construction Management Domain

Generally, the term “visualization” is associated with visualization of either 3D/4D product models or pictures and videos of projects in construction industry. For instance, Tory and Staub-French (2008) evaluated the use and value of different visual representations including 3D models, charts, texts, and boards during collaboration meetings. Aside from collaboration, the focus of this research study on the topic of information visualization for performance monitoring has been studied by many

researchers. Example studies in this area include visualization of project performance data over product models (e.g., Songer et al. 2000, Srivastav et al. 2009), as 4D models (e.g., Golparvar-Fard et al. 2009, Rohn et al. 2010), with animations over process models (e.g., Bohn and Teizer 2010), and over dashboard environments (e.g., Wu and Hsieh 2012).

In relation to the research studies that proposed integrating performance metrics with product models, Songer et al. (2004) proposed integration of 3D CAD models with workflow management concept to dynamically monitor the processes, and get visual notifications for design related problems to have better understanding of these problems. Similarly, Srivastav et al. (2009) proposed the use of information layering, color coding on 3D drawings, and use of charts for the representation of the site analysis data for the designers. In relation to integration of performance metrics with 4D models, Golparvar-Fard et al. (2009) superimposed daily progress photograph logs on the 4D as-planned models in order to compare the construction progress with the metaphor of traffic light. Similarly, Rohn et al. (2010) integrated BIM with as-is photographs to provide progress feedback. To support integrated process models with certain animation capabilities for performance monitoring, Bohn and Teizer (2010) integrated project schedule and BIMs with automated cameras, Huhnt & Richter, (2010) used color coding for activities based on trades, components, and structure information, and in relation to construction dashboard visualization; in order to visualize the present stage of ongoing projects, Wu & Hsieh (2012) proposed a tool to visualize 3D ,4D, and nD visualization for ease of communication among the project participants, and similarly Kuo et al. (2011) proposed

a framework to visualize the 3D Model with temporal (schedule), hierarchical (organizational structure), relational (adjacent building components), and spatial information.

Although most of the previously discussed studies facilitate the visualization of construction projects, the main focus is the representation of the building components associated with pictures or schedule. However, construction performance monitoring is related with information visualization (visualization of quantitative or categorical performance data) as much as related with scientific visualization (3D, 4D models). With some exceptions, the majority of the studies do not enable quantitative performance measurement through product models. In other words, for monitoring the progress of the projects, the required visualization is not always about 3D models. For instance, visualization of schedules or productivity of foremen through Gantt Charts and Scatter Graphs are not 3D in nature.

Several promising research studies started to combine the information visualization concerns of civil engineering with the principles of HCI. These studies that focus on the visualization of performance information include;

- Presenting guidelines for selecting effective visual representations: Lee and Rojas (2009) and Russell et al. (2009) studied the process of and identified a set of principles to generate effective visual representations.
- Incorporating interaction techniques: Ye and Froese (2009) mapped a set of interaction intents (e.g., comparison and clustering) and visual representations (e.g., histograms

and scatter plots), Liston et al. (2000) studied information overlaying and highlighting techniques that are applied to 2D and 3D models to foster understanding and collaboration.

- Assessing the strengths and limitations of different visual representations: Zeb et al. (2008) compared the strengths and weaknesses of a few visual representations and HCI features like zooming and brushing for as planned and as-built comparison of project schedules. Songer et al. (2004) evaluated and compared four graphics; including scatter plot, linked histogram, hierarchical tree and treemap, to represent structured cost control data.
- Performing quantitative analysis: Tory and Staub-French (2008) performed field studies to observe the interaction of project members with visual representations while planning and collaborating.

These approaches provide foundational work in identifying the information visualization needs within the construction domain; however, these studies are high-level and broad, which require them to be further elaborated. Moreover, they do not focus on the project members unique performance monitoring needs with the exception of the study done by Tory and Staub-French (2008). However, this study only focuses on collaboration meetings and no performance metrics were investigated.

2.4 Research Method

The research described in this chapter targets identification of commonalities and differences in information visualization needs of project members for monitoring the

progress of construction projects. In order to do that, I have incorporated Contextual Design process as a research method. Contextual design method is developed by Beyer and Hotzblatt (1998) in order to collect, understand, and interpret user-centered information. Fouskas et al. (2002) identified three main methods for the identification and elicitation of requirements; (1) Requirements Engineering, which involves analyzing requirements regarding implemented systems and people involved, (2) User and Task Analysis, which involves learning about ordinary users by observing their actions, and (3) Contextual Design, which involves user-centered process for learning about users, their needs and works. Currently, since a project performance visualization environment is not available, contextual design approach is selected in this study to identify the requirements.

This process helped me to redesign the desired performance monitoring structure for each of the participants through performing interviews with participants and developing work models as described in detail in the following section.

2.4.1 Contextual Design

The process begins with contextual inquiry as a technique for understanding the ways project members would like to visualize project performance. Contextual inquiry is an exploratory research to gain insights about the user needs in their context of work. Each inquiry is followed by developing paper prototypes of the participants' requests and gathering the data collected from each of the participants. Specifically, the contextual

design method provides explicit steps and deliverables with the following parts (Beyer and Holtzblatt, 1998):

- i. Contextual inquiry; gathering user data by one-on-one interviews.
- ii. Work Modeling; representing the work of each interviewed user with concrete graphical representations using five work models.
- iii. Consolidation; combining the data collected from different users together and comparing multiple users to identify possible commonalities.
- iv. Work Redesign; creating storyboards from consolidated data to understand the ways to improve the users' work.
- v. User Environment Design; structuring the system work model, and defining its parts and their inter-relationships.
- vi. Mock-up and Test with Users; testing the developed ideas with users through prototypes.

Each of these steps will be described in detail in the following subsections.

2.4.2 Pilot Study and Recruitment

The initial step of contextual design was to create and test contextual inquiry process with professors and PhD students in construction management domain. Improving and finalizing the design scenario, in this study, fourteen interviewees from six different organizations and working on six different projects were engaged. Table 2.4 summarizes their background. The organizations, with which the interviewees are associated, range from construction companies, to universities and governmental organizations. The types

of projects focus on commercial and educational buildings, highway and dam constructions.

Table 2.4: Demographics of the project members that are interviewed

Responsibility	# of Participants	Highway Projects Experience (years)	Building Projects Experience (years)
Area Manager	1	18	NA
District Manager	1	30	NA
Head of Estimation Department	1	30	NA
Owner Representative	1	NA	20
Project Manager	3	15	9, 25
Project Engineer	5	4, 8, 10	2, 8
Superintendent	2	30	1

2.4.3 Contextual Inquiry

Each contextual inquiry is composed of semi-structured interviews to learn interviewees' background, the projects that they are working on, their responsibilities, and current practice on how they access to the required information. The semi-structured interviews have been conducted in the form of a Card Game. During the interviews, I provided a blank paper, and various cards of visual representations to the interviewees. Each interviewee was asked to generate a visualization environment that they will use to monitor the performance of the projects that they are responsible from, by using a collection of these cards. Upon the selection of the cards, the specifics (e.g., fields, axis and labels) of the visual representations were written with the LoD of information asked by the interviewee. In the following sections, graphical representations of these interviews are presented.

2.4.4 Work Modeling

The next step is the generation of work models of the interviews with the participants. This process is important in order to represent the participants' inputs in visual form to identify the major components, possible breakdowns in the current process and area of development. There are five work models two of which will be introduced in this report.

1. *Flow model*; representing the communication and coordination necessary, irrespective of time. For my research, the flow models capture each interaction that an interviewee requested; such as interacting with visual representations and calling other people. Six entities are used in this model; nodes representing the responsibility of individuals, artifacts representing the requested information, actions representing the interaction tasks of users, nodes representing other project members, nodes representing the decisions of the interviewee and breakdowns representing the problems in the information flow. Figure 2.2 demonstrates an example of a flow model for a project engineer.
2. *Artifact model*; showing the physical things required to support the needs of an interviewee including documents, tools, or electronic things. This model captures the structure, usage and intent of the required information. Following the interviews, I generated the collection of visual representations, in other words the visualization environment, using MS Visio and Adobe Illustrator. An example for artifact model is presented as Figure 2.3 for a project engineer working for a building construction project.

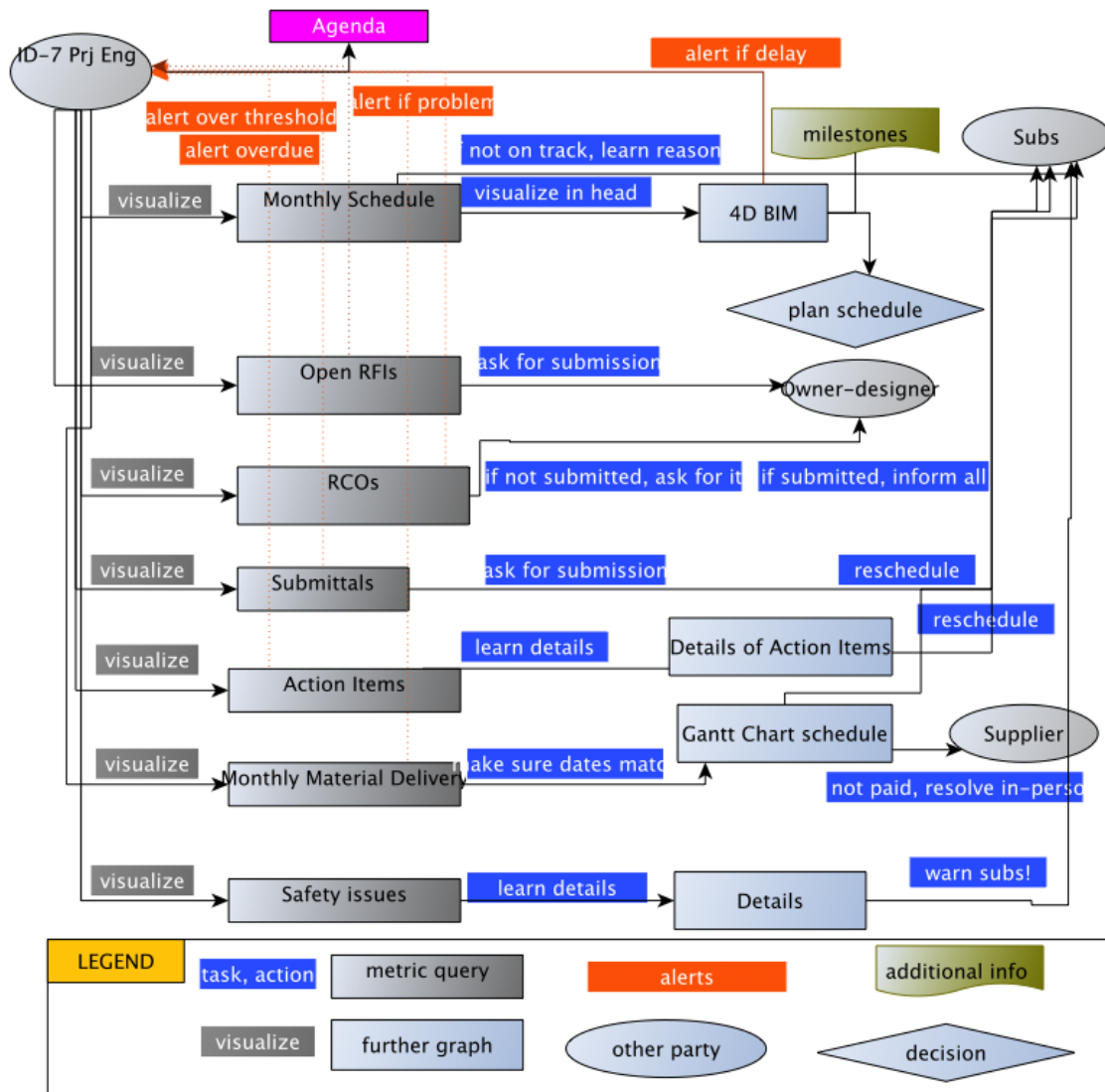


Figure 2.2: Flow Model for a project engineer (interviewee #7)

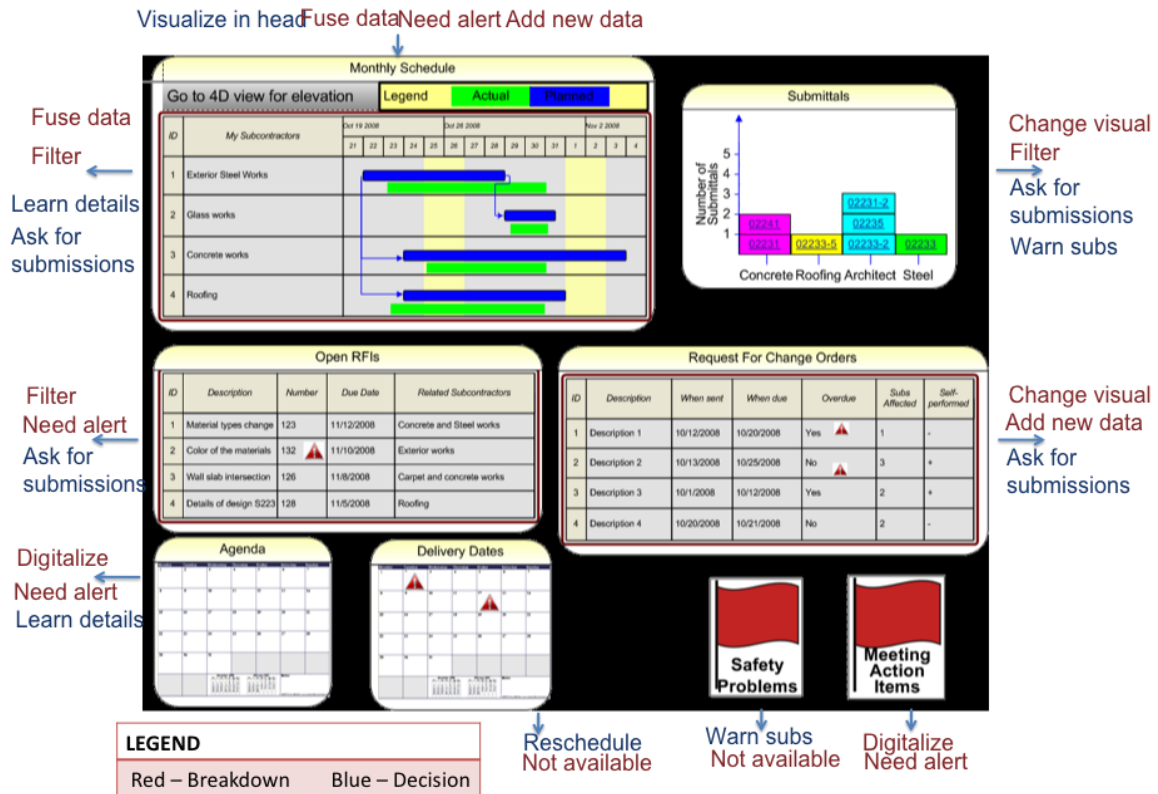


Figure 2.3: Artifact Model for a project engineer (interviewee #7)

The core of the artifact models also represents a mock-up of the designed visual environment per user needs that is referred to as “dashboard”. Two examples of these dashboards, belonging to a superintendent and owner representative are presented in Figures 2.4 and 2.5.

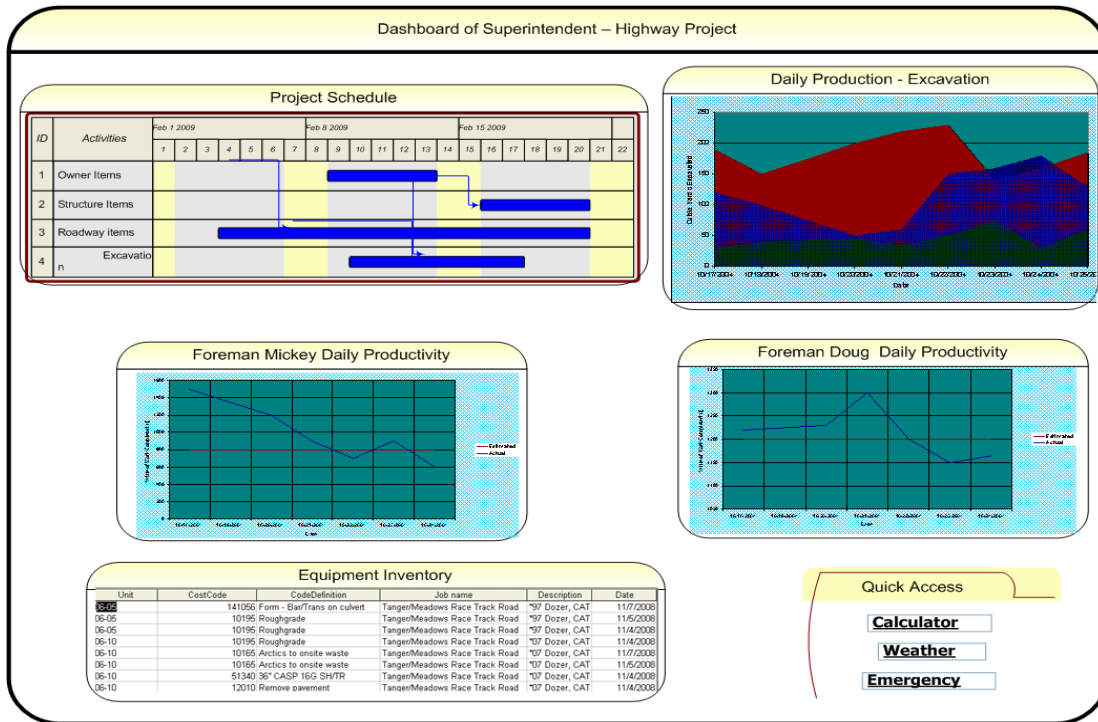


Figure 2.4: Example dashboard: Dashboard of a superintendent of a highway project

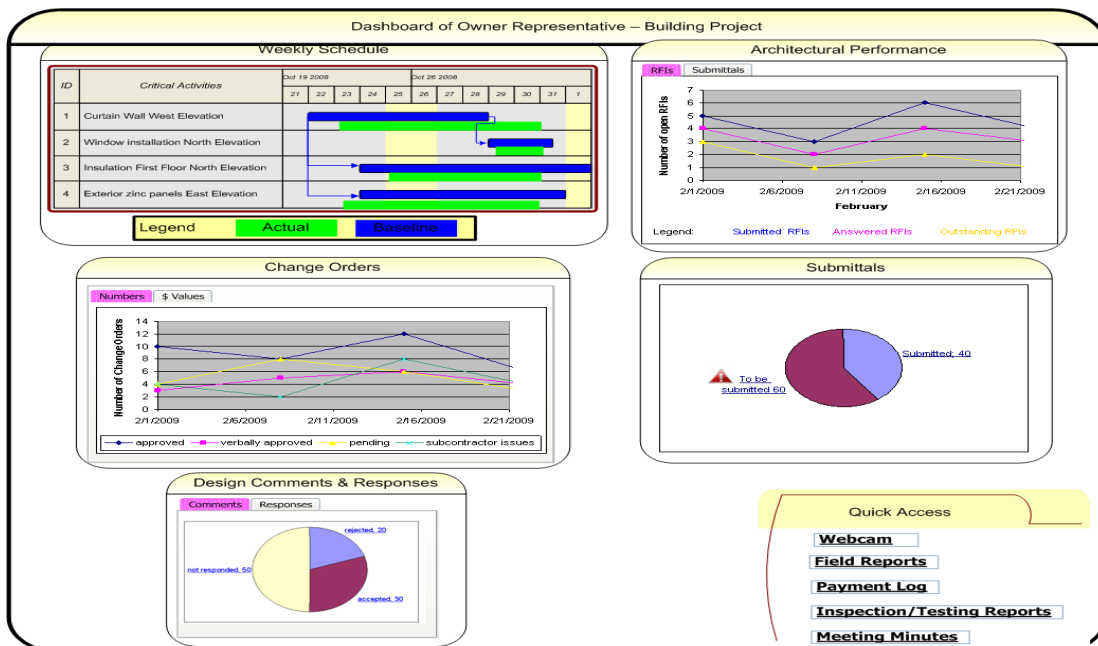


Figure 2.5: Example Dashboard: Dashboard of an owner representative of a building project

There are three more work models in contextual design approach; Sequence model (detailed work steps necessary to accomplish a certain activity), Cultural model (the norms, and policies of work environment) and Physical model (the physical environment where the work tasks are performed). Since the aim of this research is to capture the individual information visualization needs, the processes of a working system, limitations of physical environment, and cultural effects are out of the scope.

2.4.5 Consolidation

The consolidation step of contextual design includes the use of affinity diagrams. Affinity diagram is the means to organize data and ideas and brings together the data collected from different interviewee in a hierarchical diagram. My strategy for this cross-classification problem involves identification of the categories of commonalities among the information visualization needs of project members, and discriminating and noting the differences among these categories. Through multiple iterations, I have identified the commonalities and differences in these needs. Figure 2.6 demonstrates the affinity diagram for the consolidated interview results. The affinity diagram shows patterns and hierarchy across notes from 14 study participants through maintaining the detail of each participant's interview. Top-level categories that emerged in through diagramming are characteristics of project members, performance metric dimension, concept of interest, level of detail, decision usage, intent, visual representations, frequency, notifications and interactions. While each of these categories is discussed in the upcoming sections, the initial category for project member characteristics is used only to capture the information about the interviewee. Indeed, these characteristics are not used as a variable of the differences in information needs.

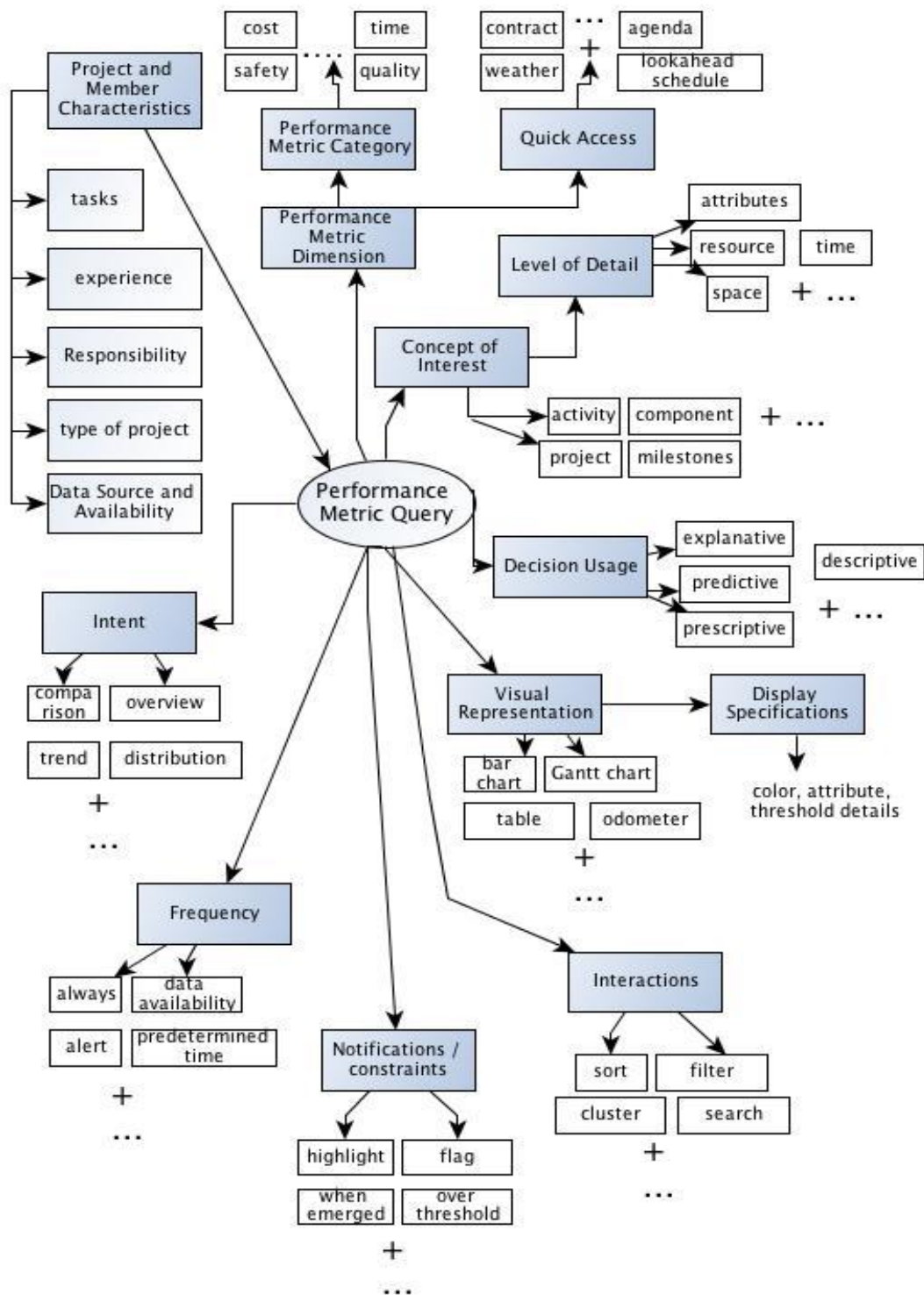


Figure 2.6: Affinity diagram of the consolidated interview results

The consolidation is followed by work redesign; creating storyboards from consolidated data to understand the ways to improve the users' work, user environment design; structuring the system work model, and defining its parts and their inter-relationship, mock-up and test with users; testing the developed ideas with users through prototypes. The last three steps of contextual inquiry are mostly related with developing a working prototype, which is out of scope for the information requirements identification purposes of this chapter.

2.5 Findings

Through the detailed analysis of the interview results, I observed that the information visualization needs of project team members change based on various factors, such as their responsibilities, decisions to be made, experience levels, and characteristics of the project. These information visualization needs have commonalities and differences in terms of the performance metrics monitored, the level of detail of the metric queries, and the frequency and ways of visualizing these metric queries.

2.5.1 Overview

The results of the fourteen interviews that capture a total of 132 *metric queries* were analyzed. The term *metric query* refers to the queries that are related with performance metrics and defined by the interviewees. The exploration of the metric queries revealed that not all of them were directly measurable with a certain method of calculation, but instead observation and visualization of the source of information to derive insights about the performance of the activities performed. Examples of measurable metrics are delta cost of projects, meaning the actual minus budgeted cost of the project, and turnaround

time of the documents exchanged such as submittals. Examples of source of information are look-ahead schedule, calendar of materials to be delivered and punch list items. Although sources of information are not literal metrics, they are tools and links project members requested as a must to help monitor project performance. Hence, they are included in the results section of the analysis.

In Table 2.5, the metric queries acquired from the interviews are presented under these two categories and further sub-grouped based on the type of the performance metrics, such as cost and schedule. Furthermore, short descriptions of the performance metric queries are provided with the associated number of queries covered. Even though these queries were consolidated within the same description, none of the queries were identical. For instance, it can be observed that four of the metric queries were related with delta cost of activities. However, they differ in terms of the time frames (activities performed for a day or week), locations (all activities or activities that are performed on certain work area), granularity (group of activities, an activity type or single activity), resources (cost of any combination of materials, crews, and equipments) and so on.

Table 2.5: 132 Performance Metric Queries combined under short descriptions ignoring specific LoDs

	Metric Query Category	Short Description	Total
Measurable Metrics	Communication	Comments on drawing and reports	1
		Details of documents	12
		Status of documents	11
		Turnaround time of documents	5
		EEO %	1
	Cost	Contingency	1
		Cost of Wasted Materials	1
		Projected Profit	1
		Delta CashFlow	1
		Delta Cost of Project	7
		Delta Cost of Activities	4
		Projected Cost overrun	2

Table 2.5 Cont'd: 132 Performance Metric Queries

	Metric Query Category	Short Description	Total
Measurable Metric	Productivity	Projected Cost overrun	2
		Causes of cost overrun	1
		Anticipated delta contract value	1
		Delta Cost of Resources	1
		Over-time % factor	1
		Delta Material procurement and usage	2
		Projected Resource Utilization	2
		Delta Units	2
		Units	1
		Delta Delivery schedule	1
		Delta Resource Utilization	2
		Units/MH	1
	Quality	Attributes of bids	1
		Cost of Rework	1
		Quality Problems	2
	Safety	Addressing inspection requirements	1
		inspection	1
		Safety related problems	5
	Time	Causes of schedule overrun	1
		Total Project Duration	1
		Delta Schedule	9
		Space conflict anchored in Gantt Chart	1
		4D Activity Progress	1
Source of Information	Communication	Action items of meeting minutes	2
		Activity To-Do	2
		Comments on drawing and reports	1
		Details of documents	1
		2D punch list	1
	Cost	Payments	1
	Productivity	Projected Resource Utilization	2
		Space usage	3
		Material procurement and usage	2
		Real time Resource utilization	1
	Quick Access	Agenda	3
		Lookahead schedule	2
		Delivery schedule	1
		Achievement of milestones	2
		Site tracking	5
		Weather	4
		Contact Information	5
		Widgets	2
		Contract and Specifications	2
		Schedule of Activities	2
	Time	Emails	1
		Total Project Duration	1

The summary of the number of the metric queries according to the job titles of the interviewees, and the performance metric categories are presented in Table 2.6. This table

shows that safety and quality are the least requested metrics, communication of documents and cost are the mostly requested metrics. It can be seen that in addition to the performance metrics, most of the interviewees requested to have access to sources of information, labeled as Quick Access information, such as weather, contract and drawings, and contact information. It can also be seen that the project member, who requested the most variety of the performance metrics, was the area manager, who is responsible from all aspects of multiple projects that he is managing. The main focus of a head of estimation department and owner representative was the communicated documents, to overview the status of the contracted works. Moreover, an interesting observation is that quality is not presented as a metric query neither by the project engineers nor by managers, but rather it is being tracked by project members with either highest or lowest status in a project organizational chart, such as an area manager and a superintendent.

Alarcon and Ashley (1996) stated that four to five performance measures are reasonable to obtain a relatively comprehensive evaluation of performance, however; in this research the average number of metrics asked by the participants was eight. The highest number of metric queries was requested by an area manager with an experience of 18 years, and one of the superintendents with an experience of 1 year. Based on all of these, I could not find a direct correlation between the number of requested information with the organizational hierarchy of the project team members in terms of job titles and the number of years of experience.

Table 2.6: Summary of performance metric categories per job title of interviewee

Job Title	Metric Category							Total
	Communic.	Cost	Productivity	Quality	Quick Access	Safety	Time	
Area Manager	2	6	2	1	1	1	1	14
Head of Estimation Department		1	3	1	1		2	8
Project engineer	1	4			2	1	1	9
Superintendent			4		3		1	8
District manager	4	2				1	2	9
Superintendent	4		2	2	6	1	1	16
Project engineer	4		2		1	1	1	9
Project manager	3	3			2			8
Project engineer	6	1			1	1	2	11
Project manager	2	1			3		1	7
Owner representative	8	1		1	1		1	12
Project engineer	2	1	1		4		1	9
Project engineer			4		1			5
Project manager	1	1	2		2		1	7

In order to observe if there is any correlation between the number of requested metric queries with other factors, I explored the variation in the number of metrics with the variation in the number of performance metric categories, years of experience of project members, and number of tasks (duties) of the interviewee. However, no correlation has been found.

2.5.2 Commonalities

Based on the developed affinity diagrams derived from the consolidated interviews, some commonalities amongst the visualization needs are identified. These commonalities include the way the information and visualization requirements are described using the

following concepts: Performance Metric Categories, Concept of Interest (CoI), Level of Detail (LoD), Visual Representations, Intent, Purpose of Use, Notifications, and Frequency. Based on how the project members queried the information during the interviews, each of these common concepts is presented in Figure 2.7 and can be described as follows:

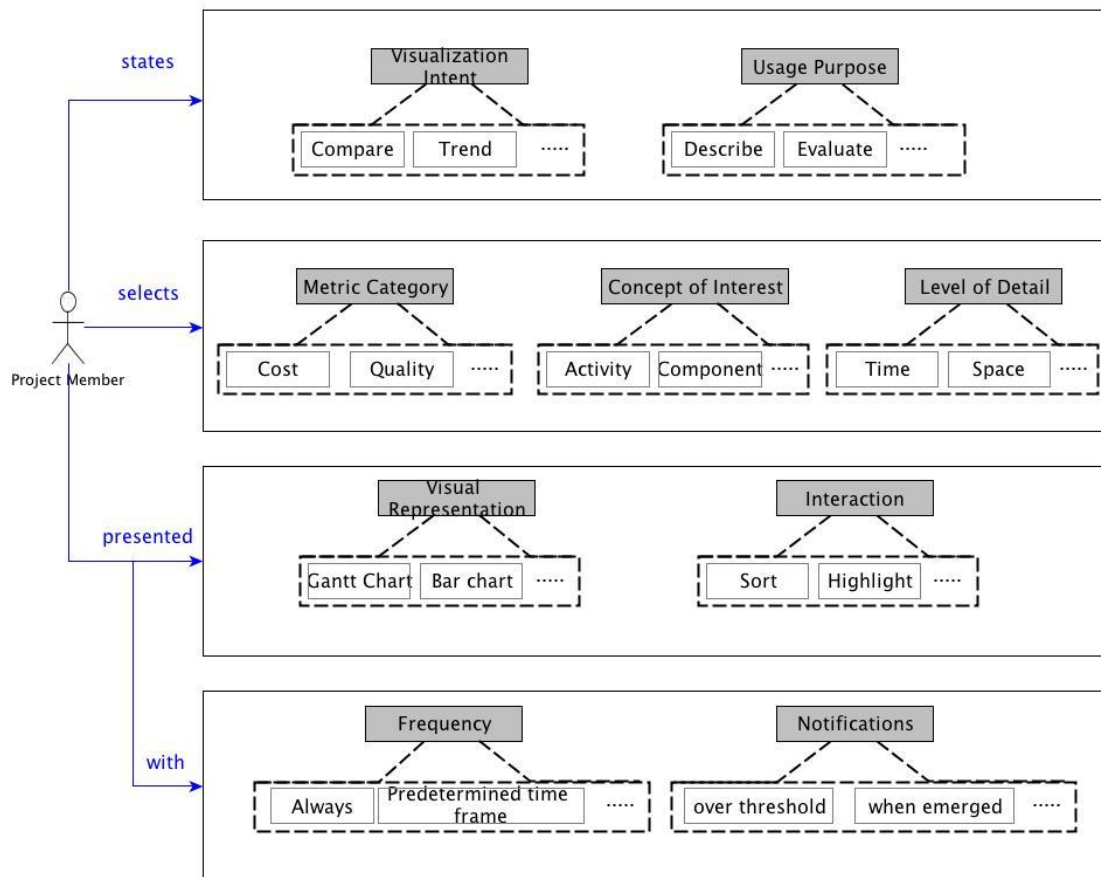


Figure 2.7: The commonalities identified based on the interviews

A project member states the type of analysis and decision s/he aims to make:

- Visualization Intent: The intent of the project member for visualizing the performance metric, such as comparison and trend analysis.

- Usage Purpose: Use of the performance information, such as checking if project is on track, and searching for reasons of problems.

A project member selects the context of performance metric query:

- Performance Metric Categories: Categories of measures against which a project's performance is measured, such as cost and safety.
- Concept of Interest: Facets of information that are the project specific concepts that a project member would like to learn more about, such as activities and building components
- Level of Detail: Hierarchies, decompositions and attributes of information items that project members desire to monitor.

A project member is presented with the visual representation and associated interaction options:

- Visual Representations: Collection of graphical symbols used to encode the properties and relationships of information, such as maps, charts and diagrams.
- Interaction: User input to perform a task with the visual representations, such as emphasizing, filtering, and overlaying extra information.
- Notifications: Alerts or highlights to remind prominent information. Unlike the other commonalities, interaction and notification are preferred by some of the interviewee.
- Frequency: The frequency with which the project member would like to monitor performance metric query.

While these are the commonalities identified through the interviews, each component has various enumerations based on the unique needs of the project members. In the following

section these differences are presented.

2.5.3 Differences in Metric Queries

In this section, the differences of metric queries are discussed regarding each of the commonalities listed above.

2.5.3.1 Performance Metric Categories

Seven categories of performance metrics are identified as a result of the consolidation of the interviews. These are: (a) cost, (b) time, (c) productivity, (d) quality, (e) safety, (f) communication of information, and (g) quick access. The distribution of the 132 performance metric categories is presented in Figure 2.8. It is observed that the communication, cost and time categories were commonly requested performance metric categories. The shares of performance metric categories are not claimed as a representative distribution regarding the AEC industry due to the fact that each organization can have different strategies that affect the selection of metric queries and related metric categories. The strategies of the organizations or the utilized construction methods and processes were not considered during the selection of the interviewees.



Figure 2.8: Distribution of Performance metric categories amongst all metric queries

2.5.3.2 Concept of Interest (CoI)

The information items requested by project members are related to specific concepts that a project member wants to learn more about or wants to monitor, such as activities, building components, budget and equipment. These concepts will be referred to as Concept of Interest (CoI). The types of CoIs requested by the project members are derived through the identification of the common interests through the consolidated interviews as follows:

- Activity (e.g., types groups)
- Component (e.g., building elements, utilities)
- Documentation (e.g., RFIs, submittals)
- Indirect/Overhead (e.g., contingency, profit)

- Issues and Inspection (e.g., problems, tests)
- Project (e.g., multiple, single project)
- Resource (e.g., material, equipment, labor)
- Space (e.g., floor, work area, room)
- Todos/Milestones (e.g., agenda items, milestones)

A list of CoIs has been generated through multiple iterations of clustering the consolidated interviews, and can be extended upon needed. For instance, documentation includes documents that are being generated and transmitted, such as submittals, RFIs and Change Orders. These items are categorized into the CoI of documents because the information items (attributes) requested per each of them have commonalities, such as due dates, status and responsible parties. However, if different analyses are performed regarding any of the documents, a new CoI can be created. Table 2.7 presents that each CoI is requested by most of the interviewees based on their job titles.

Table 2.7: The CoIs selected by different responsibilities of interviewees

CoIs	Area Mng	District Mng	Head of Est. Dept.	Owner Rep.	Project Eng.	Project Mng.	Super- intendent	Total
Activity	3	2	1	1	7	8	2	24
Component				1	2	1	1	5
Documentation	2	4	1	8	15	7	8	46
Indirect/ Overhead	2			1		1		4
Issues/ Inspection	1	1		1	4	1	5	13
Project	2	2	3		2			9
Resource	3		3		8	3	3	20
Space					1		3	4
Todos/ Milestones	1				3	1	2	7

2.5.3.3 Level of Detail (LoD)

The granularity of the performance metric queries is defined as its Level of Detail (LoD). For instance, measuring the cost of project has different LoDs in the following two examples:

- Actual versus budgeted cost of **major activity groups** performed to-date *detailed with labor, equipment and material costs.*
- Actual versus budgeted cost of **activities** performed last week.

In this example, it can be observed that there are differences in the performance metric queries in terms of the granularity of the activities (single versus groups), time frames (whole project duration versus last week), and attributes (detailed cost breakdown versus sum). By analyzing each of the performance metrics, I identified common LoD groups to represent these changes after a number of iterations. 11 groups are identified. These are activity, component (building elements and facilities), document (exchanged project related documentation), indirect/overhead (company or business level monetary information), issues and inspection (safety and quality related issues), project (number and status of projects), resource, space, time, and attributes (the attributes of information tuples captured within CoI categories, such as for CoI of activity has attributes of actual cost, budgeted cost, monthly, weekly values, etc.) The differences in the metric queries in terms of CoIs and LoDs, that are identified collectively through interviews are presented in multiple tables in the appendix, categorized based on CoIs. The higher level enumerations of LoDs and CoIs are presented in Table 2.8.

Table 2.8: Enumeration and Types of CoIs and LoDs

Category	Vocabulary / Enumeration
Activity	Activity, activity group, activity type, activity on CP, self-performed, subcontracted, require inspection, interesting operations
Component	Component, utility, unit, all components
Documents	Bid, CO, contact #, contract, design and drawings, EEO Report, Field Report, meeting minutes, punchlist, RCO, RFI, submittal, schedule versions, transmittals, outstanding items, contact information, contract and drawings, emergency contact numbers and MSDS, new e-mails, subcontractor contacts information,
Indirect/Overhead	Cashflow, contingency, payment, profit, projected profit
Issues and Inspection	safety issues, quality issues, inspection and testing, weather
Project	Multi-project, current project, targetted project
Resource	Craft, crew, equipment, estimator, foremen, labor, material, subcontractor
Space	Elevation, floor, room, site, work area
Time	3 weeks, biweekly, calendar, daily, minutely, monthly, weekly, whole project
Attributes	actual, budgeted, rejected, accepted, commented, submitted, without accidents, threshold over a value, % completion, affecting schedule, require inspection, status, due date, MH required number and so on..
ToDo/Milestones	Action items, agenda, milestones

2.5.3.4 Visual Representations

While stating the performance metric queries, the project members indicated the types of the visual representations that they would prefer to track the specific particular performance metrics. I have observed that project members used some terminologies interchangeably or mistakenly resulting in requesting a visual representation while referring to another one. For instance, some of the project members used the wordings spreadsheet, table, list and summary to refer to the need to present information in a tabular form with no interaction requirements. During the elicitation of the results, considering the needs of the project members, the mistakenly selected terminologies are replaced with the correct ones. Hence, organizing the visual representation definitions, I have identified a total of 22 visual representations that were requested by the participants to visualize their metric queries, as presented in Figure 2.9.

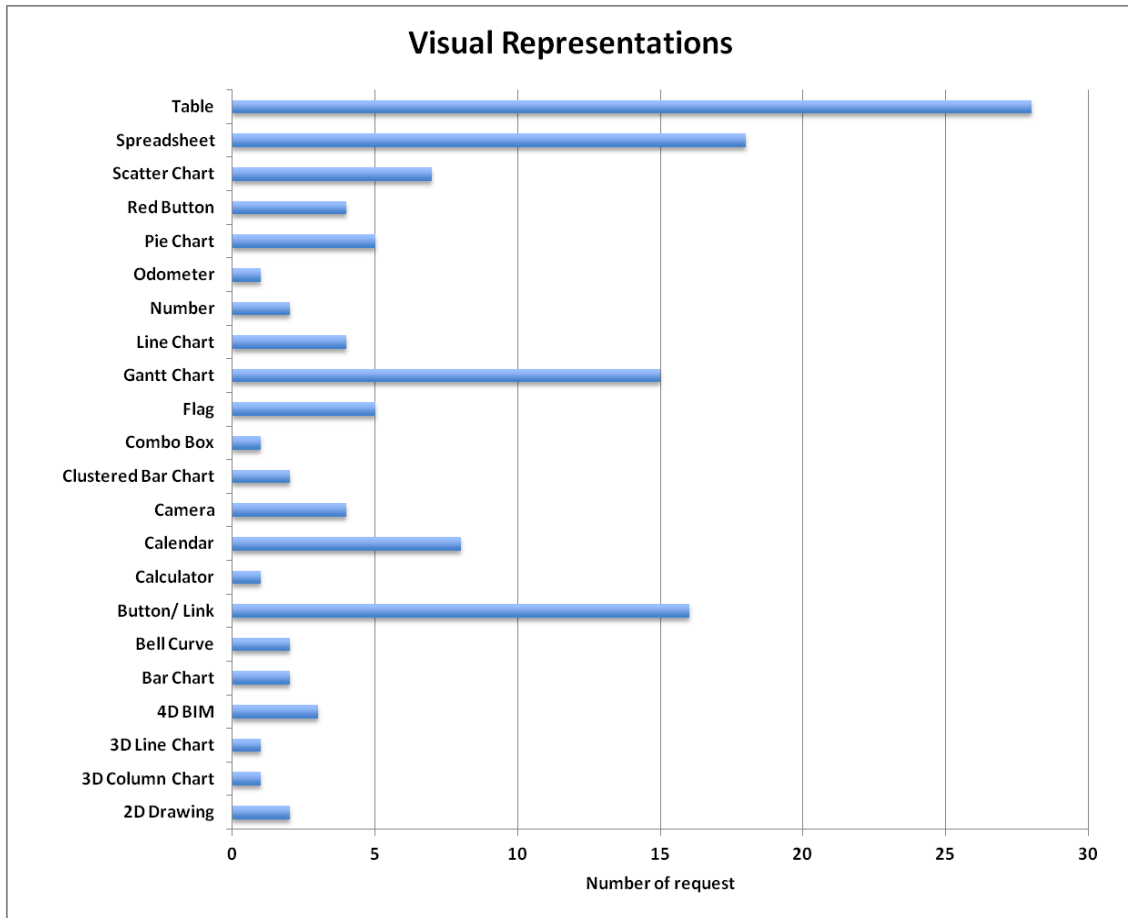


Figure 2.9: Visual Representations selected by project members

The variation of the requested 22 visual representations with performance metric category is presented in Figure 2.10. It can be understood that mostly requested visual representations are buttons/links for quick access to information, Gantt Cart for time, table for communication of documentation and spreadsheets for cost. Hence, it can be deduced that the construction project members are used to the traditional methods of information visualization.

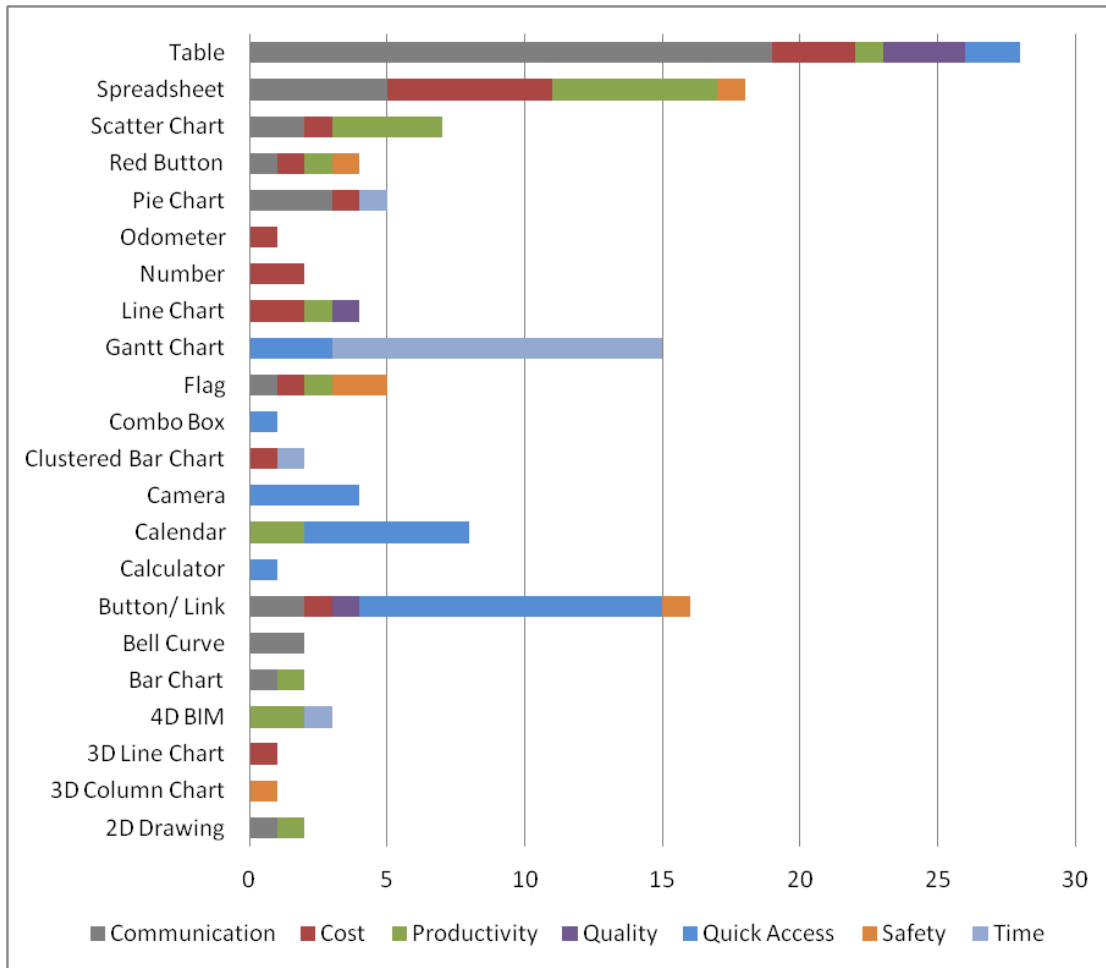


Figure 2.10: The distribution of the visual representations amongst the performance metric categories

2.5.3.5 Intents and Usage Purpose

Project members have certain intents and associated decisions while monitoring the performance of the projects. I observed that the most common intent is comparing the actual and estimates to see if something is out of track; and the decisions are mostly related with prescriptive purposes, meaning getting into contact with other project members (calling other project members or accessing to more detailed information).

The decisions that project members make with the visualization of performance metric differ in terms of the usage purpose of the performance metric. I categorized the usage of information as *descriptive*, *explanative*, *predictive*, *evaluative* and *prescriptive*. These five categories are used by O'Shaughnessy (1972) as types of decisions and the first four by Liston et al. (2000) as types of tasks performed in construction review meetings. *Descriptive* usage describes observing the current status and needs; *explanative* usage is corresponding with somebody or looking at detailed reports to learn about the reasons of the current situation; *prescriptive* usage is notifying other team members of discrepancies; *predictive* usage refers to performing planning and future projection; and *evaluative* usage is for the purpose of making an overall self-performance evaluation.

While the usage of purpose describes the types of decisions that project members would like to make, the intent describes the initial purpose of analyzing the data. Chengzhi (2003) identified a set of intents, such as overview-query (get an overall review), comparison (compare two or more things), cluster-classification (group two or more things), distribution pattern (shares something among multiple things), and dependency-correlation analysis (relationship btw two or more things). During my PhD research, I used the list of intents identified by Chengzhi (2003) as the foundation, and extended it according to the needs of construction industry. For instance, the projection cannot be categorized into any of the intents defined for the given example above. The intents used in this study are clustering, comparison, composition, correlation, detail on demand, distribution, overview, overview and detail and trend. Some of the metric queries are composed of more than one intents. For instance, projected delta of the project cost is composed of both comparison of actual vs. budgeted cost and trend over time. Figure

2.11 demonstrates that nearly half of the metric queries were related to getting an overview of the performance, followed by comparison and detail on demand.

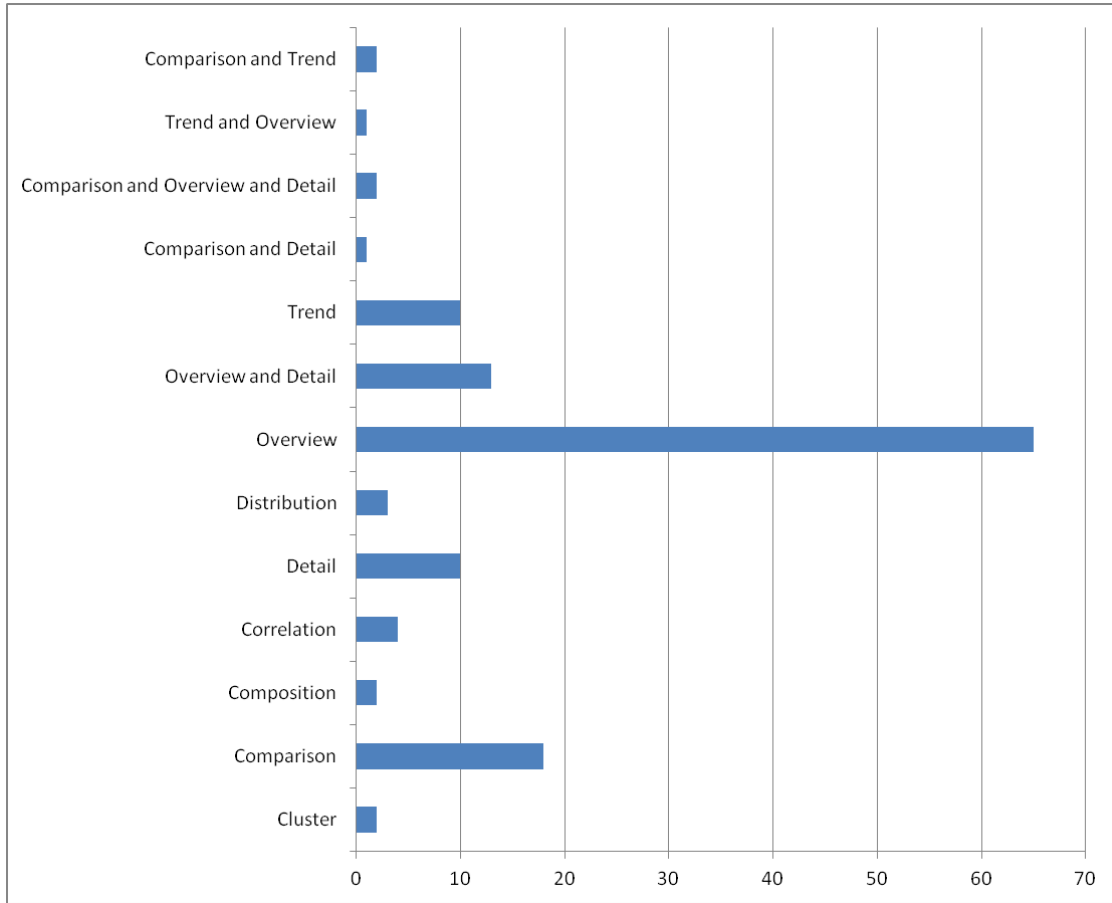


Figure 2.11: Distribution of intents

2.5.3.6 Interactions, Frequency and Notifications

Dynamic display of information is defined as allowing a user to interact with the output on the screen (Pfitzner et al. 2000), and also as the data that is presented in a visualization environment to change (Doolin 2010; Glossary:D 2010). Building on these two definitions, such a change can be based on two phenomena: (1) the information changing constantly or over a pre-defined cycle, and (2) the information requested is generated on-

demand. The first one, which can also be referred as information push, happens when a project member wants the information to be automatically provided based on his/her predefined frequency preferences. For instance, in the case of an area manager, he wants to be notified once a week if there is a cost overrun over a threshold value for all projects that he is responsible from. To facilitate that, the pertinent data defined by the user should be checked and notifications should be provided accordingly. For the second phenomena, which can also be referred as information pull, a project member accesses the information when s/he needs by clicking on a visual representation and getting further details, or sorting and filtering information.

Amongst 132 performance information requested by the interviewees, 79 % were performance metrics and 21 % were quick access information. Figure 2.12a shows that 82% of the performance information was initially requested, and the rest were further representations. The term “further representation” refers to a visual representation that is generated when a user selects an item on a displayed visual representation. For instance, one project engineer wants to be able to see the 3D as-is model of the project (which is the further representation) when she visualizes the weekly schedule and clicks on an activity, similarly a district manager wants to visualize the causes of schedule overruns on a pie chart when visualizing the schedule of the on-going projects.

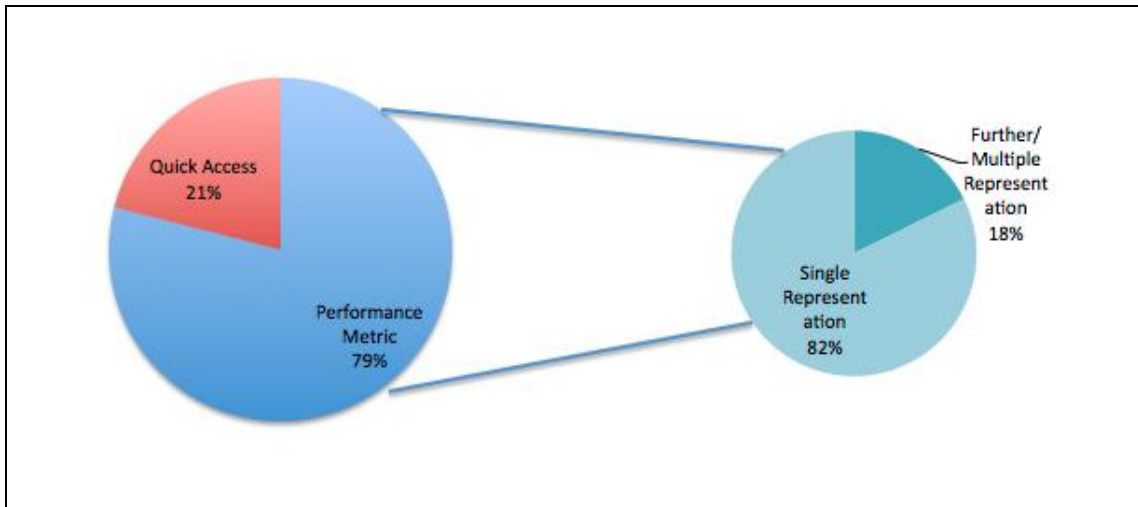


Figure 2.12.a: Distribution of Performance Metrics

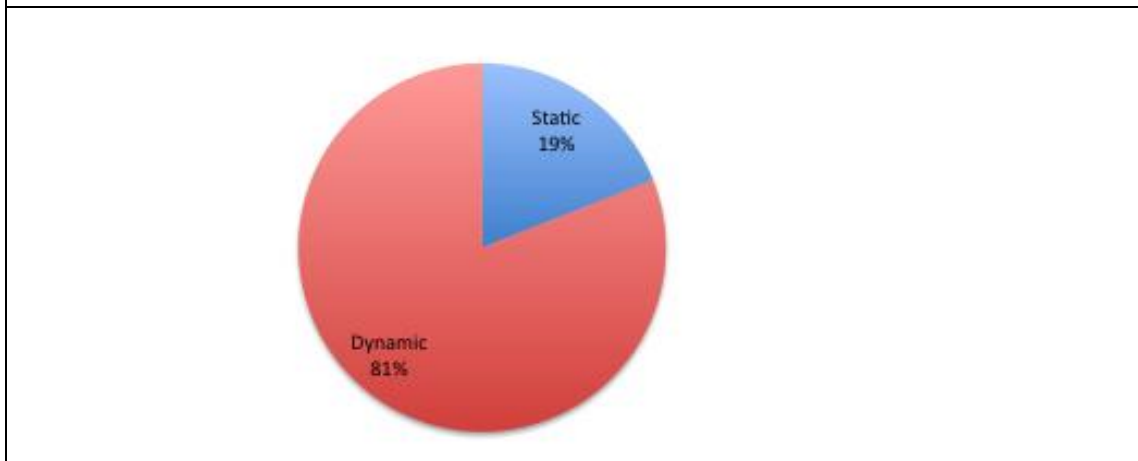


Figure 2.12.b: Dynamic vs. Static Interaction Needs

Figure 2.12: Distribution of performance metrics based on all metric queries

In Figure 2.12.b, I present the distribution of performance metrics according to dynamic interaction needs. For 19% of performance metrics, the project members prefer not to have any dynamic interaction, labeled as “static” in this figure. However, for the rest of the performance metrics, they would like to dynamically interact with a visual environment through either on-demand information requests (including sorting, filtering, zooming, quick access shortcuts, and further representations) or through pre-defined

notification and frequencies. Frequencies show how often the project members want to observe a performance metric. More detailed discussions on such interactions, frequencies and notifications can be found in the following subsections.

When the performance metrics are visualized, the ways that the project members would like to interact with it differ. While some project members ask for visualizing additional information, some want to perform several tasks on the current data, based on their intents. Amar et al. (2005) identified a set of ten analytical tasks that are independent of the visual representations. These tasks include retrieve value, filter, compute derived value, find extreme, sort, determine range, characterize distribution, find anomalies, cluster, and correlate. Wehrend and Lewis (1990) defined eleven tasks that are more closely related to the intents of the user. These tasks include identify, locate, distinguish, categorize, cluster, distribute, rank, compare within entities, compare between relations, associate, and correlate. Building up on these research studies, I have identified the interaction tasks that the project members requested as: Alert (set notifications), Cluster (group variables in a spreadsheet), Filter (filter out based on a criterion), Highlight (emphasize a variable), Link (acquire extra information through further representation), Navigate (navigate through a 2D or 3D plan), Retrieve value (retrieve some statistical values), Sort (sort variables based on a criterion), Search (look up for certain variables or values), and Simulate (animate a 3D model).

Frequency

Frequency of accessing to the information: Every project member had a certain frequency for monitoring the metric queries. I have grouped the frequencies under five types:

- *Level 1-Always*: For ~48% of the metrics, the project members want to have the relevant information to be available at all times.
- *Level 2-Always with notifications*: For 13% of the metrics, project members requested to have the relevant information available at all times with an additional criteria on highlighting or notifying prominent information;
- *Level 3-Based on data being available*: For 26% of the metrics, project members preferred to be informed, when information is updated, or has become available. This level also captures the quick access information needs.
- *Level 4-Based on pre-determined time-frame*: For 4% of the metrics, project members want to be informed about a performance metric within a pre-defined time-frame; such as once a week, or once a month.
- *Level 5- Further Graphs*: For 9% of the metrics, project members want to interact with metric queries (such as drill, link, integrate etc.) to visualize relevant information in a new visual representation. The frequency depends on the frequency of the visualization of the initial metric query and the intent of the project member to further the analysis.

Notifications

For the 28% of the performance metrics, project members stated that they would like to receive notifications on the displayed metric when it does not meet certain criteria. Such notifications differ in three major aspects:

- *The context of the information that they are applied to*: time-frame of the information (e.g., past, present, future), the CoI (e.g., activity, component, resource), and the associated performance metric category (cost, safety).

- *The criteria to find prominent information:* due/overdue items (e.g., delay of 7 days in submittals), information items over a value (e.g., schedule variance over 25%), and information items that recently surfaced (e.g., if there is a space conflict).
- *The visual methods selected:* using an icon (e.g., flashing flag), or highlighting information within its context (e.g., color-coding according to magnitude of a problem).

I explored if the notifications are assigned based on the power a project member can assert to improve the performance metric measurement. The project members can (a) Control: change the conditions to improve the performance, (b) Manage: cannot change the conditions, but can manage the outcomes, (c) No Control: cannot control or manage the metrics, can only expect the results. Figure 2.13 depicts that most of the notifications are requested for the controllable performance metrics as expected. For the metrics that cannot be controlled, project members are only interested in instant notification in the form of buttons and flags, not highlighting prominent information.

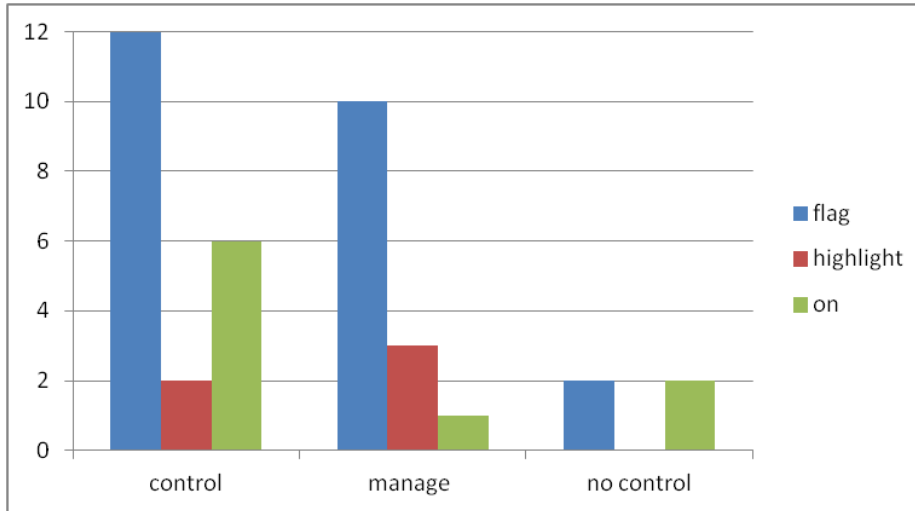


Figure 2.13: Notifications status based on controllability of the performance metric queries

In this section, the differences and enumerations of the interviews are discussed based on the previously identified categories of commonalities. In the next section, the differences will be explored across demographics.

2.5.4 Major Differences Based on Demographics

The variation in the performance metric category requests based on the responsibilities of project members can be found in Figure 2.14. It can be seen that while an area manager is mainly interested in the cost aspect of the project, he tracks the project from all aspects; as opposed to district manager, who ignores productivity and quality issues due to his responsibilities. However, I have observed some patterns in the selection of performance metric categories. It can be seen that the most frequently requested performance metric is exchange of information followed by the cost and time categories. Moreover, while all of the project members were interested in the time performance of the projects, safety and

quality were of interest to only six out of fourteen interviewees. Additionally, except for the superintendents, all interviewees wanted to track the cost performance of the projects. The exchange of information and productivity categories showed a trade-off for most of the interviewees. In other words, the interviewees, who requested one category of information item, did not request the other for most of the cases. Hence, I hypothesized that this trade-off could be due to the responsibilities of project members in the organization or the type of projects that they were engaged in. First, I investigated the project members' performance metric category selections based on their responsibilities. Figure 2.15 shows the total number of productivity and exchange of information related metric queries requested based on different responsibilities of project members. It can be seen that the project members with the responsibilities of area manager, project engineer and project member asked for similar numbers of two metric queries. However, while district manager and owner representative were only interested in exchange of information, head of estimation department and superintendent were interested in only productivity.

Second, I investigated the distribution of the number of the metric queries formed by interviewees according to the type of projects, as shown in Figure 2.16. Figure 2.16.a shows the results of the metric queries acquired from seven project members experienced in heavy/civil infrastructure projects, whereas Figure 2.16.b shows the results of interviews with seven project members engaged in commercial projects. It can be seen that while the highest interest is on exchange of information for commercial projects, the major interest is productivity for highway projects. The main reason for this difference

arises from the fact that majority of the works commercial and residential projects are subcontracted out, where in highway projects much less is subcontracted out and more is self-performed.

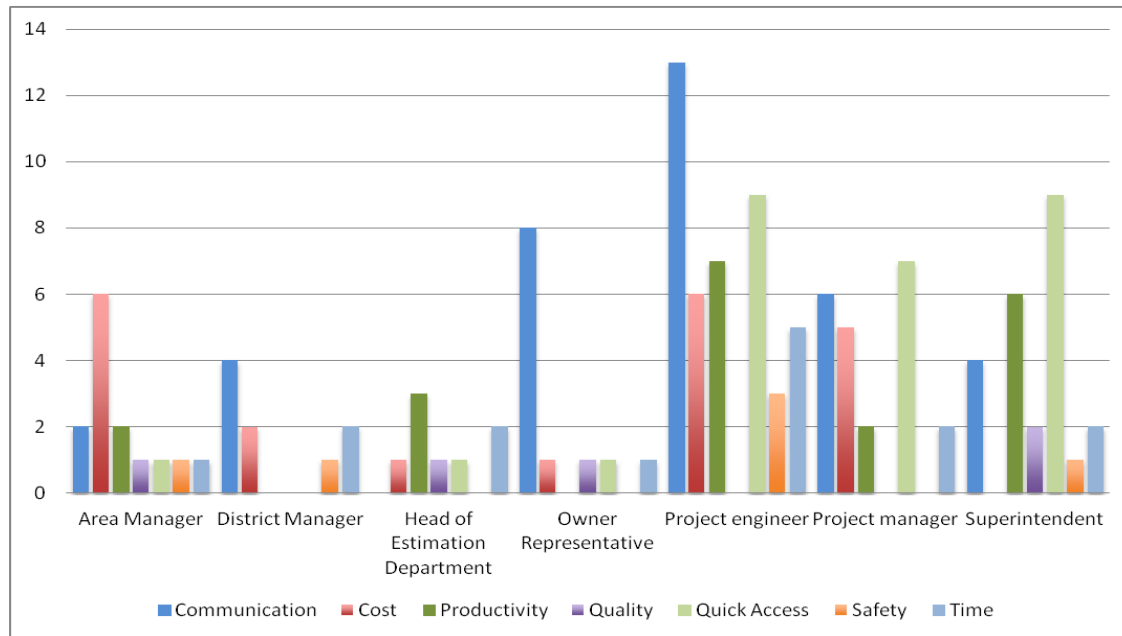


Figure 2.14: Performance Metric Categories based on project member responsibility

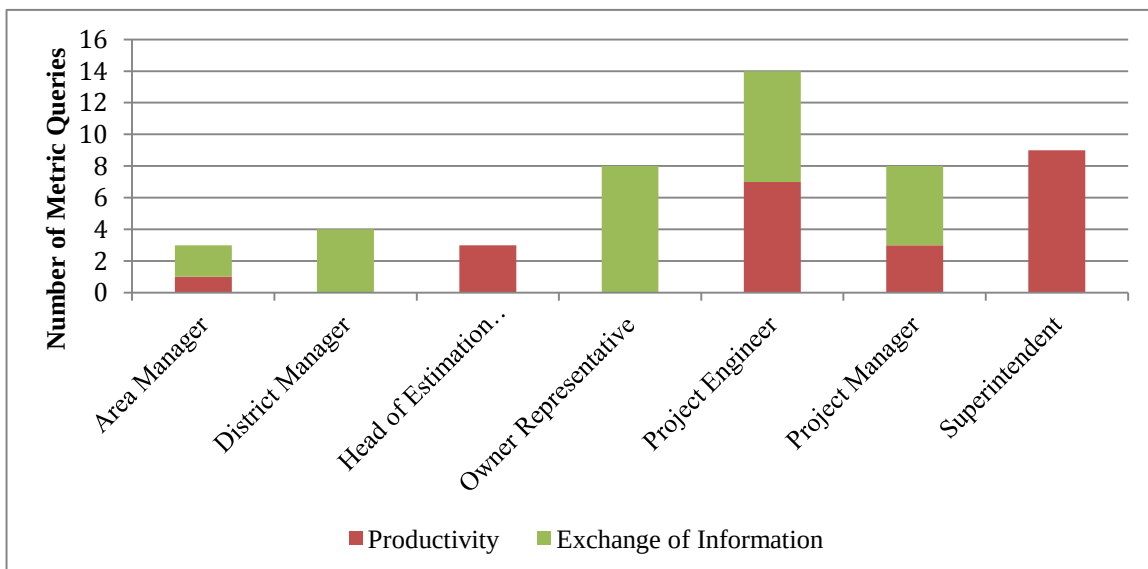


Figure 2.15: Trade-off between exchange of information and productivity performance metric categories based on responsibility of the project members (the numbers are not normalized for the number of interviewee)

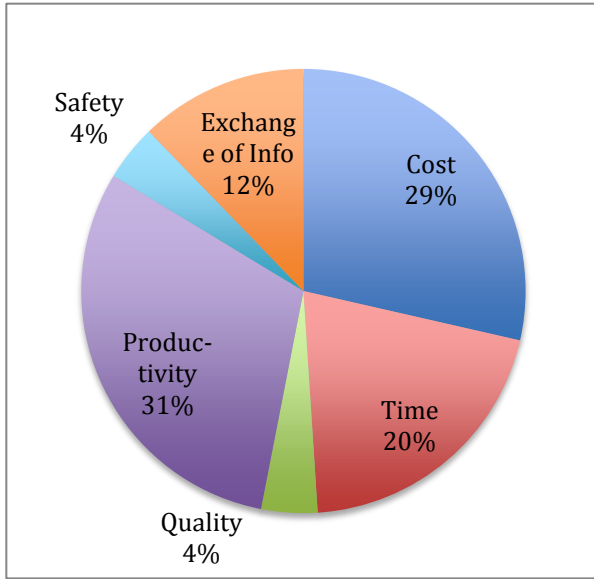


Figure 2.16.a. Highway Projects

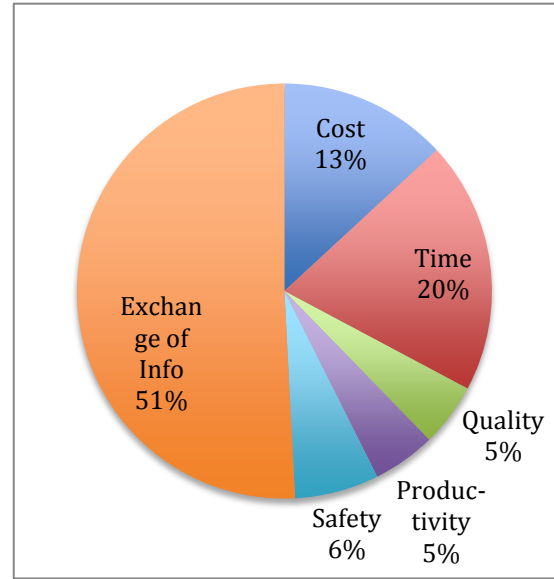


Figure 2.16.b. Commercial Projects

Figure 2.16: Distribution of the identified metric queries based on their performance metric categories and the types of projects interviewees are experienced with.

By investigating the distribution of the metric queries according to the CoIs through building and highway projects, I identified that for four out of nine CoIs, there is no significant difference according to the type of the projects. However, as shown in Figure 2.17; while the activity, and project CoIs are more widely preferred for the highway projects, component, documentation and space CoIs are more widely preferred for the building projects. Similarly, the comparison of the selection of the CoIs amongst different responsibilities of the project members is demonstrated in Figure 2.18. It can be seen that while area manager, district manager and head of estimation department requested similar number of CoIs, the main focus of owner representative, project manager, project engineer and superintendents is the documents exchanged as expected due to the fact that

they are mostly working for commercial projects.

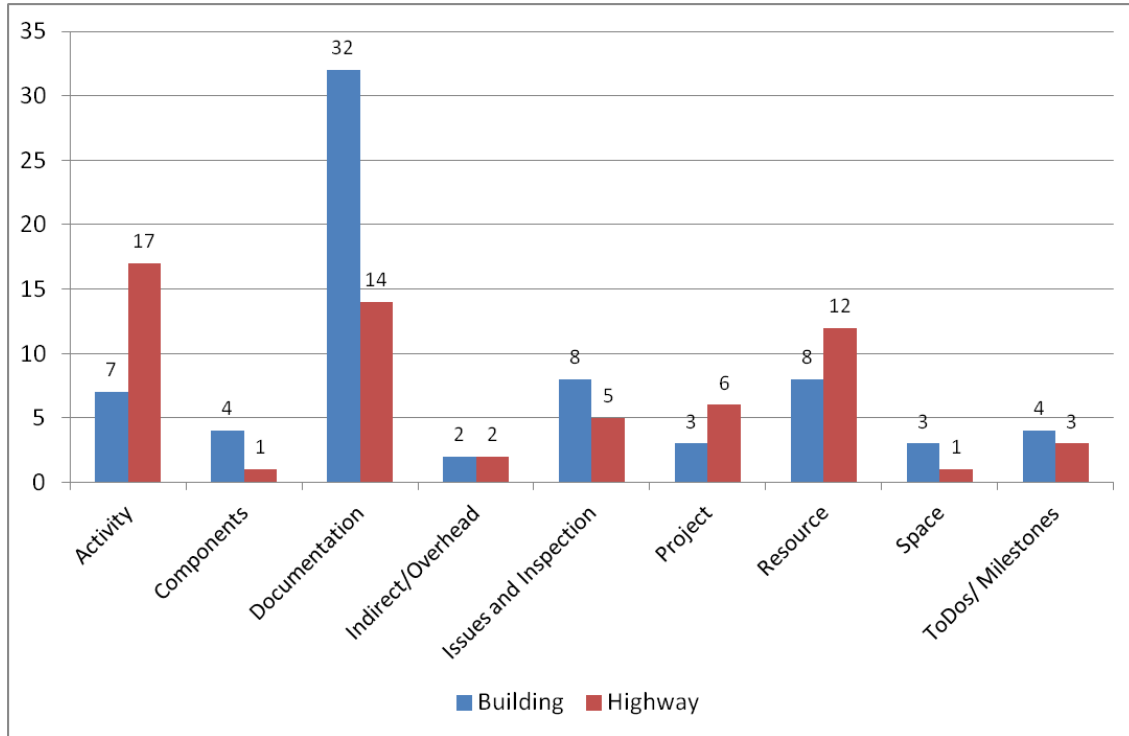


Figure 2.17: Distribution and Comparison of metric queries based on CoIs and the types of projects interviewees are experienced with

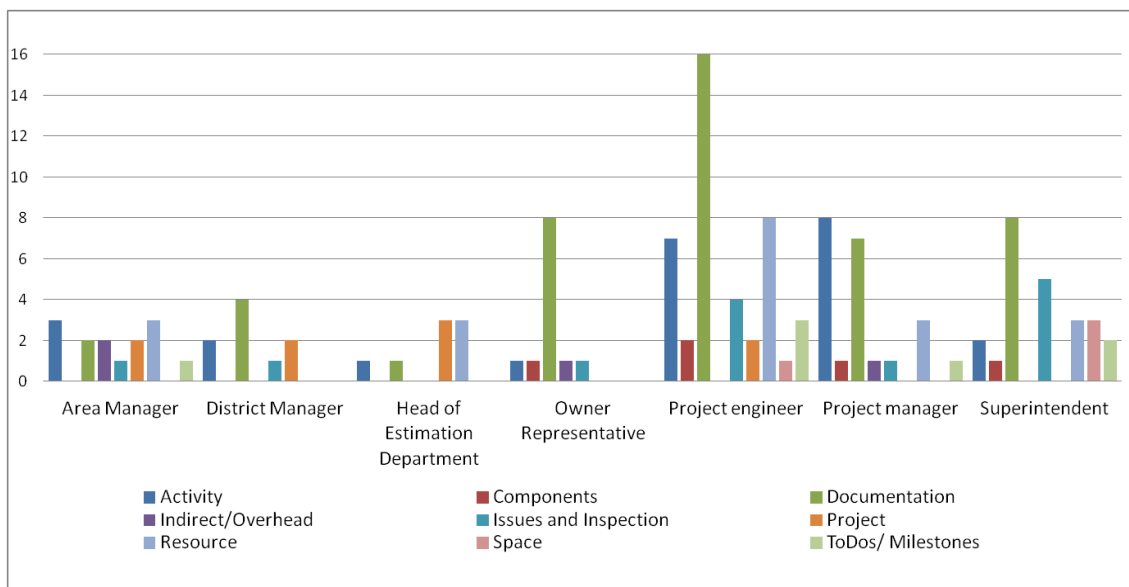


Figure 2.18: Distribution and Comparison of metric queries based on CoIs and the responsibilities of the interviewees

2.6 Analysis Results and Validation

The discussions of this research study are presented based on the findings that each performance metric query of the interviewee is different and can be categorized under certain commonalities. The commonalities are validated using principal component analysis (PCA) and the differences are validated using similarity analysis.

Principal Component Analysis

The validation of the commonalities is performed with the use of a mathematical clustering approach, as principal component analysis (PCA). PCA analyzes a data table representing observations described by several dependent variables, which are, in general, inter-correlated. Its goal is to extract the important information from the data table and to express this information as a set of new orthogonal variables called principal components. PCA also represents the pattern of similarity of the observations and the variables by displaying them as points in maps (Abdi and Williams 2010).

In Appendix B, the interviews' results spreadsheet is provided. For the validation of the differences, I used all 132 metric queries identified through the interviews as the observations. The dataset composed of a 132x26 matrix, with the variables defined as: interviewee ID (to characterize the user perspective), visual representation, performance metric category, CoI, intent, interactions, notifications, frequency, usage purpose, and fifteen variables that define the LoDs including space, activity, documentation, time, actor, component, project, overhead, inspection and issues, and attributes.

Since the method employs quantitative data, all of the categorical data is converted into numerical numbers based on the total number of different inputs in each of the variable categories. For instance, the quantification of the performance metrics category is presented in Table 2.9.

Table 2.9: Numerical values assigned to metric categories for the PCA analysis

Metric Category	Assigned Value
Cost	1
Time	2
Productivity	3
Communication	4
Quick Access	5
Quality	6
Safety	7

The list of the twenty-six original variables and their column numbers are presented in Table C.1 in Appendix C. For instance the first column is interviewee ID and the second column is the visual representations. The PCA analysis provided me with twenty-six principal components (PCs) that are a linear combination of the twenty-six original variables. The coefficients are provided in Tables C.2 and C.3 in Appendix C. Table C.2 captures the initial 16 columns and Table C.3 captures the columns through 17 to 26. These coefficients are presented in a color-coded map in Figure 2.19, with original variables in the x-axis and PCs in the y-axis. In order to identify a total of how many PCs are enough to represent the 90% of the variation in the dataset, I have calculated the latent of PCAs, in other words the eigen-values of the covariance matrix. The latent values for each of the PCs and their cumulative values are presented in Table 2.10.

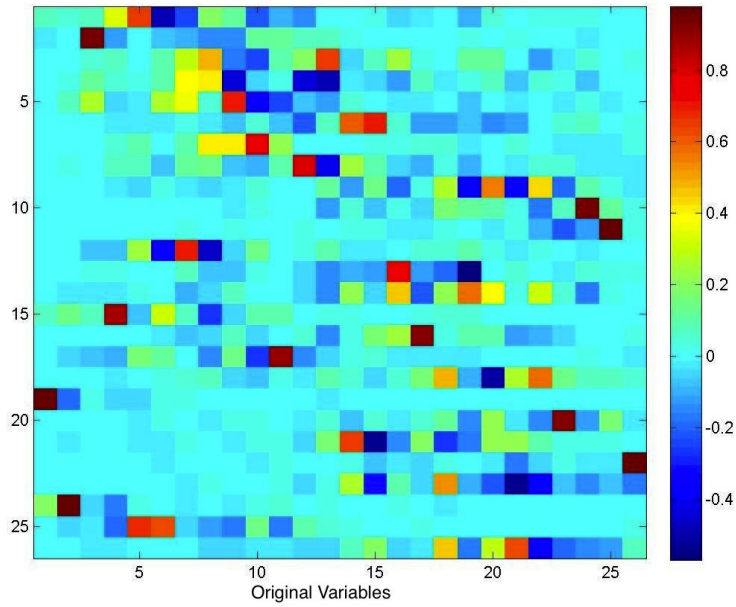


Figure 2.19: The color-coded representation of coefficients of original variables in forming PCs.

Table 2.10: Latent and cumulative latent values for the PCA analysis

Latent Values	Cumulative Latent Values
358.2367	66.7994
71.8871	80.2040
33.6300	86.4749
21.5270	90.4890
17.7080	93.7909
8.3378	95.3457
6.1704	96.4963
5.0005	97.4287
3.5319	98.0873
2.8851	98.6252
2.3074	99.0555
1.4749	99.3305
0.7124	99.4633
0.5325	99.5626
0.4634	99.6491
0.3635	99.7168
0.3226	99.7770
0.2884	99.8308
0.2584	99.8790
0.1758	99.9117
0.1389	99.9376
0.1355	99.9629
0.0713	99.9762
0.0635	99.9880
0.0407	99.9956
0.0235	100.0000

Table 2.10 presents that while the first two PCs can represent the 80% of the variation; four PCs are required to account for 90% the variation in the dataset. Figure 2.20 plots the dataset on the first two PCs. As can be seen from Figure 2.20, the data points are located with a certain trend according to the PCs, with the exception of a few points, as highlighted with red and blue colors.

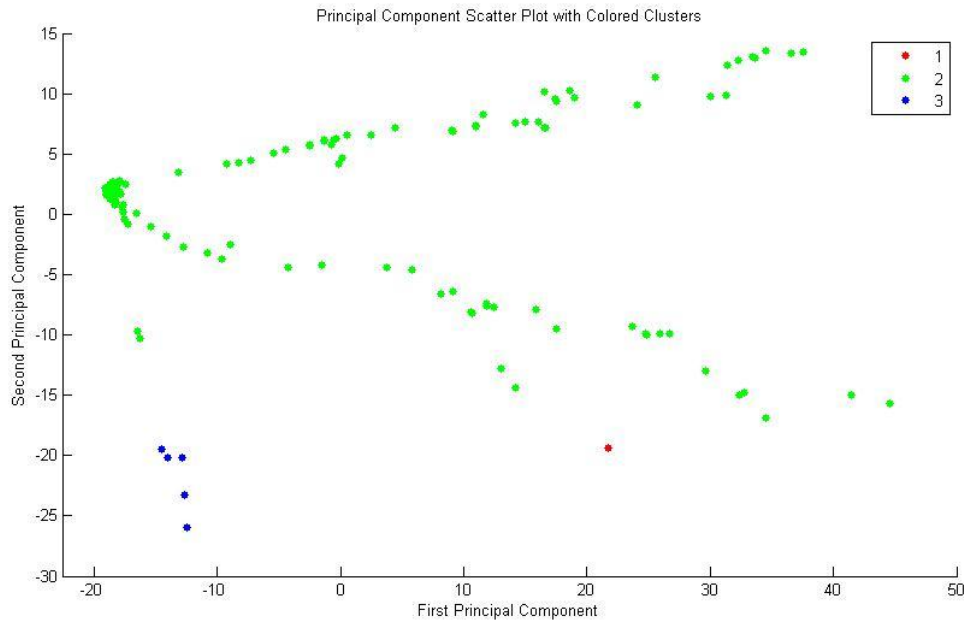


Figure 2.20: The dataset plotted against the first two PCs.

As a next step, I identified the original variables within these four PCs with the highest coefficients in order to determine the most influential ones. Coefficients smaller than 0.1 are considered as not contributing to the major variance, whereas the rest are considered as influential. The distribution of the original variables within the PCs, grouped according to the coefficients being smaller or greater than 0.1 is presented in Table 2.11. The column rank of the variables (e.g., interviewee ID is 1st and visual representations is 2nd) are presented in this table, and Table C.1 in Appendix C can be referred to learn the

associated variables. From this analysis, it can be seen that the variables in the columns 3,4,5,6,7,8,9,10,11,12,13,16,19,20 and 22 are considered as influential. The same result can also be derived from Figure 2.19 presented above. These components are metric category, CoI, intent, notifications, use, frequency, and LoD in terms of project, component, actor, resource, space, overhead, and attributes.

Table 2.11: The distribution of variables according to the magnitude of their coefficients in the representation of the PCs

	Coeff >0.1	Coeff < 0.1
PC1	4,5,6,7,8,9,10,12	1,2,3,11,13,14,15,16,17,18, 19,20,21,22,23,24,25,26
PC2	3,4,8,9,10,11,	1,2,5,6,7,12,13,,14,15,16,17,18, 19,20,21,22,23,24,25,26
PC3	7,8,9,10,12,13,16,19,20,22	1,2,3,4,5,6,11,14,15,17,18,21,23 24,25,26
PC4	3,7,8,9,12,13,16,	1,2,4,5,6,10,11,14,15,16,17,18, 19,20,21,22,23,24,25,26

The variables that are identified as commonalities but not selected as influential components are visual representations and interaction. The project members have selected only twenty-two visual representations, within which visuals, such as tables, spreadsheets and flags, were requested with high frequency. Hence, it is understandable that the visual representations accounted for only trivial amounts of variance in the analysis and were not identified as an influential component. The reason why interaction is not identified as an influential component could be due to the fact that a binary (yes, no) identifier is used for this variable. Through the interviews a comprehensive list of interaction options are not identified, hence instead of using a quantification of the interaction types (sort, highlight, zoom), whether interaction is requested by an interviewee or not is used. Since 81% of the project members requested to interact with the visual representations, they did not contribute to significant variance, resulting in having a lower coefficient. Moreover, it is known that PCA is not very robust when there are outliers in a dataset. It can be interpreted that in a case with absence of the outliers

identified in Figure 2.20, and intents assigned with their values, the PCA analysis could have identified all of the major variables as influential. In order to explore and validate the differences within the metric queries, similarity analyses are performed.

Similarity Analysis

There are numerous similarity measures that are used to explore similarity within a categorical dataset. Boriah et al. (2008) compared the performance of fourteen measures and identified that *overlap* is the most frequently used method and *goodall* (used as *goodall3* in Boriah et al. 2008) is one of the best measures for a categorical dataset.

- *Overlap* measure: For two categorical data rows in a multivariate dataset, the similarity between them is proportional to the number of variables in which they match (Boriah et al. 2008). The similarity scores take values btw 0 and 1, where 0 is assigned when there is no match, and 1 is assigned when the variables match.
- *Goodall* measure: While *overlap* measure gives equal importance to matches and mismatches, the *goodall* measure takes into account the number of occurrence of values within the variables. This measure assigns higher similarity to a match if the value is infrequent than if the value is frequent.

The formulae used to calculate the similarities are presented in Figure 2.21. Similarity of the dataset is calculated as a cumulative of multiplication of weight and similarity between two rows. Weight is $1/d$ where d stands for the total number of variables, and similarity between two rows for the m^{th} variable ($S_m(X_m, Y_m)$) is calculated differently for *Overlap* and *Goodall* measures. In *Overlap*, if the variables are equal, then the similarity

score is 1. In *Goodall*, a $P^2(x)$ value is calculated which is a function of frequency ($f_m(x)$) of the x value in m^{th} column. The frequency defines the number of times the x value occurs in the m^{th} column.

$$S(X, Y) = \sum_{m=1}^d w_m S_m(X_m, Y_m)$$

$$S_m(X_m, Y_m) \xrightarrow{Overlap} = \begin{cases} 1, & \text{if } X_m = Y_m \\ 0, & \text{otherwise} \end{cases}$$

$$S_m(X_m, Y_m) \xrightarrow{Goodall} = \begin{cases} 1 - P_m^2(x), & \text{if } X_m = Y_m \\ 0, & \text{otherwise} \end{cases} \quad \text{where } P_m^2(x) = \frac{f_m(x)(f_m(x) - 1)}{N(N - 1)}$$

Figure 2.21: The Overlap and Goodall similarity analyses using the equations in Boriah et al. 2008

In this study, both overlap and goodall measures are utilized and 132x132 similarity matrices are acquired. The results are evaluated in two folds: (1) overview of the results with color-coded charting, and (2) details of the results with statistical values for the matrices. Figures 2.22 and 2.23 present the color-coded similarity matrices for the *Overlap* and *Goodall* measures. Since the similarity of a row with itself is always the highest score (1 in the case of *Overlap*, and 19 in the case of *Goodall*), the diagonals in the matrices bear the highest scores. However, the similarity of a row with itself is not relevant and the scores of the diagonals being highest is an artifact of the analyses. Hence, to eliminate this artifact, I replaced the values of the diagonals with the highest similarity scores acquired from comparison of two different rows. It can be seen that while most of the metrics have a similarity value of 0.5 or higher for the *Overlap* measure, the scores are lower than seven with the Goodall measure. The reason why

Overlap measure gives higher scores stems from the fact that the frequency of the variables are not considered in that analysis.

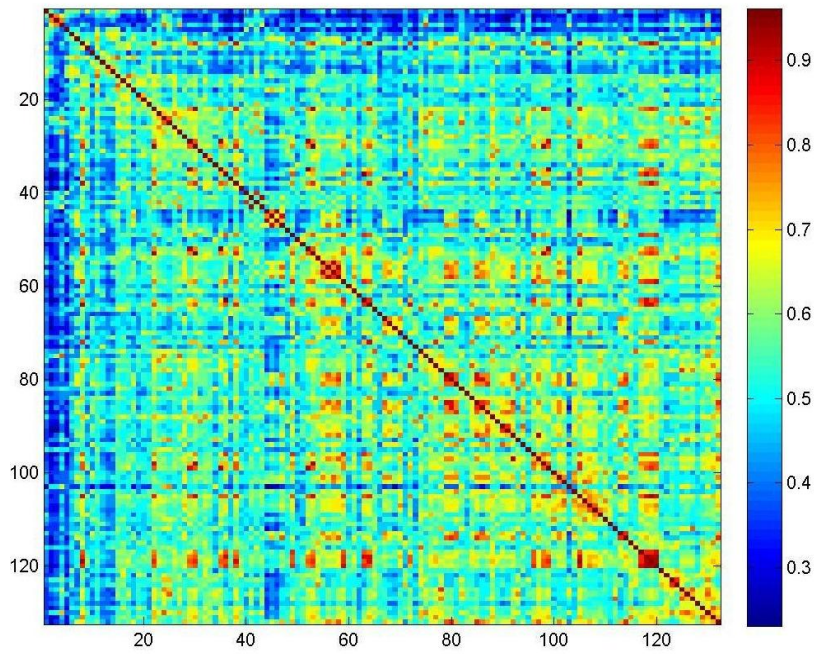


Figure 2.22: Similarity matrix for Overall measure

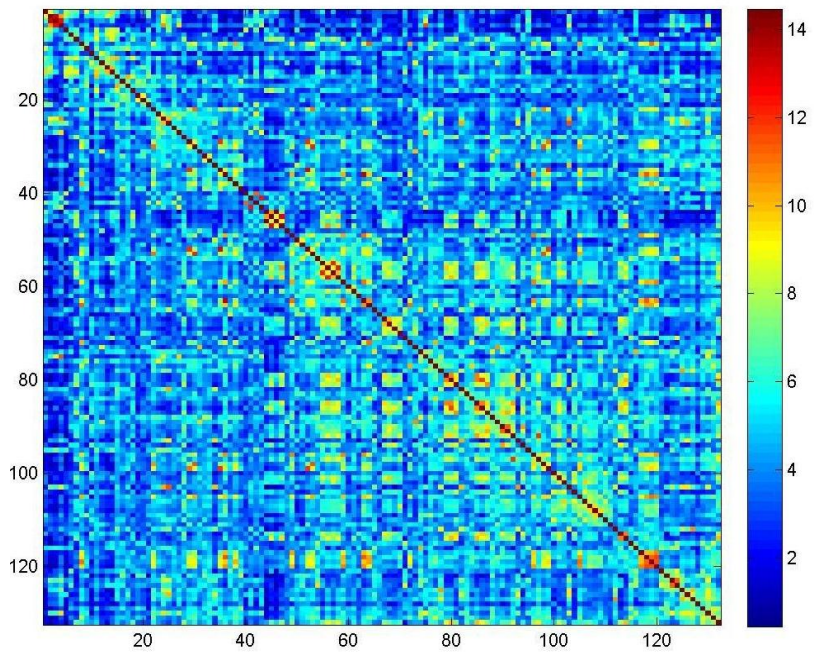


Figure 2.23: Similarity matrix for Goodall measure

Values concentrated in the left most columns and top rows that belong to the metric queries of area manager have a higher dissimilarity ratio than the rest of the metrics. Similarly, the metric queries that belong to district manager and owner representative have a higher concentration of dark blue values, indicating a higher variance than the rest of the metrics. On the other hand, the data values, which have a higher red concentration near the central part of the figures, indicate that metric queries, which belong to the project engineer and the superintendent, have a higher similarity with the rest of the metrics.

Table 2.13 presents the statistics of the similarity measures such as minimum, maximum, median values and the standard deviation. For the overlap measure, while the median is 0.54, the standard deviation is 0.11. Hence, it can be concluded that while most of the metric queries show resemblance amongst fourteen variables (~half), the rest has significant similarities and differences.

Table 2.13: The statistics of similarity matrices

Values	Overlap Measure	Goodall Measure
Min	0.23	0.45
Max	1	18.97
Median	0.54	4.34
Standard Deviation	0.11	2.01

In terms of Goodall measure, it can be seen that the median and standard deviation fall under the dissimilarity end of the evaluation measures. Hence, it can be concluded that a significant percent of the metric queries have variations from one another.

In this section, PCA and similarity analyses are performed to explore the metric queries in terms of their similarities and differences. PCA analysis showed that 90% of the

variation in the dataset can be represented by four PCs, which are a linear combination of original twenty-six variables multiplied by different coefficients. The magnitudes of the coefficients are used to validate the variables that stand for the similarities. Additionally, similarity analyses using *Overlap* and *Goodall* measures are performed to explore the differences within the dataset. It is validated that regardless of the frequency of the values within variables, significant percent of metric queries differ amongst each other.

2.7 Conclusions

In this Chapter, I presented that there are numerous performance metrics to assess the progress of construction projects and the information visualization needs of different project team members vary regarding the metrics they use to monitor the projects. In order to identify the commonalities and differences in such needs, I have utilized the Contextual Design method. As a part of this method, I have explored the project performance visualization needs for fourteen project members from seven projects through semi-structured interviews.

132 metric queries were identified through fourteen interviews. Safety and quality were the least requested metrics, whereas communication of documents and cost were the mostly requested metrics. The total number of metrics chosen by each interviewee ranged from five to eighteen with an average of eight. The highest number of metrics were requested by an area manager with an experience of 18 years, and one of the superintendents with an experience of one year. Based on these findings, I could not find a direct correlation between the number of requested information with the organizational

hierarchy of the project team members in terms of job titles and the number of years of experience.

Analysis of the interview results revealed top-level categories of commonalities that emerged through modeling and consolidation. These commonalities include the way the information and visualization requirements are described using the following concepts: Performance Metric Categories, Concept of Interest (CoI), Level of Detail (LoD), Visual Representations, Intent, Purpose of Use, Notifications, and Frequency. While commonalities are identified through the interviews, each component has various enumerations based on the unique needs of the project members.

- Seven categories of performance metrics are identified as a result of the consolidation of the interviews. These are: (a) cost, (b) time, (c) productivity, (d) quality, (e) safety, (f) communication of information, and (g) quick access.
- Nine types of CoIs are identified, such as activity, components, documentation, indirect/ overhead, issues/ inspection, project, resource, space, and milestones. For each of the CoI, various levels of details are also identified.
- Regarding the 132 metric queries, 22 different visual representations were requested. Mostly requested visual representations are buttons/links for quick access to information, Gantt Chart for time, table for communication of documentation and spreadsheets for cost.
- The intents used in this study are clustering, comparison, composition, correlation, detail on demand, distribution, overview, overview and detail, and trend.

- The usage purpose of the metric queries are identified as one of the five types, such as descriptive, explanative, predictive, evaluative, and prescriptive.
- It is observed that 81 % metric queries were requested as interactive as opposed to static information. For 28% of the metrics, notifications were requested in the form flashing flag or highlighted information.
- Five levels of frequency for information visualization is identified as always, always with notifications, based on data being available, based on pre-determined time frame, and further graphs.

When the interview results are analyzed based on the types of projects, it is observed that the highest interest is on the exchange of information for commercial projects and productivity for highway projects. The main reason for this difference arises from the fact that majority of the works commercial and residential projects are subcontracted out, where in highway projects much less is subcontracted out and more is self-performed.

The validation of commonalities and differences are performed with the use of principal component analysis (PCA) and similarity analyses. The PCA analysis provided me with the most influential variables in the metric queries dataset and validated the common categories of variables within these queries. In order to explore the differences, similarity matrices are generated using Overlap and Goodall measures. With the use of these matrices, it is validated that a significant percent of the metric queries are different than one another.

From this study, I can conclude that in order to support the decisions made by the project team members and to prevent them from information overload they should be provided with individually relevant information to increase their understanding of the projects.

CHAPTER 3

3 Construction Project Performance Visualization: A Framework towards the Development of a Performance Visualization Environment

3.1 Introduction

Monitoring and assessing the performance of a project is claimed to be one of the most important tasks of construction project management (Memon et al. 2007). During performance monitoring of the projects, project members interact with large amounts of project related data. These large amounts of data in construction projects may result in information overload (Leung et al. 2003). However, recently many research studies have targeted visualization of information in numerous domains, which can minimize the effects of information overload. For instance, Korde et al. (2005) have highlighted the potential of visualization techniques in overcoming the challenges associated with making meaningful deductions from data, and with respect to the construction industry and Russell et al. (2009) have highlighted the role of interactive visual aids in the evaluation of project performance.

Generally, the term “visualization” is associated with visualization of either 3D/ 4D product models or pictures and videos of projects in construction industry. The product

models are being used for many purposes, such as design reviews, cost estimating, and analysis of time space conflicts (e.g., Arayici and Aouad 2005, Staub-French et al 2003, and Akinci et al. 2002). However, Hartman et al. (2008) have indicated that 3D/4D models are not widely used for analyzing performance. The published research efforts on visualization of project performance have focused mainly on using vision techniques on pictures, such as time-lapse movies (Abeid and Arditi 2003), and automated cameras (Bohn and Teizer 2010), or integrating product models with as-is photographs (Golpavar-Fard and Pena-Mora 2007).

Although such studies provide means for visualizing the project performance; the needed visualization is not always about 3D models and pictures. For example, evaluating the performance of a project by exploring and analyzing the data gathered through schedules, time cards, sensors or daily reports requires usage of visual representations (e.g. line charts, and scatter graphs) that are not necessarily 3D in nature. Based on the review of past research on construction performance visualization, I have observed that there is a need for numeric multi-dimensional data visualization. However, to date limited number of studies has been performed by the construction academic community on this topic (Chiu and Russell 2011).

I performed a set of initial studies/interviews to understand whether project members have access to visualizing project performance information according to their unique needs. Through fourteen interviews I identified that currently 60% of the project performance visualization requirements of project members were not met in a format

required by them. The processes required to provide needed format include adding alerts, adding more contextual information, changing the visual representation, changing the level of detail, and enabling interaction, such as filtering, sorting, abstraction and categorization. It has also been observed that project team members do not always pick an ideal form for data visualization that support the types of decisions that project members are trying to make. There are numerous research efforts that focus on information visualization (e.g., Lohse et al. 1994; Keim 2002; Shneiderman 1996), and study the effects of data types, and intents of users on the applicability and effectiveness of visual representations. However, some of these studies are semi-structured, and the breadth of the research on these topics has not been collectively studied and unified. Hence, there is a need for development of a synthesized and formally represented knowledge on effective visualization of project performance information.

The research on development of visualization framework described in this chapter has been driven by the vision of an environment effectively visualize the project performance data associated with monitoring of construction projects. By analyzing the interview results (i.e. metric queries) that I performed with different project members, I observed that their information visualization needs change based on various factors, such as their responsibilities, decisions to be made, experience levels, and characteristics of the project. These information visualization needs have commonalities and differences in terms of the performance metrics monitored, the level of detail of the metric queries, and the frequency and ways of visualizing these metric queries. In other words, in developing a construction project performance visualization framework, I built on the findings of

commonalities and differences in information visualization requirements of project members discussed in Chapter 2 and formalized them with the addition of the synthesized research efforts on applicability and effectiveness of visual representations. This chapter discusses the developed project performance visualization framework to define and represent the information visualization requirements of construction project members.

3.2 Case Studies and Needs

3.2.1 Availability of the Information in the Required Form

As discussed in Chapter 2, currently, the required project information is not always available to the project members. In relation to the fourteen interviews, the required information is only readily available for 25% of the requested metrics, and for the rest, either the information is not available or is spread across different data sources with the need of processing the data. Examples of data processing that need to be performed are filtering or aggregating information, changing the format of how the data is represented, and integrating data from different sources. For instance, during an internship at a construction job site, I observed that a project engineer was spending 5 % of his time on aggregating and visualizing the timecard information in a spreadsheet, which is the view that he prefers to track this information. Such wasted time can add up to significant amounts when incorporating all the metrics that he would like to track and when considering the fact that all other project team members have many different metrics that they would like to track as well.

As opposed to wasting time on customization of the information, some project members prefer to use the information as-is; during which they might end up losing some important insights from the data. For instance, during an interview, a project engineer emphasized the importance of tracking the performance of activities over a period, specifically the ones with 10% cost overrun. Moreover, every week, he identified the top five activities with the highest cost overruns, and their productivities to inform the foreman about their performance. The cost reports, estimated and actual productivities are retrieved from different sources and the project engineer did not spend extra time on combining these two interests to track the productivity of the foreman who had been causing continuous cost over-runs over a period of time. Using the data from that construction project, I identified that if such an analysis were to be done, it could have been identified that a single foreman was responsible for 40% of the cost overruns and hence corrective actions could have been taken in time. Since the project engineer did not have time to do the processing, he missed this important insight that could potentially have resulted in an early intervention and achievement of higher productivity rates.

3.2.2 Software tools designed for Construction Projects and Information Visualization

I have explored commercially available project management tools and research efforts to assess how well these tools and developed approaches support customizing, visualizing and interacting with project performance monitoring data/information. I investigated seven commonly used construction project management software systems and recent research efforts targeting visualization of construction project performance. The

following paragraphs provide a synopsis of my assessment of the existing commercial systems and previous research studies in this area.

With the advent of visualization techniques, information visualization toolkits and dashboards are becoming a corner stone for decision-makers to explore, and analyze data. Hence, most of the commonly-used commercially available construction project management and monitoring software systems provide dashboards to enable the users to visualize the progress and performance of the projects. Even though such dashboards present critical information in a graphical way, only a limited number of visual representations and interactions with the data are offered to the users and they do not necessarily align well with the different needs of the project team members. Hence, these tools serve a general audience, and lack necessary means and methods to generate customized views per individuals. Moreover, Chiu (2011) performed an extensive search into the Construction Management (CM) domain oriented visualization tools and concluded that: “Currently neither academia developed nor commercial CM information systems have provided a visualization environment that allows users to generate and interact with a wide variety of visual representations of CM data that support common and unique analytic tasks associated with the diversity of tasks/functions that comprise CM” (Chiu 2011).

Other than CM domain specific visualization, I explored the information visualization toolkits that are designed for a larger audience. There are several general-purpose toolkits that target supporting information visualization. The main functionalities of these toolkits

are to display an imported document by using a predefined set of visual representations, and provide interaction options to the users. These toolkits vary in terms of the types of input data that they can handle (e.g. structured, unstructured), visual representations that they output (e.g. scatter plots, trees) and interactions that they support, such as filtering, layout, applying graph primitive properties, searching, semantic and geometric zooming. Although these toolkits support a range of functionalities and visual representation options, they are still limited in supporting the types of visual representations that are accepted as natural idioms to construction management domain, such as Gantt Charts, and Building Information Models. Moreover, such toolkits do not incorporate a framework that defines which visual representations can plot what types of data, and which representations better support users' decision needs.

3.3 Research Objective and Scope

The objectives of the research presented in this chapter are: (a) To investigate and synthesize previous research studies related to information visualization, and (b) To develop a project performance visualization framework and identify its major components to support generation of customized information visualization needs of project team members.

The scope of this research is the information visualization and interaction needs of project members, and it does not focus on addressing how comprehensive and/or generic the project data needs to be represented and stored. The reason arises from the fact that

such issues are large-scale research problems within themselves and are beyond the scope of this research. Hence, it is assumed that the project data required to generate performance metrics is available, and a project data metadata that states the required information is to be developed within the scope of this research activity.

3.4 Background

The research presented in this chapter builds on previous studies done in the areas of construction project performance monitoring and visualization, and information visualization related frameworks. In Chapter 2, I have presented detailed background research on performance metrics, and visualization of construction project performance studies. The literature provided in this section focuses on the studies that formalized the requirements, processes, or context of visualization systems in a framework and the studies that help in formalizing the process.

3.4.1 Construction Project Performance Monitoring Frameworks developed for the AEC Domain

I have reviewed over 70 papers published over the last 20 years in relation to performance metrics for construction projects. Some research studies focused on defining performance measurement with the perspective of project success at the completion, and considered the satisfaction level of the clients (e.g., de Wit 1998; Chan and Chan 2004; Yang and Peng 2008). Other studies considered the performance during the construction phase with a focus on Project-level (e.g., Alarcon and Ashley 1996; Korde et al. 2005) or Company-level (e.g., Takim and Akintoye 2002; Cheah et al. 2004) performance. While

the project-level performance measurement mainly considers processes and operations performed on site, company-level performance measurement targets strategic management and financial aspects. Several of these studies suggested the use of or developed performance measurement framework for construction (e.g., Alarcon and Ashley 1996; Cheung et al. 2004). Three major frameworks that are identified are Key Performance Indicators, Balance Scorecards, and European Model of Quality Management Excellence (Yang et al. 2010):

1. Key Performance Indicators (KPIs) are the indicators that enable measurement of project and organizational performance throughout the construction industry (DETR 2000). This approach has been proposed and evaluated by researchers to identify the key characteristics of indicators and assess their application (e.g., Cox et al. 2003; Beatham et al. 2004; Bassioni et al. 2004; Chan and Chan 2004).
2. Balanced Scorecard (BSC) is a performance measurement system that incorporates four main measurement categories; such as financial, internal business process, customer, and learning and growth. This method is used within cross-domains; however, numerous researchers studied identification of potential sub-measures and their application for the construction industry (e.g., Kagioglu et al. 2001; Stewart and Mohamed 2001; Mohamed 2003; Yu et al. 2007).
3. European Foundation of Quality Management (EFQM) Excellence Model is designed to define and assess continuous improvements of organizations based on a set of predefined criteria (Beatham et al. 2004). Major research performed in this area is towards studying the implementation of EFQM by construction industry (e.g., Robinson et al 2001, Watson 2002; Robinson et al. 2005). However, Bassioni et al.

(2004), and Lin and Shen (2007) highlighted the fact these fixed criteria are not comprehensive and limit the flexibility of application to different processes.

A common feature of these frameworks is the fact that they incorporate some metrics against which a project's performance is measured and evaluated. When dealing with construction project performance monitoring, these frameworks appear to be well known and utilized methodologies (Bassioni et al. 2004) Nevertheless, while providing a vast array of quantitative indicators, these frameworks do not address the breadth of performance visualization and interaction needs of individual project members.

In addition to these performance frameworks, as presented in Section 3.2.3, several research studies have been performed focusing on performance visualization frameworks. These frameworks are discussed in the following section with the other research studies on information visualization frameworks.

3.4.2 Construction Project Performance Visualization Frameworks

In addition to the performance frameworks described in the previous section, a number of research studies explored and developed frameworks for the visualization of the project management and control data within the Construction Management domain. A list of these research studies is presented in Table 3.1. It can be observed that most of the existing frameworks focus on exploring the process of data acquisition to visualization mainly focusing on visualization in 3D. Indeed, Hays (2002) stated that the limited number of work performed on construction data visualization remains exploratory instead of structurally examining the guidelines, and information items. In other words, while

exploring data visualization for construction projects, the research studies remain limited in structurally formalizing the breadth and depth of components required to streamline the visualization of construction projects.

Table 3.1: Construction Project Visualization Frameworks

Framework For	Focus and Audience	Components	Reference
Project Management and Coordination	Components	Information views, interaction approaches, display layouts	Kuo et al. 2011
Web-based project control	Process	Data collection, variance analysis, work evaluation, trend analysis, display, corrective actions	Chou and Chong 2008
Multi D. visualization of project control data	Process	Data structure, user operations for LoD, data graphics editing, communicating and evaluating	Songer et al. 2004
Project Monitoring	Process for Steel fabrication	Project baseline, data acquisition, visual reports, performance evaluation, performance forecasting, corrective action, simulation	Azimi 2010
Project Management Performance Visualization	Process and Componets	Performance measure, context dimension, data status, data view, interact with data granularity, interact with data range	Chiu 2011
	Components	Integrated data model, data access interface, data manager, data viewer	Wu and Hsieh 2012

In Azimi (2010) and Wu (2012) the benefits of having a visualization framework for construction performance monitoring are discussed. The proposed methods are task-centered implementations; with abstract level discussions. The work by Chiu (2011) presents some common aspects to what we have identified through contextual inquiry, but they do not have a comprehensive approach to handle the diversity of factors that we have identified and to address the observed requirements.

3.4.3 Generic Information Visualization Frameworks

Several conceptual and taxonomical frameworks have been proposed (e.g., Pfitzner et al. 2000, Aigner et al. 2007) to represent information visualization and interaction processes. However, the proposed frameworks are not comprehensive and lack formalized

definitions, detailed enumerations or mapping of components that make up the framework. Moreover, they are limited in enabling multiple representations of information according to distinct perspectives of project members. They cannot be directly applied to construction domain since they do not cover the natural idioms (e.g., Gantt Chart, Building Information Models) and project data characteristics (e.g., point cloud data, different abstraction and details of information) used during construction projects

I identified nine frameworks that are developed for information visualization. Although the titles of these studies are labeled as information visualization or its synonyms (e.g., visual exploration and visual analytics); the focus and audience, the domain and components of these frameworks vary as shown in Table 3.2. These frameworks mostly target general-purpose information visualization, with two exceptions of specific focus on certain data types. When individually evaluated, these frameworks either provide higher-level categories of components with vague definitions, or detailed enumerations targeting only certain data types. Moreover, these frameworks do not address the needs of construction industry due to their limitations in terms of the range of data types and decisions considered, and their incomprehensive context. I refer to these frameworks as non-comprehensive since they do not reflect/contain numerous findings on visualization of information that are developed within the same domain.

Table 3.2: Existing Information Visualization Frameworks

(...Framework for...)	Focus and Audience	Components	Domain	Reference
Information visualization	Taxonomy of development of graphic representations for users	User skill level, visualization/display dimensions, interactivity, data types and task types	Informatics and Engineering	Pfitzner et al. 2000
Visual exploration	Internal process for programmers	Visualization Session, results, derivation, transform, parameter type, parameter value, result type, result value, operator, and value	Computer Science	Jankun-Kelly 2003
Information Visualization	Internal representation for programmers	Object, shape, figure, view renderer, and layout	Software and Informatics	Meyer et al. 2006
Visual analytics of time and time-oriented data	Design space for designers	Data structures, Visualization and interaction components, analytical/mining components	Information& Knowledge Engineering	Aigner et al. 2007
Analytical Reasoning Process	Reasoning Processes for users	Visual representations in data view, knowledge view and navigation view	Information and Knowledge Engineering	Shrinivasan and Wijk 2008

These existing frameworks will be used as benchmarks in order not to overlook a required component, and as a baseline to compare their applicability and effectiveness to construction project management requirements. The framework developed within this research study incorporates existing framework components and their enumerations to form a more detailed and structured body of knowledge in terms of the types and relationships of framework components.

3.4.4 Taxonomies and Findings on Visual Representations, Data types, and Interaction Tasks

The importance of selection of visual representations, as well as, the characteristics of and associations amongst data types, visual representations, users' purposes, interaction tasks, and visual primitives (e.g., color, shape) is a well discussed subject within the research community (Chengzhi et al. 2003; Chuah and Roth 2002). These existing studies on information visualization include formalizing:

- Data types according to their dimensionality, such as 1D, 2D, 3D, and structure, such as temporal, hierarchical, relational (e.g., Chengzhi et al. 2003; Keim 2002; Shneiderman 1996; Tidwell 2006; Fayyad et al. 2002);
- Categorization of visual representations based on their types, such as charts and diagrams, purpose of usage, such as concept visualization and strategy visualization, and tasks of users, such as comparison, and clustering (Chengzhi et al. 2003; Chin et al. 2009; Dull and Tegarden 1999; Jung 1999; Heer et al. 2009; Lohse et al. 1994; Kosara et al. 2003; Shneiderman 1996; Fayyad et al. 2002, Eppler and Burkhard 2007);
- Task taxonomy to interact with visual representations, such as categorize, distinguish, and rank (Chengzhi et al. 2003; Chuah and Roth 2002; Li and North 2003; Zhou and Feiner 1998; Pirolli and Card 2005; Shrinivasan and Wijk 2008; Yang et al. 2003);
- Visual primitives according to data characteristics, such as priority of usage of color and shape for quantitative and qualitative data (Cleveland and McGill 1987; Kasik et al. 2009; Elmqvist and Fekete 2010; Pfitszner et al. 2000).

Although the majority of the studies have different focuses on exploring information visualization, they are important studies to build up on in order to formalize the information visualization that can be tailored for construction project performance monitoring.

The frameworks and taxonomies that are developed within the HCI domain do not address the visualization requirements in the construction domain, as these frameworks

(1) target general-purpose information visualization, and are composed of vaguely defined components (e.g., visual representations are defined as a continuum from text to graphics rather than represented according to their types and characteristics), (2) have only specific focuses (such as visualization of spatial data, or interacting with time-series datasets), and (3) limited in addressing the natural idioms used in the construction domain, such as Gantt charts, design model, 4D, and bill of quantities. Hence, there is a need for synthesizing and formally representing the studies performed in HCI domain, and customizing them to apply in the construction domain, to enable identification of effective visualization of construction project performance data. In this chapter, I focused on identifying the components of a framework for enabling visualization needs of construction project members.

3.5 Research Method

In this section, I present the research method that I have used when developing a construction project performance monitoring framework. This method includes examining the previous and current research in the AEC and HCI domains, discussing a conceptual framework for performance monitoring environment, and proposing a framework for project performance monitoring related information visualization.

I have performed a set of interviews as a part of contextual inquiry as discussed in Chapter 2, and performed literature review as a part of my research method. In Figure 3.1, I have demonstrated the means to and steps of this research activity. Through

consolidation of the interview results and formalization of the previous research studies, I have developed a construction project performance visualization framework.

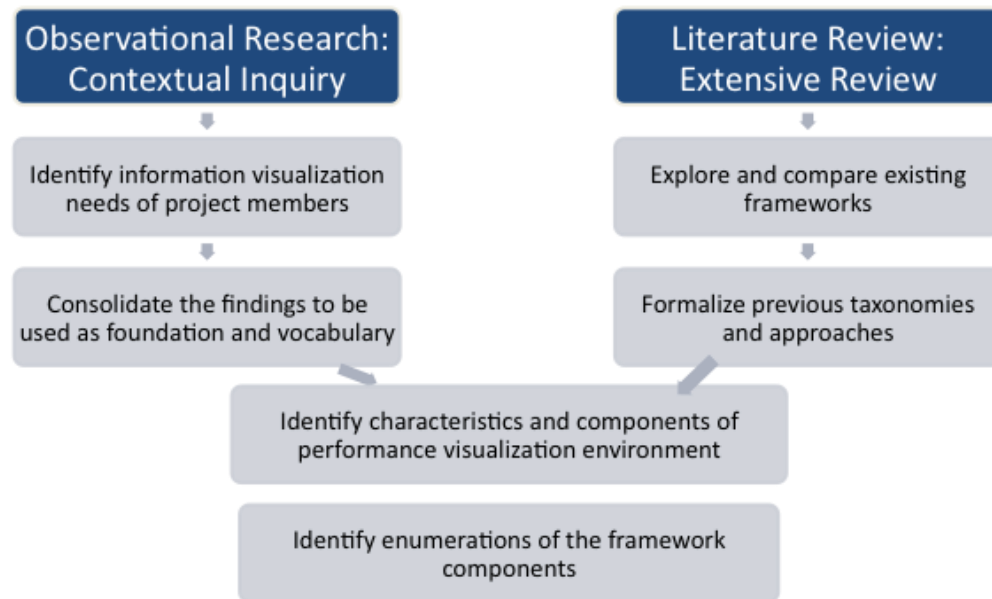


Figure 3.1: Research Method

The developed framework targets at capturing and representing the information items that are required to define, generate and visualize performance metrics for construction projects. My approach for monitoring project performance is through the correlation of information visualization techniques with the project performance metrics that are unique to project members' needs. This is a project member oriented approach; for each project member, the framework enables the identification of the visual representations that apply to the performance metric query. The objective of the framework is to support project members perform informed decisions about the progress of construction projects. The framework describes the information that needs to be considered for visualization of project performance information. As mentioned, though being limited, several existing

studies identified general frameworks, and/or information items to be considered when generating visual representations. Hence, the approach used for developing a project performance visualization framework was to synthesize the existing information visualization literature, develop a common language, identify the required framework components, their relationships, and enumerations, and represent these findings in a general framework.

3.6 Findings and Discussions

3.6.1 Overview of the Construction Project Performance Visualization Environment

Analysis of the interview results showed that information to be visualized should be tailored towards individual needs of project members, and when formalizing a project performance visualization environment, three aspects need to be considered as shown in Figure 3.2:

(1) The factors affecting the usage and context of the environment: These factors arise from the characteristics or needs of project members, such as experience and responsibility, and the characteristics of the projects that they are working for, such as scope and size. The components of the visualization environment might be subject to changes depending on project and technology related factors, such as the technologies, which are used to collect and monitor the project data.

(2) Components of the visualization environment: Exploring the commonalities and differences in information visualization needs of project members, I derived that the visualization environment should have components that serve for: a) The context of the information; the users should be able to visualize and interact with different categories and concepts of performance metric information and its levels of detail (e.g., activity versus project level information); b) Visual Representations; the ways in which to effectively represent the relevant information depend on the type of the data, and intent of the user. Hence, various visual representations and methods to display information and notifications should be enabled; (c) Interaction; the ways in which users should interact with visual representations. Indeed, I observed that for the 49% of the performance metrics, the project members requested to directly interact with data and visual representation ranging from sorting, filtering, projecting information and performing risk analysis. Hence, tasks to support visual analysis, including the alert adding capabilities, should be part of a visualization environment. d) Purpose; the factors that are related with the intent and decisions of the users help identify effective visual representations for the performance metric query at hand. Hence, these factors are an important component of a visualization environment.

(3) Principles for the visualization environment: There are publications and research efforts towards identification of principles that need to be satisfied (e.g., maintainability, security) in order to enable the users of a system (e.g., software, visualization environment) to achieve their goals (Barbacci et al. 1999). These principles to be

considered in a visualization environment can be up to twenty attributes. In addition, Svahnberg et al. (2002) state that new attributes need to be adopted as the requirements arise. In the light of the literature on principles and interviews, I identified several mandatory principles needed regarding both the quality of the data and the visualization environment. Such factors include; (1) accessibility: easy access to information, (2) accuracy: the correctness of the information; (3) dynamic interaction: supporting dynamic and real-time queries; (4) customization: generating and visualizing individually relevant information; (5) integration: data presented in an integrated form; (6) reliability: the reliability of the information provider; (7) timeliness: the time frame that the information is updated; and (8) usability: a lower learning curve and seamless communications when using the environment.

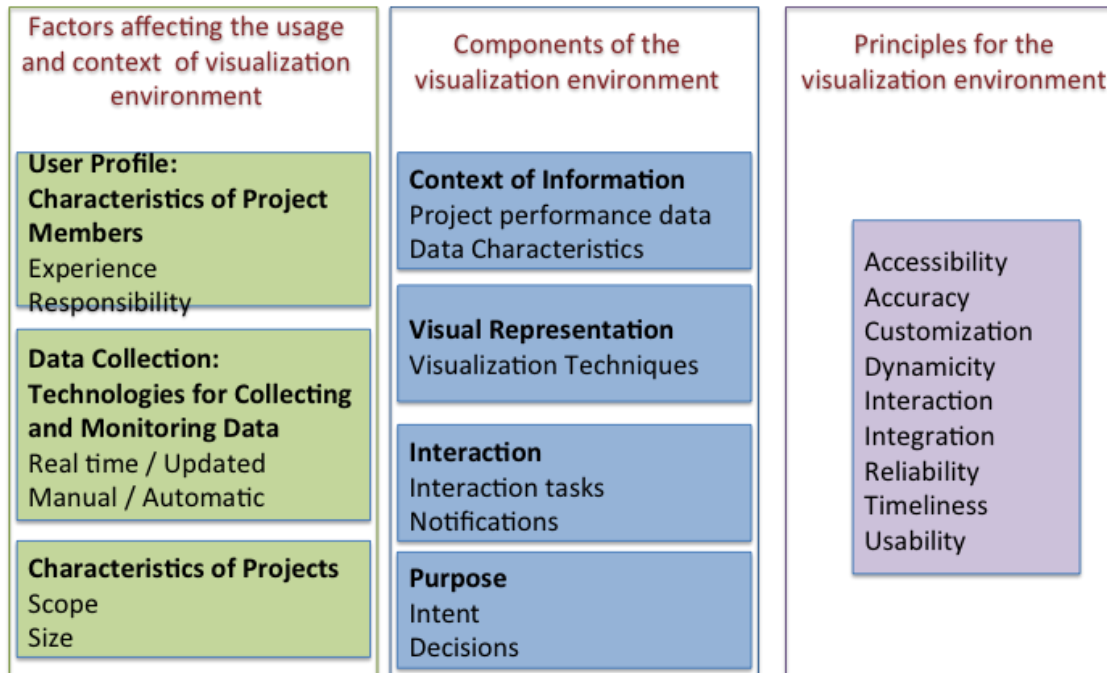


Figure 3.2: A Conceptual Model for Project Performance Visualization Environment

Although each aspect of the visualization environment is an indicator of its characteristics, the components of the environment will be the focus from this point on in this chapter. The factors and the principles of the environment are further research topics. Through consolidation of the initial fourteen interviews and building on the literature review on the requirements for information visualization, I have formalized a PROject PERformance VISUALization frameWORK, and named it as *ProPerVisualWork*.

3.6.2 Construction Project Performance Visualization Framework

A component is usually organized in hierarchies supporting different levels of data aggregation as well as multiple attributes. The proposed framework, called *ProPerVisualWork*, aids construction project members to explore the hierarchical nature of project performance data and visualize it through the use of the framework components. Visual representations and related constructs, such as data characteristics and users' intents, are key concepts/ components to establish appropriate representation for project performance monitoring.

In Chapter 2, I have discussed the process I followed in order to identify the required components for representing the requirements of project members. These requirements are depicted in Figure 3.3 in the form of an affinity diagram. The same figure was used in Chapter 2 as Figure 2.6, and reused here for the ease of referral. It can be noted that the diagram represents the components that are defined by the users. In order to enable the generation of the visual representations that are applicable to the metric query of the project member, the need for identification of data characteristics of the metric queries is

discussed in the background research section. The combination of these components makes up the framework developed in this research to provide a structured environment for exploration and visualization of construction project performance data.

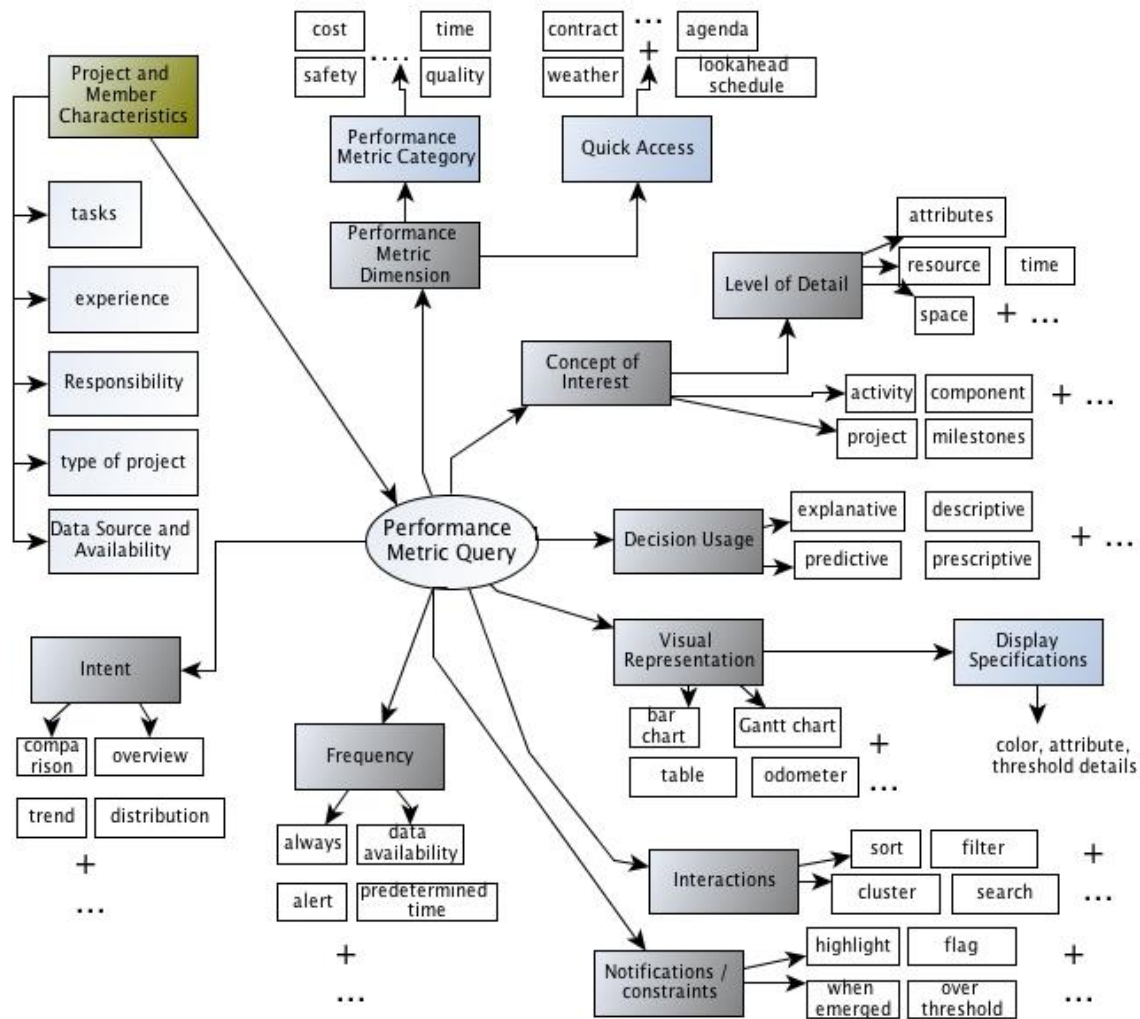


Figure 3.3: An affinity diagram depicting the requirements of project members for monitoring the performance of construction projects

By building up on the user studies and literature review, the UML representation of the *ProPerVisualWork* is developed and presented in Figure 3.4. The color coding is used to separate the components that are acquired through user input (purple) from the others.

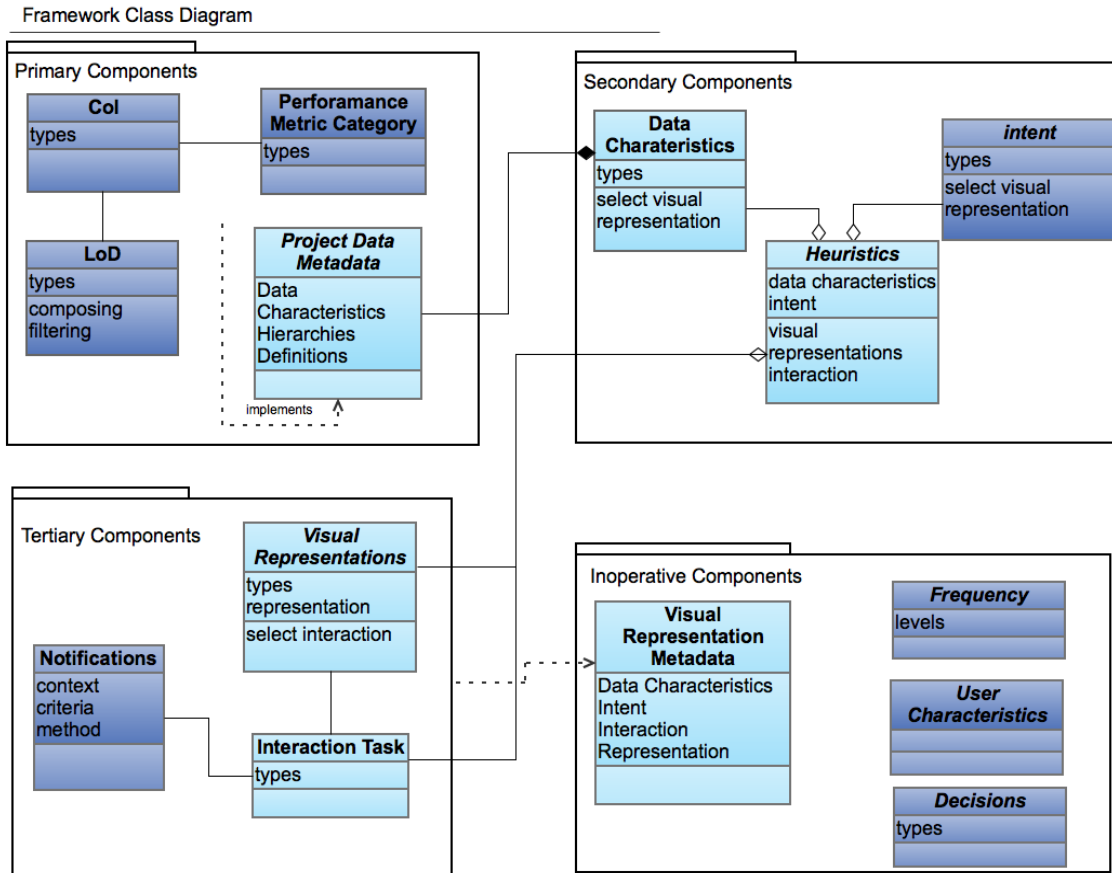


Figure 3.4: UML Class diagram representation of *ProPerVisualWork*

ProPerVisualWork is composed of fourteen components, categorized under four groups such as: (1) Primary Components that define the context of the project information used in the framework; (2) Secondary Components that define the context of the visualization information that is used in the framework; (3) Tertiary Components that define the displayed and interacted information; and (4) Inoperative Components that help characterize the framework; however not actively utilized in project performance visualization. In Table 3.3 the means of identification of each component along with short descriptions and examples are presented. These components are: (1) Decisions of

users, such as checking if a project is on track, and searching for reasons of problems, (2) Intents of users in visualizing the performance metric, such as comparison and trend analysis, (3) Performance metric categories, such as time, cost and quality, (4) CoIs, such as activity, material, and foremen, (5) LoD of information, such as a task, an activity and work group of activities, (6) Project data metadata to represent the information and relationships to be captured, such as LoD hierarchies and characteristics of data, (7) Frequency, the frequency of project members to visualize the performance information, (8) Notifications, the types and representations of the notifications that help acquire information push, (9) Visual representations used to present data, such as maps, charts and diagrams, (10) Data characteristics, such as categorical and quantitative, (11) Interaction tasks with the visual representations, such as emphasizing, filtering, and overlaying extra information, (12) Visual representations metadata, such as the graphical objects and space, and (13) *Heuristics* on visualization techniques, such as performance comparison of visual representations according to different data types, and (14) User profile, the characteristics of project members such as experience and responsibility. Each of the components is discussed in detail in the following paragraphs.

Table 3.3: Framework Components

Framework Components	Definition and Examples	Means of Identification	References
Concept of Interest	Facets of major subject of interest: e.g., Activity, component, actor	Interviews Literature	Jung and Woo 2004
Data characteristics	Categories of formats of data e.g., Types: Quantitative, Qualitative e.g., Structure: Spatial, temporal,..	Literature	Chengzhi et al. 2003; Fayyad et al. 2002; Keim 2002; Pradhan 2009;
Levels of Detail	Granularity of the requested information e.g., Project, building, floor, section e.g., Project, activity, task e.g., Component, foreman, equipment	Interviews, Literature	Aouad et al. 1995; IFC; Kang et al. 1998; Navon and Sacks 2007; Reinhardt et al 2005;
Performance Metric Categories	The categories of measures against which performance is measured e.g., Cost, time, quality, safety, scope	Interviews, Literature	Korde et al. 2005; Cheung et al. 2004; Cox et al. 2003;
Project Metadata	Data about data indicating how it is used e.g., Data formats: txt, image e.g., Data indexing: date, type, title, creator e.g., Elements & their relationships: RFI category, impact, action, element	Deduction, Literature	Chang et al. 2001; Hajjar and AbouRizk 2000; Leung et al. Lin et al. 2008; Mao et al. 2007
Decisions	Usage purpose of the performance metrics e.g., Descriptive, Explanative, Predictive, Prescriptive, Evaluative	Interviews, Literature	Liston et al. 2000
Interaction Tasks	User input to perform a single task e.g., Associate, derive, eliminate, emphasize, filter, identify, locate, rank..	Interviews, Literature	Amar et al. 2005; Zhou and Feiner 1998; Pike et al. 2009
Intents	The planned ways of analyzing the data e.g., Comparison, Correlation, Categorization, Trend Analysis	Literature	Chengzhi et al. 2003; Pfitzner et al. 2000; Schneiderman 1996
Notifications	Criteria to define prominent information e.g., setting threshold	Interviews	
Heuristics	Findings from prior work on relationship and characteristics of visualization techniques, e.g., representing data dimensions by position, length, area. Comparison of 2D and 3D line charts in terms of accuracy	Interviews, Literature	Cleveland and McGill 1987 Elmqvist and Fekete 2010; Kasik et al. 2009; Kumar and Benbasat 2004; Pfitzner et al. 2000
Visual Representations	Graphical expression of data e.g., Table, 2D-3D charts (e.g. bar chart), multiD charts (scatter plot matrix), Diagrams (static and process), Network Diagrams (pert),	Interviews, Literature	Chin et al. 2009; Chengzhi et al. 2003; Greinstein et al. 2001; Keim 2002, Shneiderman 1996
Visual Representations Metadata	Data about visual representations indicating its components and use e.g., Graphical Properties: color, size e.g., Graphical Space: timeline, coordinate system,	Literature	Aigner et al, 2007; Engelhardt 2002; Jung 1999; Wickham 2009, Wilkinson and Wills 2005;
Frequency	Frequency to visualize information e.g., Always, Predetermined time frame e.g., weekly, daily	Interviews, Literature	Navon and Sacks (2007)
User Profile	Characteristics of user and project e.g., experience, job title, project type	Interviews	

The first set of components of the *ProPerVisualWork* is Primary Components, which is depicted in more detail in Figure 3.4. Primary Components are used to acquire performance metric queries from the user with the required level of detail. The importance of these components lies in the fact that they help search and filter project database so that the unique performance metric queries are defined and related data is retrieved. Each of the components is explained as follows:

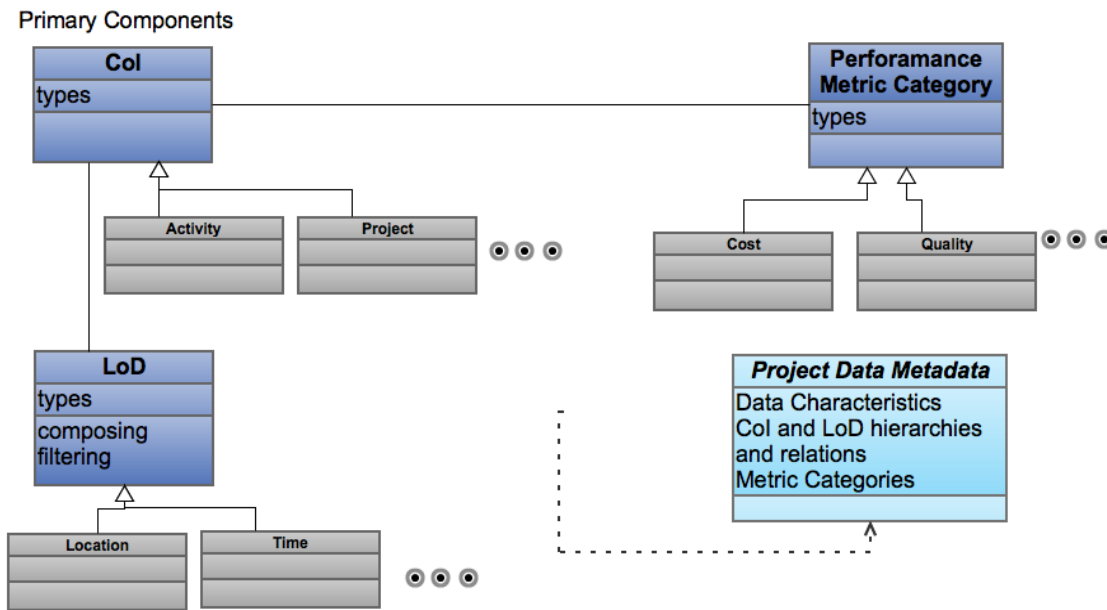


Figure 3.4: Primary *ProPerVisualWork* Components

(1) Performance Metric Category: Categories of measures against which a project's performance is measured. I incorporated the existing research efforts that studied performance metrics with the categories of performance metrics that I identified as a result of consolidating interview results. As described in Chapter 2, the major categories of performance metrics requested by the interviewees were cost (16%), communication

(28%), productivity (15%), quality (4%), time (11%), safety (5%) and quick access (21%).

(2) Concepts of interest (CoI): Facets of project specific concepts that a project member would like to learn more about. In the interviews, I observed that project members were interested in different CoIs, associated with metric queries. For instance, if a project member is interested in visualizing the daily productivity of a foreman (i.e., quantity of component built/time), the major concept of interest is foreman. On the other hand, if he is interested in visualizing the outstanding submittals categorized according to the subcontractors, the major concept of interest is submittals. Through the interviews discussed in Chapter 2, I have identified nine types of CoIs:

- Activity (e.g., types groups)
- Component (e.g., building elements, utilities)
- Documentation (e.g., RFIs, submittals)
- Indirect/Overhead (e.g, contingency, profit)
- Issues and Inspection (e.g., problems, tests)
- Project (e.g, multiple, single project)
- Resource (e.g., material, equipment, labor)
- Space (e.g., floor, work area, room)
- ToDos/Milestones (e.g., agenda items, milestones)

(3) Level of Detail (LoD): Hierarchies, decompositions and attributes of information items that project members desire to monitor. For instance if the CoI is an activity, the

LoD can be a single activity or a group of activities. I grouped the LoDs in two categories: composing LoD and filtering LoD. Composing LoD represents the data items that make up the requested information, and filtering LoD helps drill down to required portions of the requested information. If a metric query is defined as “productivity of all rebar-installation activities performed by foreman A”, in this query the CoI is activity. The granularity of the CoI is identified by the composing LoD and for this query, it is “a group of activities” that are related with rebar installation, such as unload and handle rebar, tie rebar, tie epoxy, etc. The filtering LoD is used to specify some criteria that will enable filtering out unnecessary information related with the composing LoD. For instance, in this specific query, filtering LoDs are the productivity as an attribute, and Foreman A, as an actor. Other filtering processes can be performed with the use of other CoIs, such as time, space and attribute (duration).

(4) Metadata about Project Data: To enable preservation of the linkages between project database and the visualization environment, metadata about project data is required. The metadata captures specific information about the structure of the project database, and the semantics of the project data it represents, indicating how it is used, and how the project related information and their relationships are represented, such as LoD hierarchies and characteristics of data. A project data metadata is developed within the scope of this research study, The metadata which captures information about the representation, relationship, composition, and characteristics of project related data.

Secondary Components of *ProPerVisualWork* is depicted in Figure 3.5. These components are the fundamental components that enable storage and reasoning of information visualization related *heuristics* identified in previous research studies to enable the generation of effective visual representations.

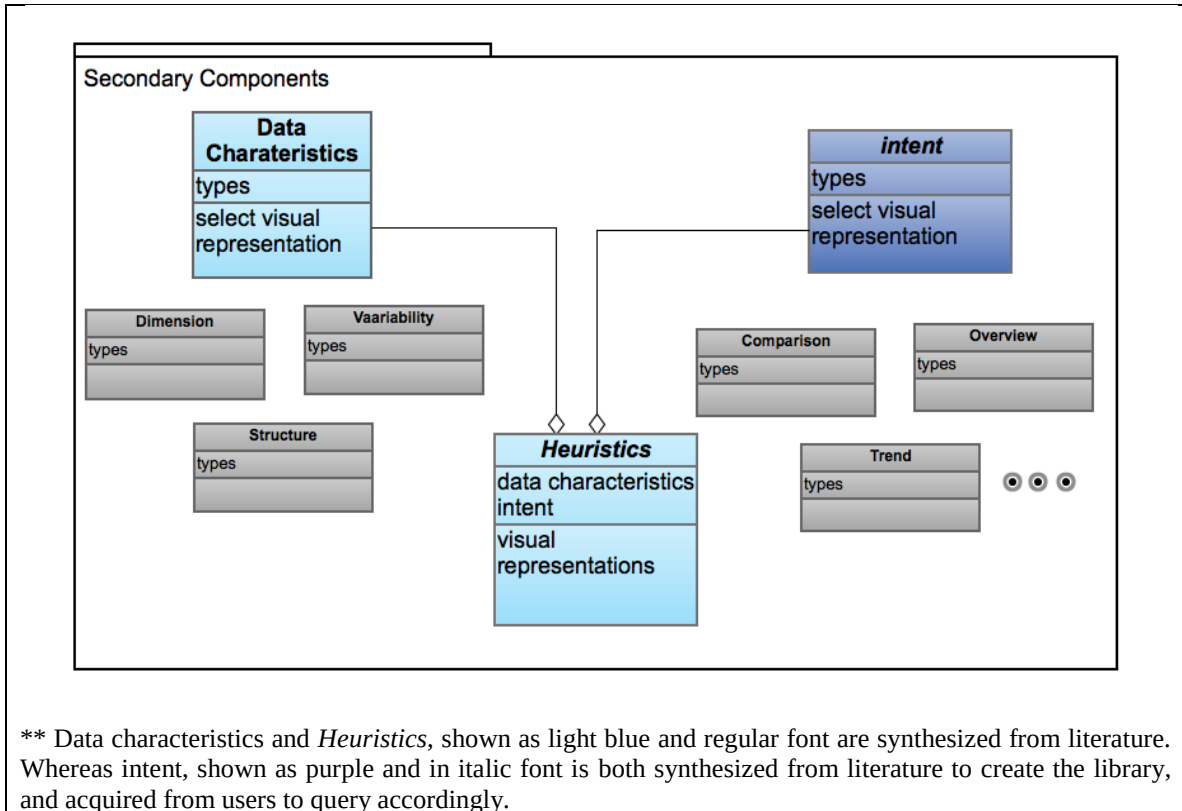


Figure 3.5: Secondary Components

(5) Data characteristics: Categories of data formats in terms of data dimensionality, structure, and variability type. Existing research studies present different number and categories of data types, and some examples are as shown in Figure 3.6a. In each of these studies, the data type enumerations lack certain data characteristics (e.g., spatial, categorical), and sometimes same data types are referred with different terms (e.g., quantitative and numerical). Instead of defining one group of data types, in the *ProPerVisualWork*, a common language is used and data types are categorized based on their characteristics into three groups as shown in Figure 3.6b. These groups are data

dimensionality (1D, 2D, 3D, and multiD), data structure (temporal, spatial, relational, causal, tabular and hierarchical) and variable type (quantitative, categorical, and text). The objective of such a categorization is to identify their common characteristics that would lead us to understand how data types and visual representations are related, and why certain interaction types would work for a certain group of visual representations.

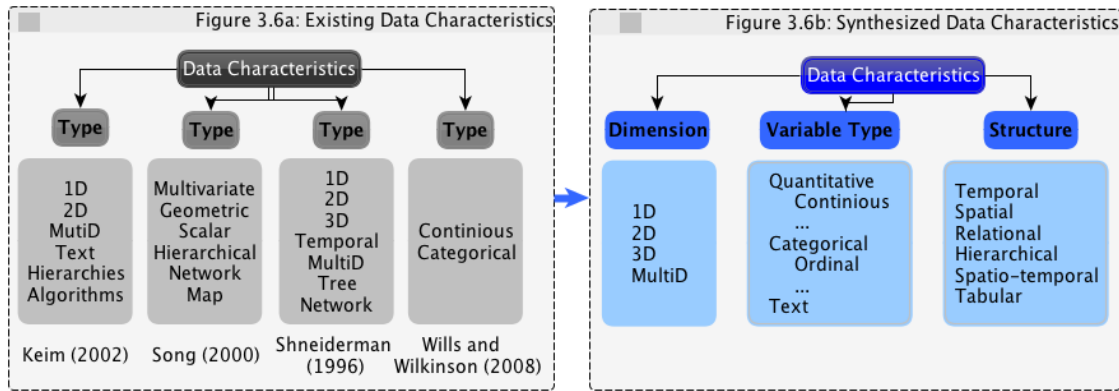


Figure 3.6: Data Characteristic defined in prior research studies and the data characteristics modeled in *ProPerVisualWork*

(6) Intents of Users: Planned ways of analyzing the data and visualizing the performance metric, such as comparison and trend analysis. For the selection of effective visual representations, identification of the intent of a user is as important as the identification of the data characteristics. Schneiderman (1996) identified seven types of intents to analyze data, such as overview, zoom, filter, details-on-demand, relate, history, and extract. However, while majority of the intents define the goal of the user, zoom and filter are interaction tasks rather than initial goals of the users. Considering this aspect, Chengzhi et al. (2003) identified another set of intents, such as overview-query, comparison, cluster-classification, distribution pattern, and dependency-correlation analysis. Building up on this research study, I formalized the intents of project members

as clustering, comparison, composition, correlation, detail on demand, distribution, overview, overview and detail, and trend. Each of these intents is described and discussed in more detail in Chapter 4.

(7) Heuristics: Findings from previous research studies about the relationship and characteristics of visualization techniques, their categorizations, and their performance comparisons. Mapping of visual representations to different categories of data characteristics and intents are represented as *heuristics* within the *ProPerVisualWork*. Other than these applicability mappings, some research efforts performed empirical studies in order to compare the effectiveness of visual representations for different data characteristics and user intents, and others defined a set of prioritizations for the use of visual primitives (color, shape, orientation) according to data characteristics (e.g., quantitative, ordinal) on visual representations. Hence, these *heuristics* are also captured within the *ProPerVisualWork*. These *heuristics* and their representations and usages are discussed in detail in Chapter 4.

Tertiary Components of *ProPerVisualWork* provide the output of the framework in terms of display and dynamicity such as visual representations, interaction, and notifications.

(8) Visual Representation: Collection of graphical symbols used to encode the properties and relationships of information, such as maps, charts and diagrams. In general, there are more than 200 visual representations that are used to communicate data. While some research efforts studied specific visual representations, others categorized them based on

their characteristics or usage area. For instance, Lohse et al. (1994) categorized visual representations under eleven groups as 2D/3D graphics, multiD graphics, maps, structure diagram, process diagram, cartographs, iconic displays, and networks based on some characteristics criteria such as abstract/concrete, temporal/non-temporal, etc. With a different perspective Lengler and Eppler (2007) categorized visual representations as data visualization, information visualization, concept visualization, strategy visualization, metaphor visualization and compound visualization. In the *ProPerVisualWork*, 163 visual representations are formalized with their characteristics.

(9) Interaction Task: User input to perform a task with the visual representations, such as emphasizing, filtering, and overlaying extra information. Amar et al. (2005) identified a set of ten analytical tasks that are independent of the visual representations. These tasks include retrieve value, filter, compute derived value, find extreme, sort, determine range, characterize distribution, find anomalies, cluster, and correlate. Wehrend and Lewis (1990) defined eleven tasks that are more closely related to the intents of the user. These tasks include identify, locate, distinguish, categorize, cluster, distribute, rank, compare within entities, compare between relations, associate, and correlate. The *ProPerVisualWork* incorporates the interaction tasks that are appropriate for the visual representations to be presented. Building up on these researches, I have identified the interaction tasks that the project members requested as: Alert, Cluster, Correlate, Extra Info, Filter, Highlight, Link, Navigate, Plan, Retrieve value, Sort, Search, and Simulate.

(10) Notifications: In the metric queries, project members stated that they would like to receive notifications on the displayed metric when it does not meet certain criteria (e.g., if an activity is behind schedule over a threshold value). I identified that the requested notifications differ in three major aspects:

- The context of the information: This aspect describes the facets of information items (CoI), the associated performance metric categories, and the time-frame when the source of information is available (or is of interest), such as past (e.g., cost overrun occurred last week), present (e.g., materials being delivered today), and future (e.g., an activity will start).
- The criteria to find prominent information refer to the conditions that identify the relevant information that the project member is looking for and have three types: a) Due/ overdue items define the assigned thresholds in terms of time frames (e.g., delay of 7 days in approval of Change Orders, or 1 week before an activity starts), (b) Information items over a value defines the assigned thresholds in terms of other numerical values (e.g., schedule overrun over 10%, or accidents more than 10), and (c) Existence of information items defines the problems that emerge expectedly or unexpectedly (e.g., if there is a quality related problem, or if there is space conflicts)
- The visual methods: They define the ways through which the notifications are communicated to the project members. Two methods that are prominently selected are the use of icons (e.g., flashing flag) and highlighting information within its context (e.g., color-coding according to magnitude of a problem).

The inoperative components of *ProPerVisualWork* are derived to help characterize the framework with the user and visualizations related factors that are not put in use to facilitate performance monitoring. Instead, these are the components that are identified as important part of an automated visualization environment that can be developed in future studies.

(11) Decision of Users: Use of the performance information, such as checking if project is on track, and searching for reasons of problems. I categorized the usage purpose of information as descriptive, explanative, predictive, evaluative and prescriptive. These five categories are used by O'Shaughnessy (1972) as types of decisions. Descriptive usage refers to observing the current status and needs. Explanative usage refers to corresponding with somebody or looking at detailed reports to learn about the reasons of the current situation. Prescriptive usage refers to notifying other team members of discrepancies. Predictive usage refers to performing planning and future projection of a given situation. Evaluative usage refers to the purpose of making an overall self-performance evaluation. The formalization of the decisions of the project members will help associate the usage purpose of the performance metrics with their intent. Intent refers to planned ways of analyzing the data such as performing comparison or trend analysis. These intents will lead to the selection of visual representations.

(12) Frequency: The frequency with which a project member would like to monitor a performance metric query. As discussed in Chapter 2, there are five levels of frequencies to monitor the project data: always, always with notifications, based on the data

availability, based on a pre-determined time frame and further/multiple graphs acquired through interaction with an existing representation.

(13) User Profiles: This is not a component effecting selection of visual representation; but a factor to store user related information including experience, type of projects, job title, and profiles of selection to be implemented in a bigger scale in the future.

(14) Metadata about Visual Representations: A common language of visual representations that enables configuration of visual representations based on the requested data characteristics, defines the associated interaction tasks, and enables implementation of new visual representations. Based on the previous studies, it has been deduced that metadata to be developed should capture attributes for defining the space, representing the visual primitives, graphical objects, data characteristics and interaction types. The details of the visual representations metadata is discussed in Chapter 5.

Except for the interaction and notifications components, all others are required for enabling the visualization of the metric queries requested by the fourteen interviewees. Formalizing the knowledge representation on information visualization would support project members to visualize and interact with appropriate and effective visual representations when analyzing datasets.

3.7 Validation

Validation of the developed framework involves evaluating its performance in two aspects: (1) its coverage in terms of the demographics of the interviewees, and literature, and (2) its usefulness in terms of the value of each of the framework components.

Coverage

The *ProPerVisualWork* framework is developed based on an integrated method of user studies and literature reviews. In this section, the coverage of the framework is validated. The scope of the research is limited to the construction project members residing in Pittsburgh/PA that work on projects within the range of \$15-100 million. While the diversity in the processes and methods of construction has not been addressed comprehensively, the types of the projects within the scope of this study include four to six storey commercial and educational buildings, multi-purpose highway projects including rehabilitation and construction and bridges, and dam projects. The diversity in the job titles, years of experience and types of projects are presented in Table 3.4.

Table 3.4: Demographics

Job Title	# of Participants	Experience (years)	Types of Projects
Area Manager	1	18	Highway
District Manager	1	30	Highway
Head of Estimation Department	1	30	Highway
Owner Representative	1	20	Building
Project Manager	1	15	Highway
Project Manager	2	9,25	Building
Project Engineer	3	4,8, 10	Highway
Project Engineer	2	2,8	Building
Superintendent	1	30	Highway
Superintendent	1	1	Building

Figure 3.7 demonstrates the distribution of the interviewee demographics, and as such scattered distribution validates the representativeness of the user tests for the demographics purposes.

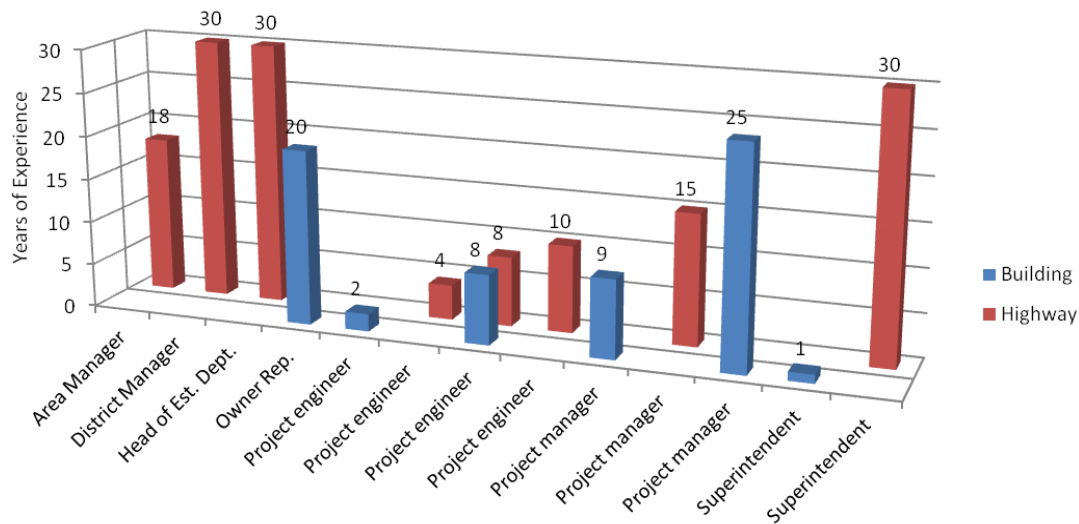


Figure 3.7: The sparse distribution of the contextual inquiry participants

In order to assess the coverage of the *ProPerVisualWork*, I have revisited the research studies that focused on project data visualization frameworks in Construction Management Domain, and information visualization frameworks in other domain. The list of the research studies and their focuses are listed as Table 3.5.

Each of the identified studies was revisited in order to validate the coverage of *ProPerVisualWork* by comparing the components and the enumerations of the frameworks. Table 3.6 presents the results of the analysis.

Table 3.5: Literature review of project data / information visualization frameworks

ID	Domain	Focus	Implement/ Not	Reference
1	CM	Coordination and management	Yes	Kuo et al 2011
2	CM	Monitor project control data	Yes	Songer et al 2004
3	CM	Monitor project data	No	Chuo and Chang 2008
4	CM	Project Performance Monitoring	Yes	Azimi 2010
5	CM	Visualize project control data	Yes	Chiu 2011
6	EM	Project Visualization	Yes	Wu and Hsiek 2012
7	IE	Information Visualization	No	Pfitzner et al. 2000
8	CS	Visual Exploration	No	Jankun-Kelly 2003
9	SI	Information Visualization	Yes	Meyer et al. 2006
10	IE	Visual Analytics	Yes	Aigner et al. 2007
11	IE	Analytical Reasoning	Yes	Shrinivasan and Wijk 2008

CM: Construction Management | EM: Engineering Management | CS: Computer Science
SI: Software and Informatics | IE: Engineering Informatics

Table 3.6: Coverage of the framework based on previous research studies

Components	1. ID of the prior work listed in Table 3.5										
	1	2	3	4	5	6	7	8	9	10	11
Metric											
Category	No	P.	P.	P.	Yes	P.	No	No	No	No	No
CoI	No	No	No	No	P.	No	No	No	No	No	No
LoD	No	E.	No	No	P.	No	No	No	No	P.	P.
Intent	No	E.	No	No	P.	No	E.	No	No	Y.	No
Frequency	No	No	No	No	No	No	No	No	No	No	No
Visual Rep	P.	P.	E.	P.	P.	P.	E.	P.	P.	P.	P.
Data Char	P.	P.	No	P.	P.	P.	Yes	P.	E.	P.	P.
Interaction	P.	No	No	P.	Yes	P.	P.	P.	P.	P.	P.
Decision	no	No	E.	No	No	Yes	No	No	No	No	No
Heuristic	No	P.	No	No	No	No	No	No	No	No	No
Metadata	No	No	No	No	yes	No	Yes	Yes	No	Yes	No
Notification	No	No	No	Yes	No	No	No	No	No	No	No
User Char	No	No	No	No	No	No	Yes	No	No	No	No

P.: Partial, E.: Exemplary

In this table, four types of labels are used: (1) No, meaning that the research study did not mention the associated component, (2) Exemplary (E), meaning that either the component name is mentioned as necessary, but no implementation/ enumeration is presented, or, a single example is introduced, (3) Partial (P), meaning a few enumerations of the associated component is presented; however no information is given on whether

the set is complete or what the other values might be, (4): Yes, meaning that a complete enumeration of the component is presented in the associated research study.

There are also a number of components identified in these research studies, but has not been defined in my framework. These components are of two types. First group of categories are related with the information visualization process, such as “view renderer” component of Meyer (2006), “visualization sessions” and “transformation” components of Jankun-Kelly (2003), and “data collection” component of Chou and Chang (2008). As can be seen, these components are mainly related with programming or data acquisition aspect of the visual representations. Since the scope of *ProPerVisualWork* is the representation of the project performance data and its visualization, the programming and data acquisition perspectives are not included. The second group of categories that has not been defined in my framework includes components, such as project baseline and simulation by Azimi (2010). These components are captured as enumerations of other components in *ProPerVisualWork*, such as project baseline under LoD attribute and simulation under visual representations. With the exceptions stated above, the coverage of the framework is validated regarding the project performance and information visualization frameworks developed in literature.

Usefulness

As introduced in the introduction section of this thesis, the representation and reasoning of the *heuristics* are discussed under Chapter 4. With the use of the *heuristics*, the usability of the framework in terms of generating effective visual representations based

on the performance metric queries of project members is tested and validated through test cases. In this section the usability and value of each of the components are discussed.

Table 3.7 presents the list of the components, and their validation status.

Table 3.7: Usability of framework components

	Components	Validated/ not	Validation by
Primary	Performance Metric Category	Yes	Synthetic Test, Usefulness
	Concept of Interest	Yes	Synthetic Test, Usefulness
	Level of Detail	Yes	Synthetic Test, Usefulness
	Project Data Metadata	Yes	Synthetic Test, Usefulness
Secondary	Data Characteristics	Yes	Literature Coverage, Usefulness
	Intent	Yes	Literature Coverage, Usefulness
	Heuristics	Yes	Literature Coverage, Usefulness
Tertiary	Visual Representations	Yes	Chapter 4, Coverage
	Interaction	Yes	Survey, Usability
	Notifications	Yes	Survey, Usability
Inoperative	Decision	No	Future work
	Frequency	No	Future work
	Visual Rep. metadata	No	Future work
	User Characteristics	No	Future work

The Primary Components are utilized to acquire and refine the context of the performance metric queries of the project members (i.e., metric categories, CoI, LoD) and map existing data structure to the framework (i.e., project data metadata). As will be discussed in detail in Chapter 4, I performed synthetic tests to identify effective visual representations for each of the 132 metric queries identified in Chapter 2. During the synthetic tests, I used the primary framework components to identify the context of the required information to derive associated data characteristics. Since each of these components help define the required metric category, in the absence of any of these

components, the required information cannot be identified with the required LoD. As a result, the framework would not be usable at all without these components due to failure in identification of information to visualize.

The Secondary Components are utilized to reflect the enumerations and relationship of the components identified and consolidated from research studies performed on information visualization. These components are the fundamental elements of the framework to visualize project performance data effectively. A total of 163 visual representations are captured in the framework. Not knowing which visual representations can better describe the data at hand, one might not be able to effectively visualize the project performance data. Given a metric query, without the use of data characteristics and intent, randomly selecting the most effective visual representation out of a pool of 163 visual representations is very unlikely as compared to making informed selections. In other words, with the use of the previous research findings on the interrelationship of visual representations with data characteristics and intent, informed selection of effective visual representations can be performed.

Figure 3.8 demonstrates the number of visual representations captured in the framework. The Venn diagrams stand for the data characteristics and intent identified from *heuristics*, and each number on the diagram represent how many visual representations are mapped to individual or combination of these components. For instance, blue “Data Dim.” component shows that in literature it is used as a total of 148 times to identify effective visual representations; amongst which it is used 28 times individually (e.g., scatter plot

used for multiD), 48 times in combination with intent (bubble graph used for overview of 3D dataset), and 9 times in combination with data variability and data structure (e.g., line chart used for 2D quantitative temporal dataset). Hence, with the use of the body of knowledge identified in literature and captured within the framework as Secondary components, the identification of effective visual representations is enabled. It can be stated that with the use of secondary components, the precision and recall in selecting visual representations are improved; thus validating the usability of these components.

Venn Diagram of Visual Representations defined by Heuristic Components

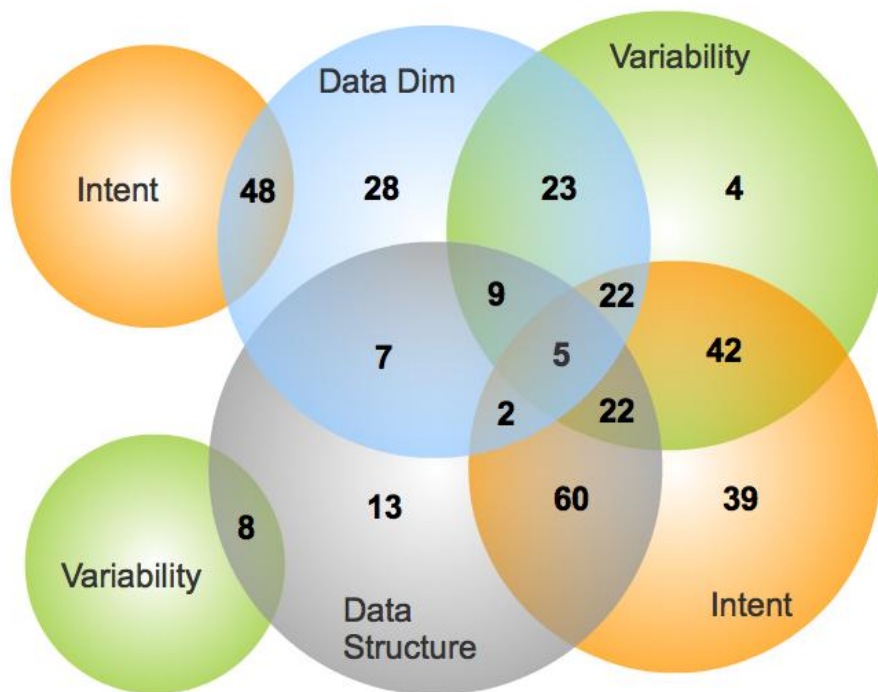


Figure 3.8: Venn Diagram of Visual Representations

The Tertiary Components are the output of the framework. The usability of the visual representations is covered in Chapter 4. As a part of Chapter 4 validation, I have

performed surveys with thirteen project members in the form of survey. The survey contained question about comparison of different visual representation for a given scenario and follow-up question about whether the project members would requested these visual representation enhanced with notifications and interactions. The results are presented in Table 3.8.

Table 3.8: Component Usefulness

Component	Not Useful	Do not know	For some of the queries	For all of the queries
Notifications	1 /13	0	9/13	3/13
Interaction	0	3/13	8/13	2/13

It can be observed that the notifications and interactions components are rated to be useful for more than 80% of the cases.

As described in the findings section, the Inoperative Components are represented within the framework for future purposes. Indeed, during the interviews and literature it was observed that these components are a part of visualization and decision making with project performance data; however since they neither change the display outcome nor the relevant interaction, their usefulness cannot be validated.

3.8 Conclusion

The advent of information visualization techniques has changed the way construction project performance is monitored. This growth in visualization has shifted project monitoring from spreadsheets to dashboards and 3D. In this chapter, I presented a framework, named *ProPerVisualWork*, which describes project members' information

visualization requirements to monitor the performance of construction projects. *ProPerVisualWork* is developed through the integration of contextual inquiry method and consolidation of literature review to meet the performance metric visualization needs of project members.

The framework is composed of fourteen components each having a primary, secondary, tertiary or inoperative function. These components are; (1) Primary Components: concept of Interest, Level of Detail, and Project Data Metadata, (2) Secondary Components: Data Characteristics, Intents, and *Heuristics*, (3) Tertiary Components: Visual Representations, Interaction, and Notifications, and (4) Inoperative Components: Decision, Frequency, Visual Representations Metadata, and User Characteristics. The coverage of the framework and the usefulness of each of the components are validated through a combination of literature review, synthetic tests, and surveys.

CHAPTER 4

4 A Formalism for Effective Visualization of Construction Project Performance through Individually Relevant Information

4.1 Introduction

Construction project performance visualization needs of project members are explored and described in detail in Chapter 2. It is discussed that a wide variety of performance metrics and their visualization is an issue to be addressed in order to facilitate project members to effectively monitor the performance of construction projects.

There are several general-purpose toolkits that target supporting information visualization. The main functionalities of these toolkits are to display an imported document by using a predefined set of visual representations, and provide interaction options to the users. These toolkits differ amongst each other in terms of the types of input data that they can handle (e.g. structured, unstructured), visual representations that they output (e.g. scatter plots, trees) and interactions that they support, such as filtering, layout, applying graph primitive properties, searching, semantic, and geometric zooming. In order to evaluate whether these toolkits address the performance visualization needs of project members, I explored the toolkits that are either designed for generic purposes or construction management (CM) domain, and can be openly accessed. The explored

toolkits include XMDV (xmdv, 2010), Prefuse (Prefuse, 2007), Piccolo (Piccolo, 2005), InfoVis (InfoVis, 2005) for generic purposes, and Constructware (Constructware, 2011), Buzzsaw (Buzzsaw, 2010), Primavera (Primavera, 2009), MS Project (MS Project, 2008), and Vico (Vico, 2010) from CM domain.

Although generic toolkits support a range of functionalities and visual representation options, they are still limited in supporting the types of visual representations that are accepted as natural idioms to construction management domain, such as Gantt Charts, and Building Information Models. Most of the toolkits focus on a small aspect of visualization, such as multidimensional data visualization and quantitative visualization and there is not a single tool available that can support all data types used in construction or the wide variety of visual representations in general.

The tools used in Construction Management domain mainly focus on planning, scheduling and estimating purposes. Therefore there is no toolkit that is solely designed for construction performance monitoring and enables the use of effective visual representations that are selected based on individual performance needs of project members. The existing tools do not support a variety of inputs, visual representations and interactions that are needed for performance monitoring. Most importantly, such toolkits do not incorporate a framework or guidelines that defines which visual representations can plot what types of data, and which representations better support users' decision needs. The existing toolkits provide a platform to visualize information; however, such a visualization does not meet the needs of project members in the ways that a) not all data

types are supported, b) not all intents are supported, and c) the most effective ways of visualizing a the queried performance metric are not supported.

In Chapter 3, a project performance visualization framework named as *ProPerVisualWork* (PROject PERformance VISualization framework), which was developed in this research, is discussed to address this need. Chapter 3 identified the major components of *ProPerVisualWork* with the purpose of finding effective visual representations that are suitable for the queried performance metric. One of the components of the *ProPerVisualWork* is identified as *heuristics*. *Heuristics* component captures findings from previous research studies that describe the appropriateness of visual representations for various types of information and user intent. Indeed, several researchers studied and pointed out the potential advantages of information visualization within the Human Computer Interaction (HCI) domain. These studies target exploring types and relationships amongst data characteristics, visual representations and interaction intents, as well as effectiveness of visual representations.

In these studies, researchers identified, categorized and mapped visual representations (e.g., Keim 2002; Hoffman and Grinstein 2002) to data types (e.g. Shneiderman 1996; Myatt and Johnson 2009), and users' interaction needs (e.g., Amar et al. 2005; Zhou and Feiner 1998), and they evaluated the effect of visual representations in augmenting human cognition (Card 2008). These studies differ in terms of their approaches and focuses. For instance, the pursued approaches include defining taxonomies (e.g., Shneiderman 1996), categorizing visual representations (e.g., Lohse et al. 1994),

comparing performance of visual representations (e.g., Kumar and Benbasat 2004), and mapping interaction tasks to specific visualization techniques (e.g., Chengzhi et al. 2003). Besides, some of the studies focused on visualization of specific datasets, or interaction types. For instance, while Aigner et al. (2007) explored time-oriented datasets, Chin et al. (2009) explored dynamic data-streams. Such studies provide guidelines on the kinds of visualization techniques to be selected for certain datasets, interactions or intents. However, this body of knowledge exists in separate sources that do not always use similar terminology or explicit implementation rules. In other words, such a wide variety of studies and findings have not been synthesized and formally represented. In order to make use of such a body of knowledge towards effective information visualization, there is a need for synthesizing and formalizing these *heuristics*. Stemming from this need, this Chapter consolidates, represents and reasons with such *heuristics* in order to identify effective visual representations for construction project performance monitoring.

4.2 Objective and Scope

The objectives of the research presented in this chapter are: (a) To investigate and synthesize the information visualization related studies in literature that would facilitate project performance visualization, (b) To represent the consolidated findings in a usable format and (c) To reason with such findings in order to generate effective visual representations to support performance visualization needs of project team members.

The scope of the proposed research is limited to the existing research studies on information visualization in both HCI and CM domains.

4.3 Background

The research presented in this chapter builds on the previous studies performed in the areas of information visualization for CM and HCI domains.

4.3.1 Research Studies on performance visualization studies in CM domain

The research studies performed on project performance visualization in CM domain mainly focus on 3D/ 4D product visualization (Building Information Models), such as 3D product models (e.g., Songer et al. 2000, Srivastav et al. 2009), and 4D models (e.g., Golparvar-Fard et al. 2009, Rohn et al. 2010). The main focus of these studies is generally the visualization of the building components associated with schedule for planning and management purposes. However, aside from scientific visualization in 3D/4D, construction performance monitoring is also related with information visualization, such as visualization of quantitative or categorical performance data. Some of the existing studies on information visualization in CM domain combine the information visualization concerns of civil engineering with the principles of HCI. Goldparvar-Fard and Pena-Mora (2007) stated that thirty to forty different visual methods are currently being used to represent construction information throughout the lifecycle of a project. Given that there are nearly 200 available visual representations, one can deduce the need for exploration of the possible usages of other visual representations within the construction management domain.

The research studies in CM on information visualization provide insights on the use of visual representations for a variety of purposes, such as scheduling and planning. For instance, regarding scheduling Russell et al. (2002) explored the use of a number of visual representations, such as bar graph and linear planning charts. For planning, Liston (2000) explored interaction techniques for Gantt Charts. Zeb et al. (2008) compared the strengths and weaknesses of a few visual representations and HCI features like zooming and brushing for as planned and as-built comparison of project schedules. Songer et al. (2004) evaluated and compared four graphics; including scatter plot, linked histogram, hierarchical tree and treemap, to represent structured cost control data. Although limited in number, such studies provide foundational work in identifying the information visualization needs within the construction domain.

4.3.2 Research Studies on Information Visualization

As described in Chapter 3, the importance of selection of visual representations, as well as the characteristics of and associations amongst data types, visual representations, users' purposes, and interaction tasks are well discussed subjects within the information visualization research community (Chengzhi et al. 2003; Chuah and Roth 2002). These research studies vary from generic taxonomies of visualization and interaction techniques, to detailed exploration of specific visual representations. A common finding amongst prior work is that the selection of effective visual representations for a given dataset is mainly dependent on the data characteristics (categorical, hierarchical), and the intent in the analysis (comparison, trend). By reviewing more than 150 research studies on visualization of information performed over the last 20 years, I have identified the ones that studied how data characteristics (structure, variability, dimension) and intents

are mapped to visual representations. For instance, some of the research studies explored visual representations based on the tasks of users, such as comparison, and clustering (Chengzhi et al. 2003; Chin et al. 2009; Dull and Tegarden 1999; Jung 1999; Heer et al. 2009; Lohse et al. 1994; Kosara et al. 2003; Eppler and Burkhard 2007); and some based on data characteristics (Jung 1999, Lohse et al. 1994, Grinstein et al 2001, Shneiderman 1996; Fayyad et al. 2002). With the existence of such a variety of research studies, there is a need for integrating these *heuristics* through developing a common language and representing them in a format to be used to streamline the selection of visual representations. However, this representation is challenging, because the majority of the *heuristics* are semi-structured, incomplete, ill-defined, or contradicting. In Figure 4.1, some examples of *heuristics* are presented to explain such characteristics.

Data dimension	Variable type	Data size	User Intent	Analysis Intent	Visual Representation	Reference	
MultiD	ND	ND	Identify correlation	Applicability	Scatterplot Matrix Dimensional Stacking	Chengzhi et al. 2003	Semi-structured
MultiD	ND	ND	ND	Applicability	2D/3D Scatter Plot, Parallel Coordinates	Schneiderman 1996	Contradicting
ND	Quantitative	ND	ND	Applicability	Scatter Graph	Mackinlay et al 2007	Ill-defined
ND	Scalar	ND	ND	Applicability	Scatter Graph	Song 2000	
MultiD	Quantitative	9-25 data points	Read value, compare, trend	Comparison in terms of time to complete task	2D line chart outperforms 2D line chart	Kumar and Benbasat 2004	Contradicting
MultiD	Quantitative	ND	Project	Comparison in terms of accuracy and time to complete task	3D line chart provides more accurate but slower insights than 2D line chart	Dull and Tegarden	

Figure 4.1: Examples of the *heuristics* on visual representations, where problems are identified with ellipses and boxes, and ND refers to not-defined

As can be seen in Figure 4.1, a set of existing *heuristics* are listed based on the studied variables, such as data characteristics, users' and analysis intent, and the visual representations. Quass et. al (1995) defined semi-structured data as data that has no absolute schema fixed in advance, and whose structure may be irregular or incomplete. Figure 4.1 shows that the *heuristics* are semi-structured because the level of analysis in terms of the studied variables that affect the selection of visual representations is different. For instance, the majority of the studies did not consider all aspects that contribute to the selection of a visual representation; these are marked as “ND”, in other words not defined in Figure 4.1. Additionally, while some of the studies only explored the availability of visual representations, some others assessed their performance using different criteria, such as time and accuracy to complete a task.

Another characteristic of the existing *heuristics* is that some of the research studies are incomplete. For instance, the first row in Figure 4.1 shows one of the results of the study performed by Chengzhi et al. (2003). In that study, some visual representations are categorized based on data types and user intents, and Scatterplot Matrix and Dimension Stacking were selected to identify correlation for a multidimensional dataset. On the other hand, Atasoy and Akinci (2010) used parallel coordinates on a multidimensional dataset in their study to identify the correlation of weather conditions with productivity of earthwork activities. Hence, it is observed that the results presented by Chengzhi et al. (2003) are incomplete in terms of not considering the parallel coordinates in their categorization of the visual representations for multidimensional datasets.

The second row in Figure 4.1 shows that Shneiderman (1996) identified 2D/3D scatter plots and parallel coordinates as applicable for multidimensional datasets. Since the user intent is not discussed in their study, it can be inferred that parallel coordinates can be used for any user intent. Whereas, Chengzhi et al. (2003) identified that parallel coordinates cannot be used for distribution pattern. Hence, some of the *heuristics* are contradicting. Finally, in order to define data types and visual representations, different terminologies are used in different studies. For instance, “scatter plot”, and “scatter graph” are used to refer to the same visual representations, as well as “scalar” and “quantitative” are used to refer to the same data type. Hence, the ill-defined characteristic of the *heuristics* reveals the need for a common language.

The semi-structured characteristic of the *heuristics* lead to assortments of visual representations, which can be categorized as applicable to certain data characteristics and user intents. Additionally, the effectiveness of these visual representations is compared for different collections of data characteristics and user intents as shown in the last two rows of Figure 4.1. Hence, the selection of visual representations requires comparison as well as categorization of such representations for a given context.

4.3.3 Formalisms Towards Generation of Visual Representations

I have found only five research studies that described formalisms for the generation of applicable and/or effective visual representations. The major limitations of these studies are that they are only data centric and/or they discuss a general process of consideration without detailing with each required step.

Mackinlay et al. (2007) had 7 predefined rules according to data types that are used for the selection of visual representations within a set of 12 predefined ones. The two data types used are categorical and numerical. Based on different combinations of these data fields, the rule sets indicate the visual representations to be used. This approach is neither generic nor extensible, since it depends on a defined set of rules for only two data characteristics. Moreover, the approach does not consider the intent and task of the users.

Similar to Mackinlay et al.'s approach, the approach developed by Wills and Wilkinson (2010) is data-centric and provides multiple views of the datasets based on the types of the data fields. In their approach, a wider variety of inputs are considered; however, with the selection of the limited number of visual representations the purpose is to only provide first glance at data before analysis. Focusing on hierarchical datasets, Stolte et al. (2003) formalized the description of multi-scale visualization of data cubes with a set of data and visual abstraction techniques. The main purpose of the paper is introducing four general patterns of performing multi-scale zooming. The collection of patterns are considered as simple but not comprehensive.

The other two previously developed formalisms are both data and task centric and are generated by Helfman and Goldberg (2007, 2010). Task centric approaches consider the types of tasks; such as identification, comparison, and comparing; and associate applicable visual representations with these tasks. Helfman and Goldberg (2007) presented their approach for limiting the types and number of data characteristics, tasks,

and visual representations. Although the selection criteria and visual representation options are discussed through an example of three-variable dataset with different data characteristics; how these guidelines are derived and how they can be applied/extended to other datasets have not been formalized in their study. In Helfman and Goldberg (2010), the authors formalized the steps towards selection of applicable visual representations by providing a decision framework, composed of data, task and visual representation attributes formalized in their 2007 publication. The steps and guidelines towards selection of visual representations are demonstrated for specific task types; however, the complete set of framework components, and selection criteria of visual representations are not discussed.

The formalization of information visualization is a very recent topic with a limited number of publications that are neither comprehensive, nor extensible. Moreover, the goals of these research studies have been streamlining the selection of applicable visualizations, using a limited number of rule sets and data characteristics, and not the selection of effective visual representations. However, these developed formalisms and guidelines provide building stones towards selection of visual representations.

In conclusion, currently neither commercially available tools, nor individual research studies fully support the visualization of construction performance according to the needs of construction project members. Existing sources are limited in terms of the breath of the visual representations, and guidelines in terms of selecting the effective ones. However, when collectively evaluated, such resources provide a rich pool of information on

visualization of information. Stemming from such need, this research described in this chapter has been performed to formally and collectively represent and reason with existing research studies on information visualization.

4.4 Elicitation and Representation of Heuristics

This section describes the research method that I have used in eliciting the information visualization related *heuristics* from prior research studies, and the findings on the representation of elicited *heuristics* in order to find effective visual representations for a given performance metric query.

4.4.1 Elicitation of Heuristics

An extensive literature review is performed in order to identify *heuristics* that can be used in selection of effective visual representations. Over 150 papers that are published over the last 20 years on information visualization are studied. The elicitation of the *heuristics* defines identification of the research studies that explored effectiveness and appropriateness of visual representations for various variables such as characteristics of data, user's intent, and interaction options. More than 600 *heuristics* are identified with the use of six major variables:

- Data dimension (D_dim): The dimensionality of the data is identified according to the number of data fields that are analyzed. Its enumeration includes 1D, 2D, 3D, and multi D: (1) 1D: One-dimensional data has one major dimension in other words one attribute field is assigned to a data series such as over-cost values for multiple projects, (2) 2D: Two-dimensional data has two major dimensions such as geographical maps

(x and y values) and actual vs. budgeted cost values for activity codes, (3) 3D: Three-dimensional data has three distinct dimensions such as 3D BIM models, and cost and due date of RFIs (Request for Information), and (4) multiD: Multi-dimensional data has more than three distinct attributes such as 4D BIM models, and cost, due date, status and responsible parties of RFIs.

- Data variability (D_var): The variability of the data defines whether the dataset is composed of quantitative, categorical, text or a mixture of these variables. The enumerations of data variability are hierarchical as shown in Figure 4.2. The major categories are identified as quantitative, categorical and text. Each of which is self-explanatory. The enumerations of these major categories are included if they have been studied by previous research studies. Hence, these sublevel enumerations might be overlapping and incomplete. Quantitative data is composed of sub-categories of: (1) continuous data that can be measured as a continuum on a scale, (2) ratio/ percent data that defines the relative values of variables and (3) non-negative data that is positive in value. Categorical data is composed of three sub-categories such as (1) Ordinal defines a categorical variable that has a clear ordering of variables such as likert scale (2) Nominal defines numbers that are used as a name for categorical variables such as activity codes, and (3) Frequency that defines the count of occurrence of a categorical data.
- Data Structure (D_str): The structure of the data defines the interrelationship of datasets, in other words how they are represented. Six data structures are used in this study including (1) Temporal data that has occurrences explicitly associated with time frames such as timecards, (2) Spatial data that has occurrences associated with

locations such as punch list items, (3) Hierarchical data that can be represented in a tree-like structure such as cost breakdown, (4) Relational data that represents the interrelationship and connections of its components such as foremen responsible from activities, (5) Spatio-temporal data that has occurrences that are associated with both time and location, and (6) Tabular data that can be represented in a tabular format, such as monthly summary cost report.

- **Intent of User (Intent)** The intent of the user refers to the type of analysis that a user performs on a dataset. Nine types of intents are used in this study (1) Comparison: examination of two or more items to identify similarities and differences, (2) Cluster: examination of two or more items to identify group of similarities, (3) Correlation: examination of two or more items to identify interdependency between items, (4) Composition: examination of two or more items to identify their share in a whole, (5) Distribution: examination of two or more items to identify how something is shared amongst them, (6) Trend: examination of two or more things to identify how they change over a period of time, (7) Detail: examination things to describe them item by item, such as visualizing a specific number or statistics, (8) Overview: general review and examination of things in summary, and (9) Overview and Detail: overseeing a dataset in summary and with certain details highlighted like top three items, and over the threshold.
- **Visual Representations:** “A collection of graphical symbols that visually encode causal, functional, structural, and semantic properties and relationships of a represented world” (Sedig and Sumner 2006, pg 2). This term captures tables, texts,

graphs, maps, diagrams, icons and networks. In this study a total of 163 visual representations are used. The whole list is presented in Appendix D.

- Interaction: User input to perform a task with the visual representations, such as emphasizing, filtering, and overlaying extra information. All of the interaction types that are identified as associated with specific visual representations are captured. For instance, Harris (2000) identified that arrange, rank and sort are the interaction techniques that can be applied to tables, and similarly, Liston et al. (2000) identified that overlay is a technique that can be applied to Gantt Charts and BIM models.

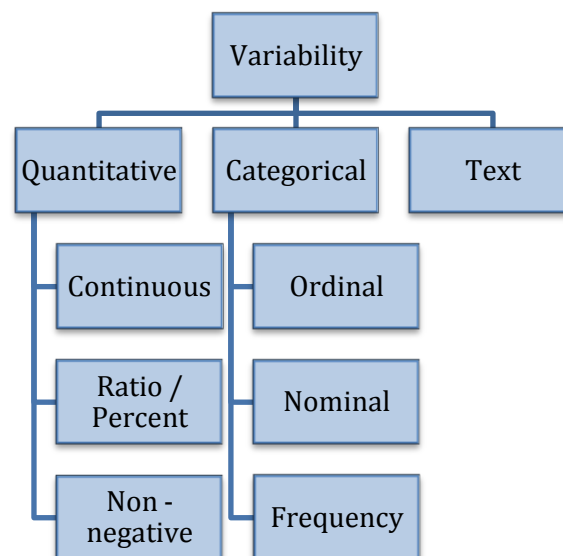


Figure 4.2: Data Variability

The previous research studies that explored visual representations according to one or more of these characteristics are recorded as *heuristics*. However, during the elicitation of the *heuristics*, a large number of *heuristics* (~100) are eliminated from the findings due to five major factors such as:

- Absence of visual representations: The relationship of data characteristics, intents and interaction tasks are discussed without tying the discussions to any specific visual representations. For instance Zhou (1999) stated that clustering can be performed for multiD datasets; however a visual representation is not associated with that discussion. Since the goal of this research task is to identify effective visualizations, the studies that do not provide any means of visual representations are not included in this study.
- Out of scope visual representations: The focus of studies is not information visualization, but scientific visualization (pixel oriented), strategy visualization, or metaphor visualization. For instance Lengler and Eppler (2007) presented that funnel, iceberg and bridge shapes can be used for metaphorical visualizations. Since these visualization forms are not a part of performance visualization, they are not included in this study.
- Generic representation modes: Instead of a specific visual representation, representation modes (the arrangement of visuals based on whether they are pixel-oriented, icon-based, or coordinates-based) that can be applicable to multiple visual representations are discussed. For instance, Jung (1999) stated that multiple y plots can be used for multiD datasets. However, various charts including column charts, line charts, and scatterplots can be represented as multi-y plots. Since a single chart is not studied in such *heuristics*, they are eliminated. .
- Algorithm based visualization methods: Visualization methods that help decrease the degree of clutter such as dimensional stacking and multidimensional scaling can be applied to multiple visual representations. Additionally one method can be implemented through different algorithms. Hence they are not associated with specific

visual representation. For instance Bratke (2005) discussed that multiD scaling can be applied to multiD datasets, and Keim (2002) discussed that dimensional stacking can be applied for hierarchical datasets.

- Distortion-based tool-specific visual representations: Some research studies explored tools that are used to change the view of visual representations as a method. For instance, Chengzhi et al (2003) presented that fisheye view can be used to get an overview of a dataset, and hyperbolic view is presented as applicable to hierarchical datasets. However, these tools can be used in different forms and on different visual representations; hence they also are eliminated from the *heuristics* pool.

With the elimination of the inapplicable *heuristics*, a total of 490 *heuristics* that are identified from 35 previous research studies are used in the *heuristics* pool. These 35 studies are presented in Table 4.1. The whole list of the *heuristics* is presented in Appendix D and discussed in section 4.4.2.

Table 4.1: Research studies used in *heuristics* identification

Type of Research	References
Research Paper for Formalisms in HCI	Schreck et al. 2010, Shrinivasan and van Wijk 2008, Icke and Sklar 2009, Mackinlay et al. 2007, Stolte et al. 2003, Stolte et al. 2002
Research Paper for Frameworks in HCI	Aigner et al. 2007, Chengzhi et al. 2003, Keim 2002, Abela 2006
Research Paper on Visual Representations in HCI	Johansson 2008, Kumar and Benbasat 2004, Lengler and Eppler 2007, Lohse et al. 1994, Tominski et al. 2004, Trickett and Trafton 2006, Weber et al. 2001, Bartke 2005, Thakur and Rhyne 2009, Grinstein et al. 2001
Books	White 1984, Chambers et al 1983, Fry 2007, Wills 2012, Harris 2000, Spence 2007, Young et al. 2006, Tidwell 2005
Research Papers for Visualization in CM	Azimi et al. 2011, Russell and Norvig 2002, Huang et al. 2009, Liston et al. 2000, Kuo et al. 2009, Chiu 2011, Hays 2002

The distribution of the number of *heuristics* acquired from research studies is presented in Figure 4.3. The papers are ordered based on the year of publication in x-axis, with the numbers identified in y-axis. The earliest research study that provided findings to be used in the *heuristics* pool dated as early as 1980s and the rate of the research increased after 2000. Although there are a few research studies that that provided a large number of *heuristics* as outliers in the chart, one to twenty *heuristics* are acquired from the research studies

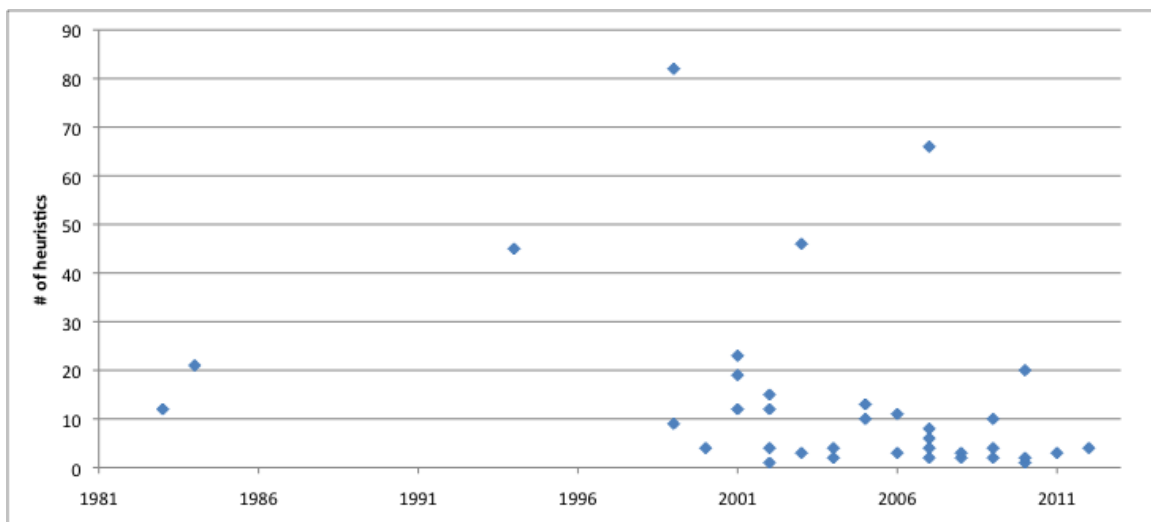


Figure 4.3: The distribution of publications describing the *heuristics* included in this study based on their year of publication

4.4.2 Representation of Heuristics

As discussed in section 4.3.2, the semi-structured characteristic of the *heuristics* lead to assortments of visual representations, which can be categorized as applicable to certain data characteristics and user intents. Additionally, the effectiveness of these visual representations is compared in relation to different collections of data characteristics and user intents. Hence, the selection of visual representations requires comparison as well as

categorization of such representations for a given context. Therefore, the *heuristics* should be represented in a format that provides simplicity in traceability. Moreover, information visualization is still a developing research area with possible new findings on the applicability and effectiveness of visual representations to be defined in the future. The identified *heuristics* are consolidated and represented in a table format to enable (1) tracing the existing terminologies and *heuristics* by direct observation and data filters, (2) extending the *heuristics* by adding new instances of *heuristics* as rows (3) modifying the *heuristics* by changing the existing terminologies, or by defining new variables as columns.

As discussed in Chapter 3, studying the research studies on information visualization, *ProPerVisualWork* is developed by consolidating such studies into components, their subcategories and enumerations. The major components that are required to enable selection (secondary framework components) of and interaction with effective visual representations are identified as data structure (D_str), data dimension (D_dim), data variability (D_var), intent, visual representations, and interaction. Hence, the variables (column headings) of the *heuristics* table are determined based on these components. The detailed enumerations of each of these variables are presented in the previous section.

Building up on these components, three more variables are derived to be used in the *heuristics* table, such as lower D_var, lower intent and higher visual representations: (1) Lower D_var defines the subcategories of categorical and quantitative variables such as nominal, ordinal, frequency, continuous, ratio, and non-negative. Since very limited

number of research studies used these subcategory assignments to D_var, the higher categories are used during reasoning with *heuristics*; however these lower categories are still captured to be used in the future; (2) Lower intent defines the subcategories of nine types of intents. There was no common terminology used amongst previous research studies to refer to intents of users. For instance, while some studies used “distribution” as intent, “relative distribution”, “absolute distribution”, and “absolute distribution” are also used. Hence, as a future reference, the list of the detailed intents are captured within the *heuristics* table. However, for the simplicity of the search and retrieval of *heuristics*, the higher level categories are used for the reasoning. Similarly, picture data, show facts, depict stages and illustrate are categorized as overview, (3) Higher Visual representations define a categorization for the visual representations that belong to same origin with mild alterations. For instance, bar chart, stacked bar chart, and aligned bar charts are all counted as different visual representations within the 163 visual representations; however, they all belong to bar chart class. The visual representations table presented in Appendix D includes the higher-level categorization of these representations although they are not used as a part of the formalism.

The *heuristics* identified through previous research studies were semi structured (ill-defined) in terms of the terminology used for defining the *heuristics* when collectively explored. Therefore, a common language was required to be developed. For instance, the terms step chart and histogram (White 1984 and Chengzhi et al. 2003) were used interchangeably referring to the same chart type, similarly the terms qualitative and categorical (Icke and Sklar 2009) were used synonymously to refer to the same D_var.

Hence, when two different terminologies are used to define the same concept, one of the concepts is selected in random to be used with the table, and the alternate terminologies are recorded for future reference.

As a result, 490 *heuristics* are identified as presented in Appendix D. A demonstrative set of *heuristics* is presented in Table 4.2 in order to explain the concepts described in this section. Table 4.2 presents the *heuristics* represented with nine column headings, where each row represents a single *heuristic* identified from a research study. For instance first row shows that 3D bar graph can be used for a 3D temporal dataset. The other variables, such as intent and interaction, are not identified/described in that particular *heuristic*; hence, they are depicted with not applicable “N/A” label in the table. This *heuristic* also shows that 3D bar graph is categorized under bar graphs in a higher category. Though not shown in Table 4.2, the first and second rows are identified from the same source, revealing that both 3D line chart and 3D bar graph can be used for the same D_dim and D_str. (3D temporal).

Table 4.2: Demonstrative set of *Heuristics*

D_dim	Lower D_var	D_Var	Lower Intent	Intent	D_str	Visual Rep.	Higher Visual Rep.	Interaction
3D	N/A	N/A	N/A	N/A	temporal	3D bar graph	bar graph	N/A
3D	N/A	N/A	N/A	N/A	temporal	3D line	line graph	N/A
3D	ratio and nominal	Quant and cat	N/A	N/A	N/A	2D line graphs	line graph	N/A
N/A	Quant	Quant	compare	compare	N/A	aligned bar graph	bar graph	N/A
multiD	Quant	Quant	correlate	correlate	N/A	scatterplot	Scatter plot	zoom in and out
Quant: quantitative Cat: Categorical								

The third row shows that 2D line charts can be used for a dataset composed of a mixture of categorical and quantitative data. It also gives an example for lower D_var column where ratio and nominal datasets are categorized as “categorical and quantitative”. Fourth row depicts that aligned bar charts can be used for comparison of quantitative datasets. It can be seen that aligned bar chart is a subcategory of bar charts. Fifth row depicts that scatter plots can be used for correlating multiD quantitative datasets. Additionally, as opposed to previous examples, this heuristic presented the zooming as an interaction technique for a scatter plot.

The reasoning process uses the major four components of the *heuristics* table for the identification of effective visual representations. These components are D_var, D_str, D_dim, intent. Using these four variables, a total of 163 visual representations can be identified within the *heuristics* table with 490 *heuristics*. As explained, most of the *heuristics* are incomplete, in other words not all variables are assigned, but one or more “N/A” fields exist. The number of enumerations of each of the variables and the number of mapped visual representations (rows without the “N/A” field for that variable) is presented in Table 4.3 to summarize the totality of the *heuristics*.

Table 4.3: The number of enumerations of *heuristics* table major variables and the number of *heuristics* that capture a value for these variables

Component	# of enumerations	# of <i>heuristics</i>
D_dim	4	244
D_var	3	211
D_str	6	197
Intent	8	343
Interactions	12	84

4.5 Reasoning with the Heuristics

Until this point, this chapter focused on the characteristics and representation of the information visualization related *heuristics*. It is stated that the effective visual representations for a given dataset is identified through data characteristics and the intent of the user. Reasoning with the *heuristics* is the final task of the thesis that has the overarching purpose of enabling project members to effectively monitor the performance of construction process according to their unique needs. Before detailing how to use *heuristics* for the identification of the effective visual representations, I will explain the process of converting project members' performance metric queries into heuristic variables.

In Chapter 2 commonalities and differences in performance metric queries of project members are identified through contextual inquiries. An example of a metric query is “comparison of actual vs. budgeted cost of activities performed last week on a scatter graph”. Through consolidation of contextual inquiry results combined with research studies on information visualization, *ProPerVisualWork* is developed in Chapter 3. The initial step of reasoning with *heuristics* is assigning the *ProPerVisualWork* components with relevant information breakdowns of the metric queries. Figure 4.4 present the process of reasoning with *heuristics* starting with assigning the *ProPerVisualWork* components.

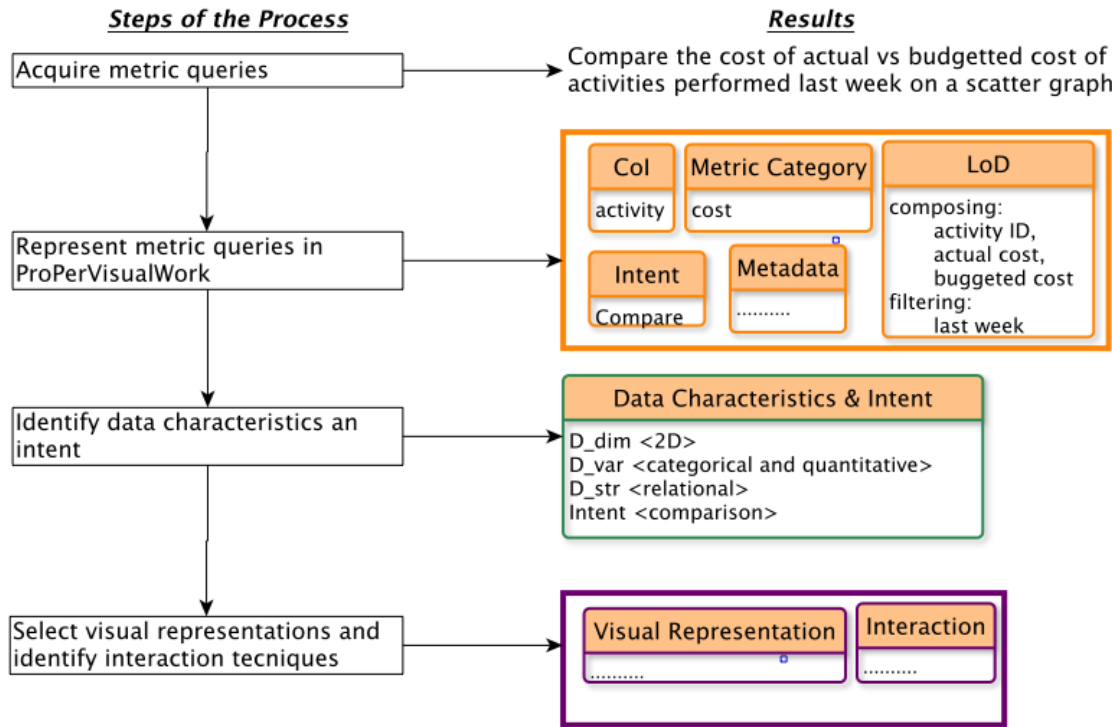


Figure 4.4: A Summary of steps of the reasoning mechanism

The CoI, intent, and metric category of metric query is directly identified as activity, comparison, and cost. The identification of the LoD is performed in two steps as composing LoD and filtering LoD. The composing LoD defines the information items that are related with the selected CoI and performance metric category. For instance, if a project member states the performance metric category as “cost”, and CoI as “activity”, the composing LoD may represents the information items that make up the cost of the activity including labor cost, material cost, equipment cost, may represent their totals or each individually, may represent the actual cost values and/or budgeted cost and etc. In our case, the composing LoD represents the total actual and budgeted cost of each activity. On the other hand, as shown in Figure 4.5, the use of the filtering LoD enables scoping down to a specific focus with the use of the other CoIs. For instance, if the

project member is interested in cost of activities that are performed on a certain floor by a specific subcontractor, then the CoIs used to scope down to the required detail are space and actor. In the case of our example, the filtering LoD is time and scopes down to activities performed last week. The use of filtering and composing LoD concepts help define the granularity of the requested information, so that data characteristics can be accurately identified.

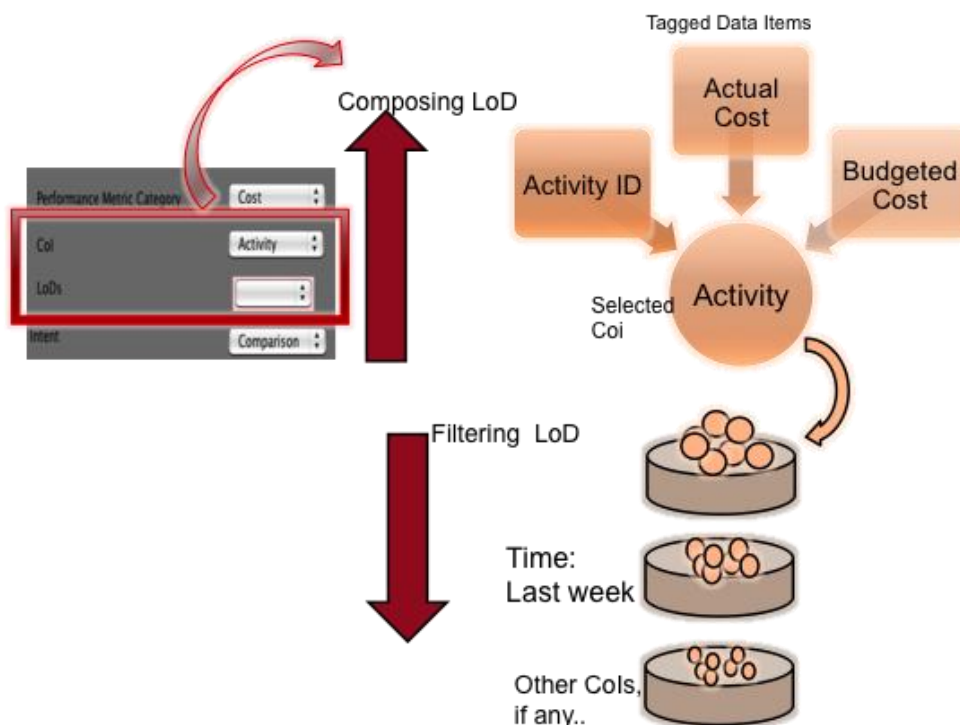


Figure 4.5: Composing and Filtering LoDs

To preserve a link between a project database and the *ProPerVisualWork* a project data metadata is required. The project data metadata captures information about the representation, relationship, composition, and characteristics of project related data that are used and reasoned within the *ProPerVisualWork*. In order to enable navigating

through project data via composing and filtering LoDs, the representational requirements of the project data metadata is presented in Figure 4.6. Two major components of the project data metadata are project data tags, and CoI & LoD hierarchies. First, CoI and LoD hierarchies are defined to identify the required information to the exact granularity required. The CoIs, and LoDs are relationally represented, hence they are interrelated. For instance, the CoI of resource has the enumerations of material and equipment, and each material is associated with the activities table (the CoI of activity), making it possible to state for which activity materials are used. Second, project related data is tagged with certain labels including performance metric categories, CoIs, data characteristics, and attributes. Since a project member performs a query via specifying the performance metric category and CoI s/he is interested in, tagging the project data with the related CoIs and performance metric categories would enable filtering and retrieval of the related data, when the process is automated. Similarly, the data items are tagged with attributes, such as actual and budgeted values (of cost, schedule, etc.), location, and format. The last requirement of the project data metadata is tagging each data item with their characteristics. In other words, each data item in the project database is assigned with their characteristics, based on dimensionality, variability and structure. For instance, given the example query in Figure 4.4, the actual cost of an activity has the characteristics of dimensionality <1D>, variability <quantitative>, and structure <hierarchical, relational>.

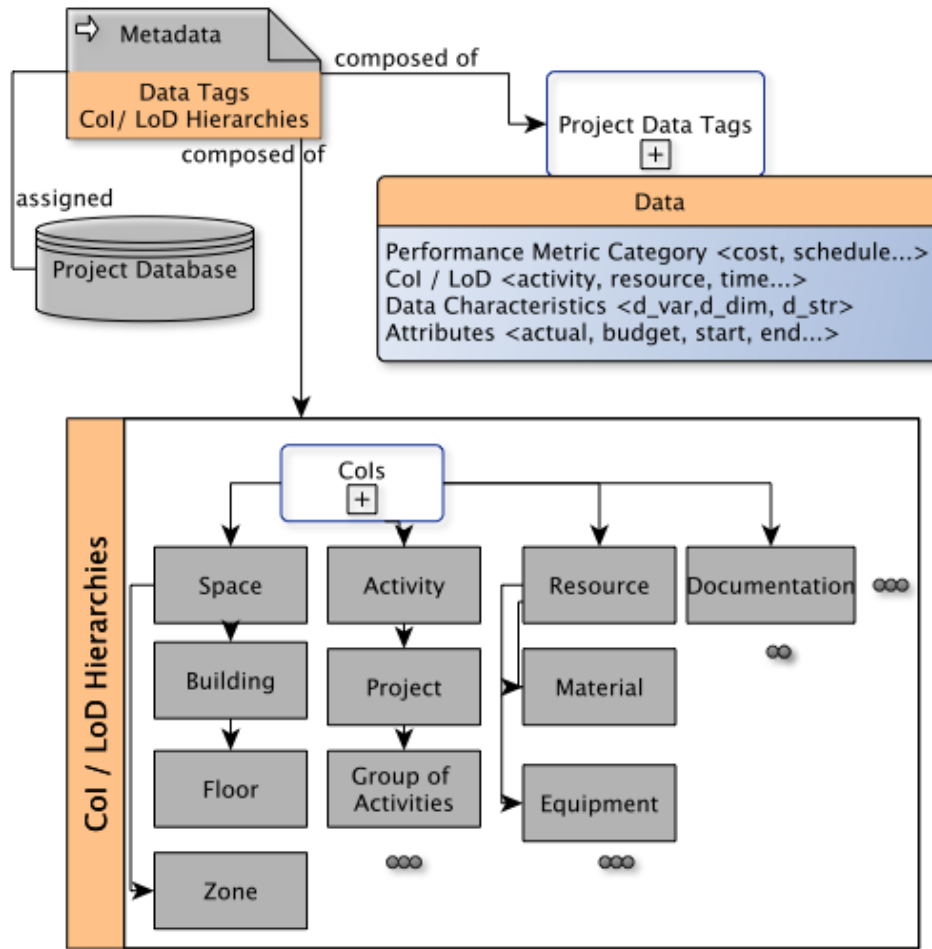


Figure 4.6: Project data metadata

When the requested data items of a metric query are identified, their data characteristics are captured. Figure 4.7 demonstrates the values assigned to *ProPerVisualWork* components for the activity cost comparison metric query example introduced in Figure 4.4. The intent acquired from user, and the data characteristics extracted through the project metadata are used in the following step to reason with *heuristics*.

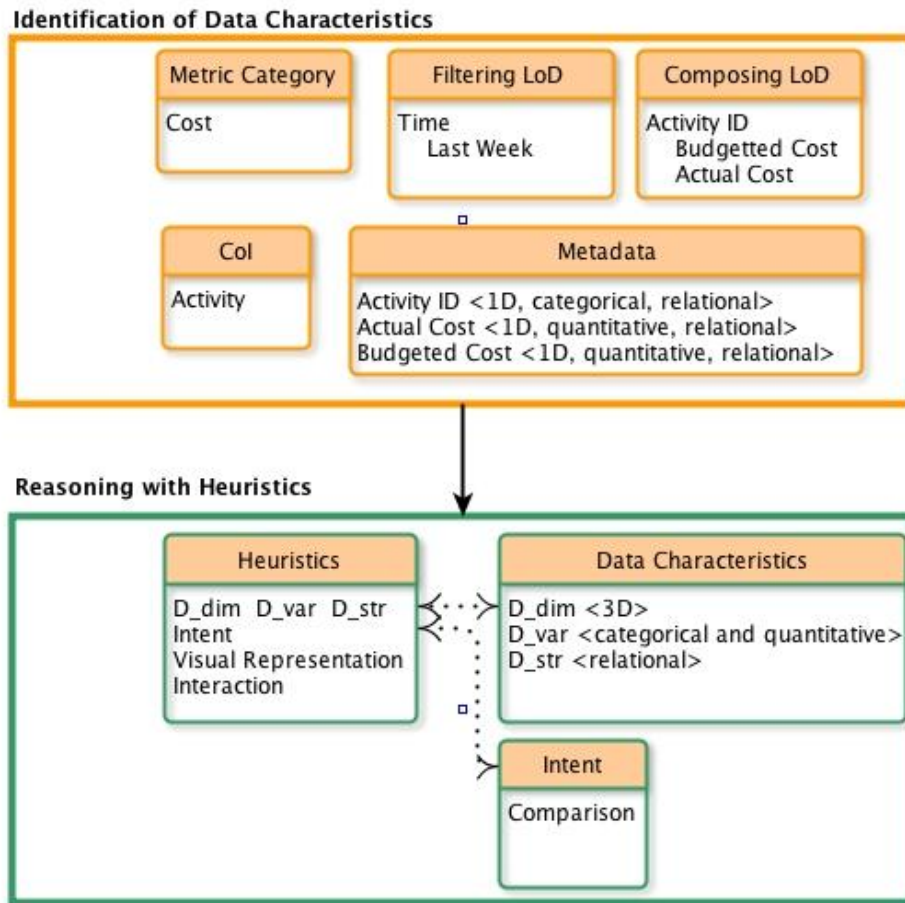


Figure 4.7: Identification of Data Characteristics

Upon the identification of the characteristics of the data items and intent queried by the user, heuristics table is used. Looking up the heuristics table for the relevant *heuristics* enables visual representation selection. Since most of the *heuristics* are incomplete (in terms of providing a full mapping between a visual representation and all characteristics of data and user's intent) a set of perfect matches based on all aspects of the framework might not be always found. Indeed, Figure 4.8 demonstrates the distribution of the represented *heuristics* based on which variables are used in suggesting the selection of a visual representation. It can be seen that only eight *heuristics* out of 490 uses all four

components of the framework (D_dim, D_var, D_Str, and intent) when suggesting a corresponding visual representation. Whereas, forty-six *heuristics* define the applicability of a visual representation by only using the D_dim component.

Venn Diagram of Visual Representations defined by Heuristic Components

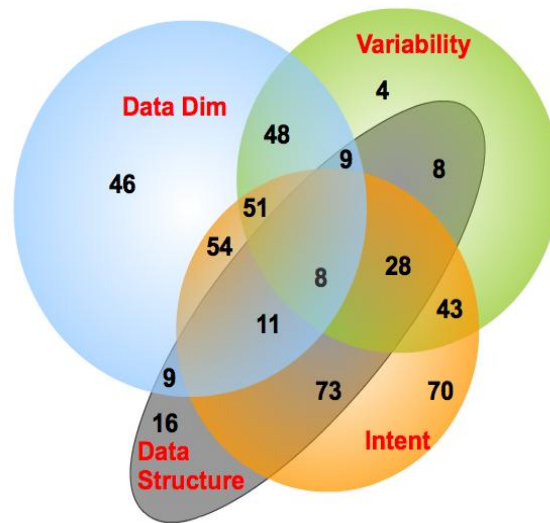


Figure 4.8: The Distribution of Heuristics based on how they use the components of the framework

Hence; applicable visual representations are selected according to different combinations of data characteristics and user's intent variables. Figure 4.9 demonstrates an example on selection of applicable visual representations. It is seen that the data characteristics and intent are used to look up matches in the *heuristics* table. However, *heuristics* are mostly not complete with all four variables assigned. Given that condition, different combinations of these variables are used to search for visual representations and stored accordingly. These facets are (a) 4s: All fours variables should match, (b) 3s: Three of the

variables should match, (c) 2s: Any two of the variables should match and (d) 1s: Any one of the variables should match. It should be kept in mind that while performing searches, mismatches are eliminated from the search space. For instance, if three of the variables (e.g., data characteristics) of the queried data match a heuristic the value of the fourth one determines whether the visual representation is selected or not. If the fourth variable (e.g., intent) is any value other than what is searched (for instance comparison is searched, and the heuristic states overview as intent) the heuristic is eliminated. On the other hand, if the fourth variable is “N/A” it means that no analysis has been performed in terms of that variable and therefore it can be applicable. In that case, that heuristic is captured as a valid one.

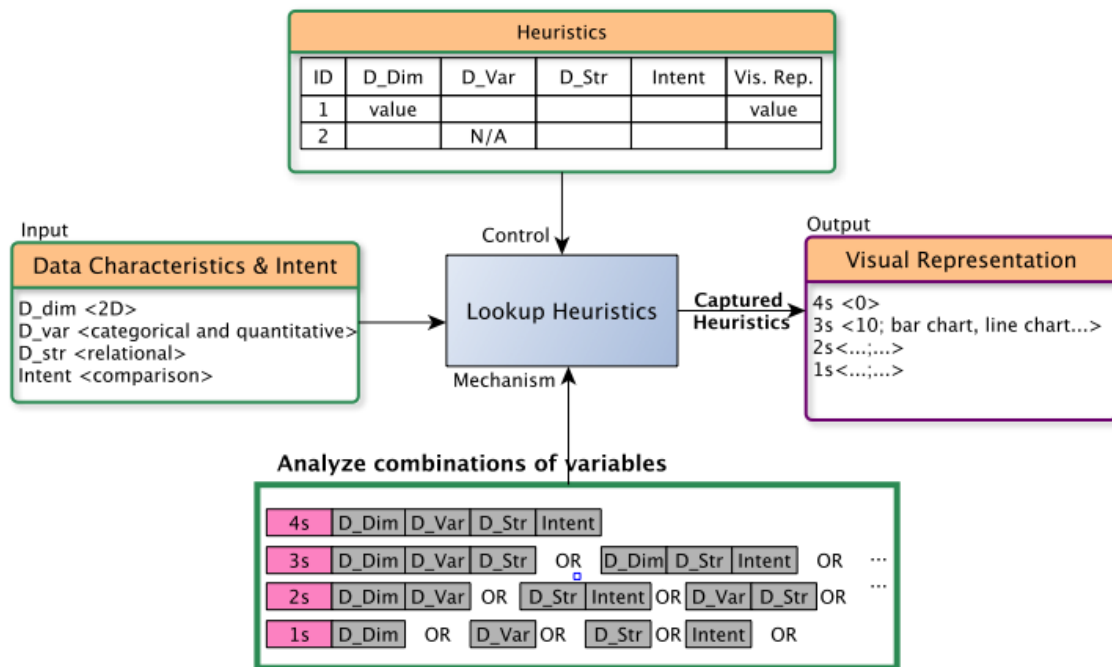


Figure 4.9: Looking up heuristics

As shown in Figure 4.9, upon reasoning with the *heuristics*, four facets of visual representations are created, where each facet represents a better match for the dataset

with the increase in the combination of the variables. In other words, 4s provides the best results, 3s provides better results than 2s and 1s, and 2s better than 1s.

When the applicable *heuristics* are identified for each of the facets, the visual representations that are captured within each facet is presented to the user with an order of more components of the framework is matched to less. As discussed earlier, the purpose of this research study is to enable project members to visualize data in an effective and desirable way. Since the cognition and preferences of users might differ in terms of visual representations, the selection of the visual representation will be left to the user among the facets of visual representations that are available to plot the requested dataset. However, in order to improve the selection process for the effective visual representations, two methods are proposed: (1) use constraint/ comparison *heuristics*, and (2) create user profiles.

Initially, “comparison/constraint *heuristics*” can be used to weigh or further sort the facets of visual representations to be presented to the users. The research studies that provide empirical findings on the comparison of the performance of visual representations or state constraints for some visual representations are stored as “comparison/constraint *heuristics*” as an additional table. Hence the *heuristics* can be categorized into two groups such as:

- Constraints: *Heuristics* that state constraints for visual representations in terms of data dimension are captured within this category. For instance, Icke and Sklar (2009) stated

that scatter plots should not display more than 3D data, and it is possible to map up to 18 data features to chernoff faces.

- Comparisons: *Heuristics* that compared the performance of visual representations with different performance metrics are captured within this category. These comparisons are mainly based on D_str, intent, and accuracy of results. For instance; Mackinlay et al. (2007) stated that line charts are best practice for temporal data. Similarly, Cleveland and McGill (1984) stated that bar charts are better than pie charts for distributive analysis. Additionally, Dull and Tegarden (1999) compared 2D and 3D charts given different interaction options, and sorted their performances.

The complete list of “constraint/ comparison *heuristics*” is presented in Appendix D. These *heuristics* has a total number of 50, and they capture 20 out of 163 visual representations that are represented in *ProPerVisualWork*. When the “comparison/constraint *heuristics*” list contains visual representations that are proposed by the reasoning mechanism, the visual representation with a higher priority can be identified using these *heuristics*.

Secondly, a user profile can be used for sorting the visual representations. A profile capturing the prior selections of the user based on the data characteristics, intent and context of the metric query, can help prioritize the visual representations individually for a user regarding the sequential metric queries.

Finally, based on the selected visual representation, the interactions that are identified as applicable to the selected visualization is identified. These interactions are stored as a match to visual representations. When the list of visual representations is identified, the associated list of interactions is presented to the user. In conclusion, in the research findings section, I have discussed my findings on the elicitation and representation of the information visualization related *heuristics* in a table format, and the use of look-up table method for selection of the effective visual representations for a given metric query.

4.6 Validation

Validation of the developed formalism involved evaluating its performance in three aspects: (1) its *precision* in terms of retrieving applicable visual representations, (2) its *usability* for construction management domain, and (3) its *usefulness* in terms of the effectiveness of the suggested visual representations. *Precision* is evaluated as percent reduction in the number of visual representations retrieved with the formalism as opposed to all visual representations in the library. *Usability* shows the relevancy of the retrieved *heuristics* to performance monitoring among all *heuristics*. *Usefulness* shows the perspectives of project members on suggested visual representations through the formalism.

In order to evaluate the precision and usability of the formalism, I performed synthetic experiments with the metric queries identified from the contextual inquiry. The contextual inquiries included fourteen project members, who requested 132 metric queries. The synthetic experiments included tests performed by me, using the *heuristics*

table, to test the performance of formalism in retrieving visual representations. I evaluated the usefulness of the formalism through a survey with additional fourteen project members. Validation results of synthetic tests and survey are provided separately, as detailed below.

4.6.1 Testbed for the synthetic tests with the formalism and its performance

132 metric queries that are presented in Appendix B are used for the synthetic validation of the formalism. In order to perform the tests, I identified the data characteristics and intents for each metric query. During the analysis of the test cases, I observed that six of the metric queries were composed of composite intents in other words contained two intents at once. Table 4.5 lists these queries with their metric IDs and short descriptions. During the elicitation of the *heuristics*, I observed that except for two *heuristics*, each heuristic is composed of a single intent. Having composite intent queries makes it impossible to map visual representations. Hence, these queries are divided into two segments with each intent.

Table 4.4: Metric queries with composite intents

Metric ID	Short Description	Composite Intents
1	Delta Cashflow: Comparison of estimated vs. actual monthly cashflow till today	Comparison & trend
2.1	Delta cost of project: Last week's actual vs. budgeted cost, where all activities are presented with the highest differences highlighted	Comparison & overview and detail
17	Total project duration: Value and data of current and targeted projects	Overview & Trend
88	Anticipated delta contract value: Original vs. anticipated contract value till the finish date.	Comparison & trend
122	Delta resource utilization: Weekly actual vs. budgeted resource usage per day	Trend & Comparison
123	Delta units: Activity cost variance for each activity with highest differences highlighted	Comparison & Overview and detail

The approach taken towards these cases is to identify the effective visual representations for both parts of the queries, and then checking for common visual representations. If exists, that should be the first to be suggested, if not the most effective ones from both categories can be combined through a common y-axis. In addition to the multiple intent exception, certain rules for the determination of the data characteristics of the datasets are utilized as part of the reasoning mechanism. These exceptions are as follows:

- Data structure being temporal, spatial, and spatio-temporal are straightforward for the data items. However, any data in a project database can be tabular or relational due to the fact that they are generally stored in databases and visualized in a tabular format. Additionally, some data items have the characteristics of a hierarchical breakdown as in the case of cost of activities composed of material, equipment and labor costs. Therefore, the solution space for the metric queries that have the characteristic of multiple structures, such as being both relational and hierarchical or temporal and hierarchical are treated as a union space. In other words, when a venn diagram notation is considered, each of the D_str is considered as a set (relational A, and hierarchical B), and the solution space is recorded as the combined area of both sets union A and B). Indeed, the visual representations that match both temporal and relational variables are accepted as valid solutions.
- During the assignment of the data dimension, the intents and interaction requirements of project members are also considered. For instance, if the project member requests to compare a multiD dataset (e.g., actual vs budget cost of each activity cost breakdown), D_dim assigned as multiD. However, if the project member requests to see an alert for this information, a flag is displayed in case there is cost overrun, and D_dim is

assigned as 1D for the same dataset. Because, rather than overseeing the all cost values, a single dimensional binary data is requested for the latter case.

- During the D_str assignments, the use of “hierarchical” is determined based on whether the higher level or the breakdown of info is requested or not. For instance, if the cost of activities is requested as the total weekly cost, the D_Str is relational only; however, if the cost of activities with the material, and equipment breakdowns included is requested, the D_Str is relational and hierarchical.

Using the above rules when applicable, the data characteristics and intent of each 138 metric queries are identified as presented in Appendix F. A short part of the table is presented in Table 4.5 for demonstration.

Table 4.5: Test Cases

ID	Description	D_dim	D_var	D_str	Intent
1	Delta Cost of Activities	2D	both	relational or tabular	Overview
2	Delta Schedule	multiD	both	temporal or hierarchical	Comparison
3	Over-time % factor	multiD	quantitative	hierarchical or relational or tabular	Overview
4	Agenda	multiD	both	temporal	Overview
5	Projected Profit	multiD	quantitative	temporal	Trend

Through the formalism, the *heuristics* that are applicable to each of the metric queries are determined and the 4s, 3s, 2s, and 1s facets of visual representations are identified for each test case. Appendix F presents the lists of the number of *heuristics* identified in

Table F.2, and visual representations in Table F.3. A set of test cases is presented in Table 4.6 for demonstration purposes.

Table 4.6: Facets of Visual Representations for the initial four metric queries

Test Case ID	Heuristics				Visual Representations			
	4s	3s	2s	1s	4s	3s	2s	1s
1	0	3	6	18	0	3	4	12
2	0	4	17	25	0	4	12	16
3	0	0	0	9	0	0	0	8
4	0	1	10	40	0	1	9	26

Table 4.6 shows that for each of the test cases, the total number of applicable *heuristics* and visual representations are identified. Since one visual representation can be used in multiple *heuristics*, the number of visual representations identified for each test case is different than the number of *heuristics*.

Precision

Precision is evaluated as percent reduction in the number of visual representations retrieved with the formalism as opposed to all visual representations in the library. The *heuristics* used and visual representations retrieved to each of the 4s, 3s, 2s, and 1s facets for 138 metric queries are identified, and the median and max values of each facets are presented in Table 4.8.

Table 4.7: Precision of formalism

(Median, Max)	Total	4s	3s	2s	1s
Heuristics	<u>490</u>	<u>0, 2</u>	<u>1, 17</u>	<u>4, 33</u>	<u>27, 43</u>
Visual Representations	<u>163</u>	<u>0, 1</u>	<u>1, 15</u>	<u>4, 27</u>	<u>20, 29</u>

As can be accepted, Table 4.8 presents that the number of *heuristics* and visual representations applicable per test case increase with less variables to be used in the

facets. It can be observed that there is only one visual representation that satisfies 4s situation through the use of two *heuristics*. That visual representation is called “data vases” invented by Thakur and Rhine (2009). It can be observed that on average one and at maximum 15 best matches of visual representations (within 3s facets) are presented for each test case. Considering the 1s and 2s matches, the retrieved visual representations are limited to 29 at max. Therefore, it can be concluded that *heuristics* provide efficient retrieval of the *heuristics* with 9 % and 18 % for *heuristics* and visual representatives respectively.

Usability

After each of the test cases are performed, I explored the *heuristics* table in order to detect the *heuristics* and visual representations that are used the least and the most. Figure 4.10 presents the usage of *heuristics* for the test cases. It is observed that while 8% of the *heuristics* are used for more than 50 times, 36 % of the *heuristics* are not used at all. The reasons for such a usage profile are explored. It is observed that the “multiD” is the most common D_dim for the metric queries with “the comparison and intent” as the mostly selected intents. Indeed, the most frequently used *heuristics* has the characteristics of the following combinations:

- multiD, relational/tabular, both quantitative and categorical with a number of 32 instances
- multiD, temporal, both quantitative and categorical with a number of 16 instances
- 1D, relational/tabular, text with a number of 7 instances

- multiD, hierarchical/relational, quantitative and categorical with a number of 7 instances

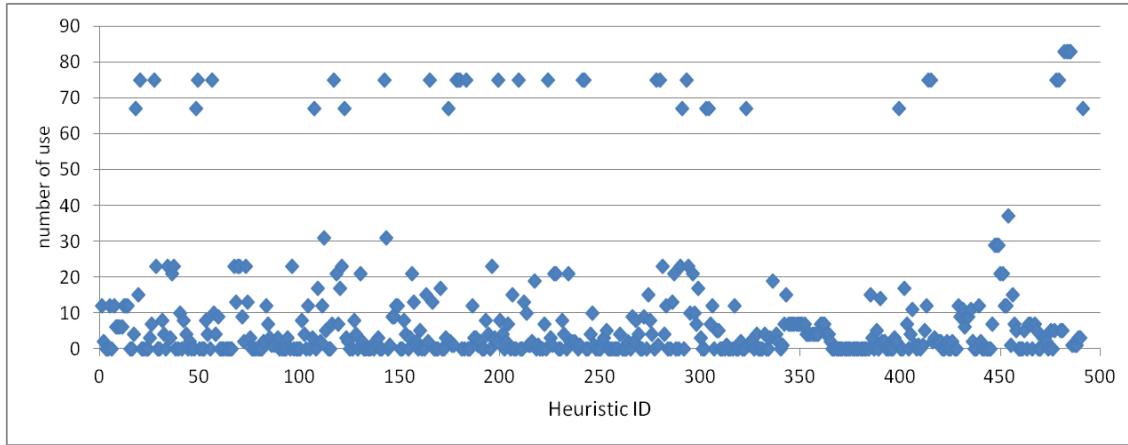


Figure 4.10: Usage of Heuristics

Additionally, the identified visual representations are evaluated, and Figure 4.11 presents the usage frequency of visual representations. Through the detailed exploration of the results presented in Figure 4.11, it is observed that while 35 of the visual representations have never been suggested based on the queries, 11 visual representations have been suggested for more than 100 times. On the average 30 visual representations for a metric query are retrieved with a standard deviation of 58. The pseudo code for the identification of the visual representations is provided below:

/ MQ: Metric query, as a total 138 MQ exist. */*

/ HV: Heuristic variables in a row of heuristics table. As a total 4 HV exist for one heuristic (D_dim, D_str, D_var, intent) */*

/ F: Facets of effective visual representations, where F1: facet one composed of one HV, F2: facet two composed of two HVs, F3: facet three, composed of three HVs, F4: facet four composed of four HVs. */*

For each MQ (from 1 to 138) do

```
{  
    Find F /*number of visual representations retrieved for F1, F2, F3 and F4 */  
    {  
        Check HV /* the process to check the HVs to see if they match the queried  
characteristic*/  
        {  
            Find F4 /* all HV satisfied, meaning each HV is checked once */  
            Find F3 /* each HV is checked 4 times (3 element subsets of a 4 element set) */  
            Find F2 /* each HV is checked 6 times (2 element subsets of a 4 element set) */  
            Find F1 /* each HV is checked 4 times (1 element subsets of a 4 element set) */  
        }  
    }  
}
```

From the pseudo code it can be seen that in order to identify visual representations for a metric query, the **Find F** function is performed once for each metric query. As shown in the code, **Check HV** process is run as a total of 15 times (1 for F4 + 4 for F3 + 6 for F2 + 4 for F1) for the **Find F** function. As a result, **Check HV** is performed 2070 times (138 * 15) to perform this validation. Within the 2070 searches, the list of most and least frequently requested visual representations are presented in Table 4.8. It can be seen that ‘chernoff faces’ is retrieved for 402 times and ‘parallel coordinates’ for 307 times out of 2070 searches, resulting in a retrieval rate of %19 and %15 as the highest rates.

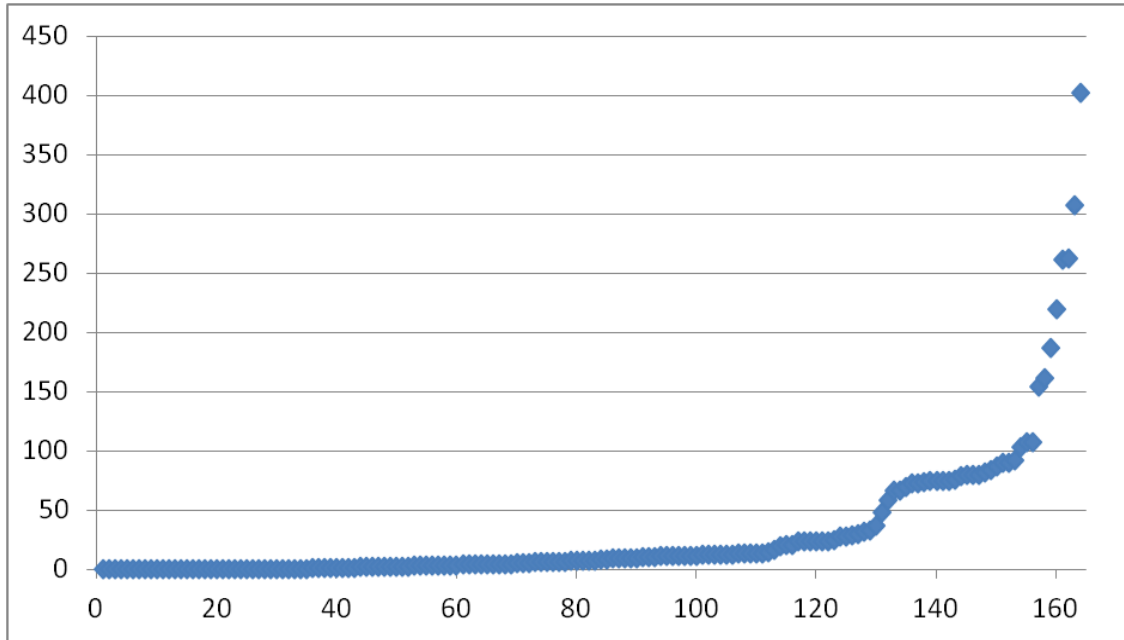


Figure 4.11: The number of visual representations used in the test cases

Table 4.8: Frequency of highest and lowest number of visual representations out 163

The Visual Representations	Number of Use
'3D bar graph', 'PolyViz', 'affinity diagram', 'bcg matrix', 'calendar', 'circular network', 'cognitive mapping', 'concept map', 'cone tree', 'cpm critical path method', 'decision discovery diagram', 'disk tree', 'document lens', 'entity relationship diagram', 'flow maps', 'geologic map', 'hyperbolic tree', 'hyperbolic view', 'information cube', 'isopleths', 'knowledge map', 'magic eye view', 'mindmap', 'pareto chart', 'radviz', 'sankey diagram', 'semantic network', 'shaded map', 'spectrogram', 'stacked 100% area graph', 'supply-demand chain', 'survey plot', 'tileBars', 'tree', 'tukey box plot'	0
'column graph'	103
'flowchart'	107
'geographic map'	107
'andrew curves'	154
'line graph'	161
'star glyphs'	186
'bar graph'	219
'scatterplot matrix'	261
'scatterplot'	262
'parallel coordinates'	307
'chernoff faces'	402

The least used visual representations are the ones designed for overview of text data, hierarchical correlation and flow charts. In Table 4.9, from the usability analysis it can be

seen that since certain data characteristics and intents are frequently used, a limited number of unique visualization options are mostly presented to the user. Within these unique options, it can be seen that only bar graph and geographic maps are the traditional construction management charts, while scatter plot, and column used traditionally for information visualization in general. The rest of the visual representations are new for CM. Therefore, it can be concluded that there are some patterns in data characteristics of the performance metric queries, which result in frequent suggestion of certain visual representations. that are not necessarily commonly used in the CM domain. This highlights the value of the formalism. Moreover, 128 out of 163 visual representations are suggested by the formalism, showing its usability for CM domain.

The patterns of data characteristics are further evaluated in order to understand if such a pattern is related with the context of the metric queries. During the analysis, the number of *heuristics* and visual representations identified within the 3s facet are considered (a total of 58 visual representations). The following observations are made:

- For most of the metric queries related with the CoIs of activity, and document, the suggested visual representations are stacked column chart and flowchart respectively.
- When the LoD is building component, the analysis is spatio-temporal and used visual representation is BIM.
- The CoI categories that has the lowest rate of being associated with a 3s match are resource and documents. Indeed out of 22 resource related metric queries, 13 of the metric queries retrieved a match for 3s facet, similarly, for the documents category 18 out of 44 metric queries provided a match for 3s facets.

- There are certain patterns for especially quick access related metric queries: for a link to safety related issues flow chart is proposed for four out of five metric queries; BIM and flow chart are proposed for all agenda and to-do related queries.
- Regarding communication metric category, 25 out of 37 metric queries did not have a 3s match of visual representations and for cost category, 10 out of 24 metric queries did not have a match for 3s facets. The most commonly retrieved visual representation is column chart for the cost category, while the most commons are column chart, and BIM for the time category.

These findings suggest that there are certain patterns for the visualization of time related performance metrics or quick access information; however, in general, the differences in LoDs of metric queries result in suggestion of different visual representations. The least number of *heuristics* identified in 3s categories suggest the need for further research on visualization of such information (documentation details and cost for the project).

4.6.2 Testbed for the Survey and its performance

A survey is conducted to validate the usefulness of the reasoning mechanism towards comparison of the user views on visual representations identified through different combinations of the components (variables). The survey is presented in Appendix E. A total of 14 project members that reside in Turkey and work on a variety of projects anticipated in the study.

The survey is composed of six questions within which users are requested to compare two visual representations for the given metric query for that question. In other words, through six questions, 12 visual representations are paired into groups of two and compared. The metric queries and associated visual representations are presented in Table 4.9.

Table 4.9: Survey questions

Question	Metric Query	Chart 1	Chart 2
1	Overview of multi-project weekly overtime factor	Heat map	Bar chart
2	Comparison of actual vs. budgeted cost of activities performed previous week	Column chart	Scatter plot
3	Trend in cost overrun for each activity performed to-date	Combined line and column charts	Area chart
4	Overview of space usage for utilities, works areas ,and etc.	2D map	Table
5	Distribution of submittals based on their submission rate and status	Scatter plot	Pie chart
6	Overview of the weekly material procurement schedule	Calendar	Gantt Chart

The selection of the visual representations is performed according to the discussed reasoning mechanism in the way that the visual representations that are retrieved as 4s, 3s, 2s, and 1s facets are identified. For instance given he first query in Table 4.9, no visual representation is retrieved for 4s facet. However, 3, 9, and 40 visual representations are identified for 3s, 2s, and 2s categories respectively. The chart 1 option is selected within 3s facets as heat map. Similarly, the Chart 2 option is selected within 1s facet as bar chart. Performing similar selections for the other five questions, Chart 1 options for each of the questions were selected from the 3s facets of the metric queries. The second charts were selected from the facets of 1s or 2s as shown in Table 4.10. Table 4.10 details whether the chart belongs to a 1s or 2s facet, and using which of the data characteristics and/or intent it is selected. The selections were not totally random; instead

a equal number of 1s and 2s options are used, where each one of the data characteristics (D_dim, D_var and D_str) and intent is used at least for once.

Table 4.10: The selection of charts within facets of 1s and 2s as the 2nd chart option

Questions	Visual Representation		Variable(s) used in the
	(Chart 2)	The Facet	Facet
1	Bar chart	1s	D_dim
2	Scatter plot	2s	D_var and Intent
3	Area chart	2s	D_dim and D_var
4	Table	1s	Intent
5	Pie chart	1s	Intent
6	Gantt Chart	2s	D_str and Intent

A 5-scale likert scale is used for the evaluation of each of the visual representations with the scales of; 1: Definitely effective, 2: Effective, 3: Neither effective nor ineffective, 4: Not effective, 5: Definitely not effective. The results acquired for each 6 questions regarding both charts are presented in Table 4.11. Table 4.11 presents the overview of the survey results for each of the fourteen participants and each of the charts presented for six questions. The results of the table are evaluated with the box plot representation of the results.

Table 4.11: Survey results

	Q. 1 Chart 1	Q. 1 Chart 2	Q. 2 Chart 1	Q.2 Chart 2	Q. 3 Chart 1	Q. 3: Chart 2	Q. 4. Chart 1	Q. 4 Chart 2	Q. 5 Chart 1	Q. 5 Chart 2	Q. 6 Chart 1	Q. 6 Chart 2
ID	1	2	1	2	1	2	1	2	1	2	1	2
1	1	4	3	1	3	1	4	1	4	3	3	4
2	4	5	5	1	3	3	5	4	4	5	4	3
3	1	1	1	4	4	1	1	4	4	1	4	1
4	1	5	5	3	4	3	5	3	4	5	5	1
5	1	4	3	4	1	3	1	3	3	4	4	3
6	4	5	5	4	4	3	4	3	3	4	3	4
7	2	5	4	3	4	3	2	4	3	4	5	1
8	2	4	4	2	4	3	5	3	4	4	5	4
9		5	5		3		2	2	1	4	1	3
10	4	5	4	3	5	3	5	4	4	5	5	5
11	4	5	5	2	5		5	3	4	5	3	5
12	2	4	4	3	4	3	5	2	4	4	4	3
13	3	5	4	2	5	3	4	3	1	4	4	5
14	4	5	5	2	4	2	5	3	5	4	5	4

The statistical results for the survey data is calculated and presented in Table 4.12 in order to acquire the box plot distributions as shown in Figure 4.12.

Table 4.12: The statistical results for the data survey data

Chart 1s	Question1	Question2	Question3	Question4	Question5	Question6
Quartile 1	1	4	3	2	3	3
Min	1	1	1	1	1	1
Median	2	4	4	4	4	4
Max	4	5	5	5	5	5
Quartile 3	4	5	4	5	4	5
Chart 2s	Question1	Question2	Question3	Question4	Question5	Question6
Quartile 1	4	2	3	3	4	3
Min	1	1	1	1	1	1
Median	5	3	3	3	4	3
Max	5	4	3	4	5	5
Quartile 3	5	3.25	3	4	5	4

I have developed two premises before interpreting the results. These premises are: (1) Premise 1: The project members will score the traditional charts, higher than their alternatives, and (2) Premise 2: The project members will score lower for the visually cluttered charts (the clutter can be measured in terms of relative and absolute data density and cluster as suggested by Bertini and Santucci (2004)).

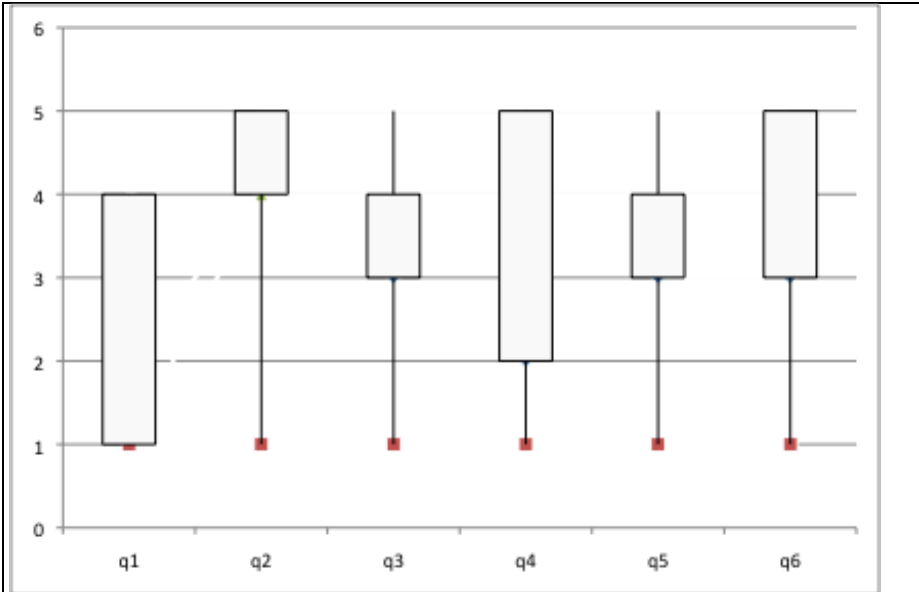


Figure 4.12a: The box plot of the survey results for Chart 1 presented in each six questions

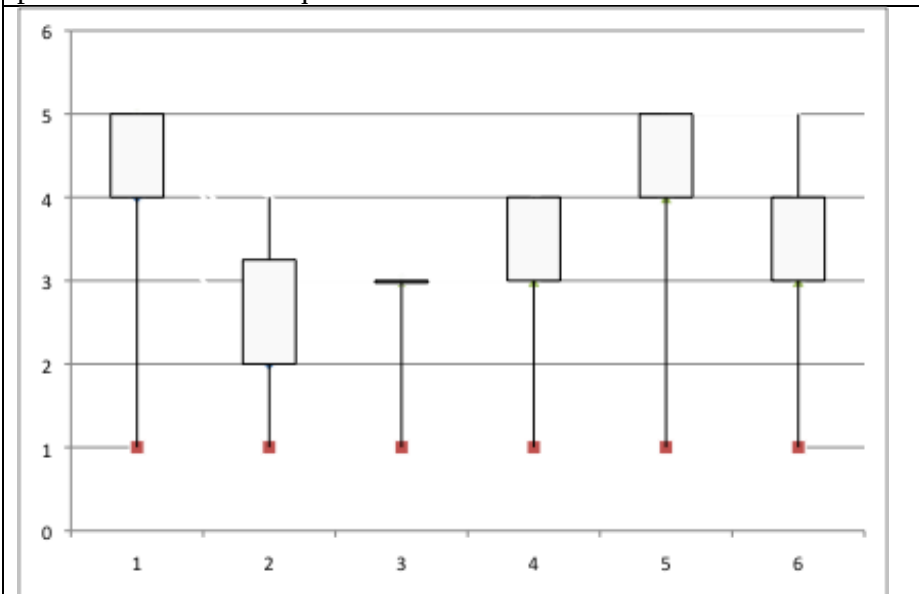


Figure 4.12b: The box plot of the survey results for Chart 2 presented in each six questions

Figure 4.12: Box plot distributions

Based on the aforementioned premises, the exploration of the charts used in the survey and their scores displayed in Figure 4.12 for each of the questions is performed. The results for each question in order are as follows:

1. Question 1: The 1st Chart is not a traditional chart and it is hard to understand the legend and interpret the chart quickly. However, 2nd chart is a traditional chart. Since the 2nd Chart has higher scores, it can be seen that premise 1 is confirmed.
2. Question 2: In terms of the readability of the data points, the 1st Chart is clear (effective) whereas 2nd chart has clutter (not effective). Since the 1st Chart has higher scores, it can be seen that premise 2 is confirmed.
3. Question 3: 1st Chart is cluttered but it is still possible to read data from these traditional charts. However, 2nd chart does not comply with the intent of the analysis (not effective) and data is hard to read from this untraditional chart. Both 1st and 2nd premises are confirmed.
4. Question 4: 1st Chart is traditional however, needs improvement in terms of providing dynamic interaction to get details. Albeit being traditional, 2nd Chart requires a map to interpret the location codes. None of the premises were tested for this question, because both charts are traditional and readable.
5. Question 5: For this question both traditional and simple charts are used. However, Chart 2 is more effective in terms of less data density ratio, confirming Premise 2.
6. Question 6: Both traditional charts are used; however 1st Chart is simpler to read resulting in confirming Premise 2.

This analysis shows that both hypotheses are confirmed. Hence, two major conclusions that can be derived from the hypotheses are: (1) Project members score more on traditional charts than innovative charts, when applied, and (2) The charts that are easier in terms of *data reading* are more effective. The means, minimum and maximum values assigned for the Charts are presented in Table 4.13.

Table 4.13: Mean and Max. scales assigned to Charts

Questions	1	2	3	4	5	6
Chart 1 mean	2.5	4.1	3.8	3.8	3.4	3.9
Chart1 (min, max)	1, 4	1, 5	1, 5	1, 5	1, 5	1, 5
Chart 2 mean	4.4	2.6	2.6	3	4	3.3
Chart2 (min, max)	1, 5	1, 4	1, 3	1, 4	1, 5	1, 5

From the results presented in Table 4.14, it can be observed that:

- For almost each chart, there is a project member who finds the chart effective
- For each chart, there is a project member who finds the chart definitely not effective
- For Chart 1 types (selected by 3s commonalities), the median is “effective” for 5 out of 6 questions regarding 14 participants. Within these 6 questions two charts are rated in the ineffective scales (1 or 2 on likert scale) due to the following reasons:
 - For the 1st question, a small dataset is used for the analysis, and the legend (D_var) is as much as the total data size. Hence, this visual representation should be used for larger datasets.
 - In order to depict the locations of the utilities, codes are used for the map presented in 4th questions. The take home is not clear without knowing the codes hence, either a better notation system or dynamic interaction is required.
- For Chart 2 types (selected by 1s or 2s), median is “don’t know” for 4 out of 6 cases.

In other words, Charts are mostly not as effective as the ones selected amongst 3s

facets but still rated as effective expect for the chart that is presented in question 2. For question 2, the visual clutter makes it impossible to read data points.

Hence, from analysis of mean, min and maximum values assigned to the charts two major conclusions can be deduced as follows:

- The formalism is useful in terms of the fact that the visual representations selected within 3s categories provide better results than the alternatives.
- All of the charts received ratings in the both extremes of the likert scale. Hence, it can be deduced that in addition to the data characteristics and intent, the performance of visual representations relies on the individual perception of the participants.

It has been stated that the second charts (Chart 2) presented in the questions are acquired from either Facet 1s or Facet 2s. Indeed, the visual representations for questions 1, 4 and 5, were retrieved with the list of 1s facets; whereas for questions 2, 3, and 6, visual representations were selected within the list of 2s facets. In order to compare the performance of 1s and 2s facets, I explored the 2nd Chart results. Higher ratings are observed for questions 1 and 5, whereas lowest ratings are observed for question 2. If to be interpreted on its own, these results would suggest that 1s selections are better than 2s selections. However, since the effectiveness is a multivariate evaluation, I realized that some of the important factors that could change these results were not include in the testing process such as providing interaction, and using larger datasets, and requesting the users to perform tasks such as read, compare rather than fill in the survey. Hence it is

believed that the results would be different if the validation process is improved with the presented suggestions.

When evaluated in overall, it can be seen the Charts that are identified as a result of 3s facets are more effective than 1s and 2s facets. This observation supports the fact that the visual representations that are identified with the highest number of matches for data characteristics and intent provide the best results for the dataset. This observation validates the value and usefulness of the formalism.

In conclusion, the synthetic tests performed for the precision and usability evaluations showed that the formalism provides high precision in terms of retrieving relevant visual representations. The usability of the formalism is limited to the extent that certain intents and data characteristics are used for performance monitoring which results is retrieval of 11 visual representations with a frequency over 100. However, albeit this fact, 64% of the *heuristics* are used for the metric queries overall. This usability would increase if more variety of project members were analyzed from different projects and locations. The usefulness of the formalism is validated due to the fact that survey participants mostly rated suggested visual representations higher than alternatives.

CHAPTER 5

5 Conclusion and Future Works

This chapter presents an overview of my PhD research on visualization of and interaction with construction project performance information, and provides a summary of the contributions and practical implications of the research. It also discusses possible directions for future research.

5.1 Summary of Contributions

This research has three main contributions that relate to the three research questions explained in Chapters 2 to 4: 1) Identification of the commonalities and differences in project performance visualization needs of project members; 2) Development of a construction project performance visualization framework; 3) Representation of and reasoning with information visualization related *heuristics*. Each contribution is described in more detail in the following three subsections.

5.1.1 Identification of the Commonalities and Differences in Project Performance Visualization Needs of Project Members

When monitoring the performance of construction projects, project team members prefer easily accessible, customizable and individually relevant information to gain project awareness. Information required to monitor the performance of projects varies amongst the project team members based on their responsibilities, objectives and experience levels. However, currently such performance visualization needs are not fully supported. Not meeting the unique information visualization needs of project members results in

wasting unnecessary time in accessing and visualizing information. During my internship at a construction job site, I observed that one of the project engineers spent 10% of his time on a daily basis to aggregate and visualize the timecard information in a spreadsheet view, which is the view that he preferred to visualize and analyze such daily production data. When accumulated over time and project members, the inability to visualize individually relevant information might be costly for projects.

Although there are research studies that identify performance metric queries that are generally used by project members, the project performance visualization needs of project members have not been identified and discussed in detail. A formal approach is needed to identify the information necessary to be captured in relation to project performance visualization. In order to address this need, I performed contextual inquiries with various construction project members. Through fourteen interviews, 132 metric queries were identified in detail. Total number of metric queries requested by each interviewee changed from five to eighteen with an average of eight. No correlation of the number of requested metric queries with the organizational hierarchy of the project team members in terms of job titles and the number of years of experience is found. Within the 132 metric queries, safety and quality were the least requested metrics, whereas communication of documents and cost were the mostly requested metrics. The contextual design resulted in understanding and identifying different information visualization needs of project members. These information needs are grouped under nine commonalities: visualization intent, usage purpose, metric category, concept of interest, level of detail, visual representations, interaction, frequency, and notifications. Within each of these

commonalities, the differences of each of the performance metric queries are explored. Seven categories of performance metrics are identified as a result of the consolidation of the interviews. These are: (a) cost, (b) time, (c) productivity, (d) quality, (e) safety, (f) communication of information, and (g) quick access. Nine types of CoIs are identified such as activity, components, documentation, indirect/ overhead, issues/ inspection, project, resource, space, and milestones. For each of the CoI, various levels of details are also identified. Regarding the 132 metric queries, 22 different visual representations were requested. Mostly requested visual representations are buttons/links for quick access to information, Gantt Chart for time, table for communication of documentation and spreadsheets for cost. The intents identified during the interviews included clustering, comparison, composition, correlation, detail on demand, distribution, overview, overview and detail and trend. The usage purpose of the metric queries was identified as one of the five types such as descriptive, explanative, predictive, evaluative, and prescriptive. It is observed that 81 % metric queries were requested as interactive as opposed to static information. For 28% of the metrics, notifications were requested in the form flashing flag or highlighted information. Five levels of frequency for information visualization were identified as always, always with notifications, based on data being available, based on pre-determined time frame, and further graphs. The numbers acquired from the interviewees is a representative set of demographics within the scope of this research study, and does not necessary represent the AEC industry due to the fact that the engaged organizations were not explored in terms of their strategies, utilized construction methods and processes.

The validation of commonalities and differences were performed with the use of two mathematical clustering approaches as principal component analysis (PCA) and k-means. The PCA analysis provided me with the most influential variables for which the highest numbers of coefficients are assigned as Eigen values. The four variables that are identified with the highest coefficients are frequency, LoD (attribute), intent and CoI. For the k-means analysis the principal axes were selected within the four variables that were identified as most influential by the PCA analysis. Clusters ranging from two to eight were created iteratively. Regardless of the used principal components and number of clusters, the sparse distribution of the observation points validated the differences identified through this research study.

5.1.2 Development of a Construction Project Performance Visualization Framework

Identification of the commonalities and differences in information visualization needs of project members revealed that there is a gap in prior work associated with representing the information required to visualize project performance. This research task has been driven by the vision that a project visualization environment is required to monitor the construction projects so as to generate effective visual representations of the project performance data to make sense of it. Current project performance monitoring frameworks only capture the metric query descriptions without any details on their granularity, visualization or interaction. Hence, there is a need to develop a framework for project performance visualization.

In this research, I developed a framework that has components representing the commonalities identified through interviews to support unique needs of project members and synthesis of previous research efforts on applicability and effectiveness of visual representations. The project performance visualization framework is abbreviated as *ProPerVisualWork*.

ProPerVisualWork is composed of fourteen components each having a primary, secondary, tertiary or inoperative function. These components are; (1) Primary Components that help define, search and filter project performance data. These include Performance Metric Category, Concept of Interest, Level of Detail, and Project Data Metadata, (2) Secondary Components that enable storage and reasoning of information visualization related *heuristics*. These include Data Characteristics, Intents, and *Heuristics*, (3) Tertiary Components that provide the output of the framework in terms of display and dynamicity. These include Visual Representations, Interaction, and Notifications, and (4) Inoperative Components that help characterize the framework with the user and visualizations related factors that are not put in use to facilitate performance monitoring. These include Decision, Frequency, Visual Representations Metadata, and User Characteristics. The coverage of the framework and the usefulness of each of the components were validated through a combination of literature review, synthetic tests, and surveys.

5.1.3 Representation of and reasoning with information visualization related heuristics

There are several commercially available project management tools that are designed to support various project tasks, such as scheduling, and planning. Similarly, various

information visualization toolkits that are designed to analyze and visualization any kind of data exist. The existing toolkits provide a platform to visualize information; however, however, such visualization does not meet the needs of project members from several aspects: a) not all data types are supported, b) not all intents are supported, and c) the most effective ways of visualizing a queried performance metric are not supported.

Arising from the need to identify effective visual representations for a given performance metric query, *heuristics* component that is introduced in Chapter 3 as a part of *ProPerVisualWork* is studied in detail in Chapter 4. *Heuristics* capture findings from previous research studies that describe the appropriateness of visual representations for various types of information and user intent. A wide range of research studies that explored effective information visualization with different contexts and focus groups are identified. With the existence of such a variety of research studies, there is a need for integrating these *heuristics* through developing a common language and representing them in a format to be used to streamline the selection of visual representations. However, this representation is challenging, because the majority of the *heuristics* are semi-structured, incomplete, ill-defined, or contradicting.

Over 150 papers that are published over the last 20 years on information visualization were studied. More than 600 *heuristics* are identified with the use of six major variables: (1) Data dimension (D_dim): The dimensionality of the data is identified according to the number of data fields that are analyzed. Its enumeration includes 1D, 2D, 3D, and multi D, (2) Data variability (D_var): The variability of the data defines whether the dataset is

composed of quantitative, categorical, text or a mixture of these variables, (3) Data Structure (D_str): The structure of the data defines the interrelationship of datasets, in other words how they are represented. Six data structures are used this study including temporal, spatial, hierarchical, relational, spatio-temporal and tabular, (4) Intent of User (Intent) The intent of the user refers to the type of analysis that a user performs on a dataset. Nine types of intents are used in this study including comparison, cluster, correlation, comparison, distribution, trend, detail, overview, and overview and detail, (5) Visual Representations: A collection of graphical symbols that visually properties and relationships of a represented information capturing tables, texts, graphs, maps, diagrams, icons and networks. In this study a total of 163 visual representations are used, and (6) Interaction: User input to perform a task with the visual representations, such as emphasizing, filtering, and overlaying extra information.

During the elicitation of the *heuristics*, a large number of *heuristics* (~100) are eliminated from the findings due to five major factors such as absence of matching visual representations, use of out of scope visual representations, use of generic representation modes, algorithms based visualization methods and distortion-based tool-specific visual representations instead of specific visual representations. With the elimination of the inapplicable *heuristics*, a total of 490 *heuristics* that are identified from 35 previous research studies are used in the *heuristics* pool and represented in the form of a look-up table.

Using the look-up table, a formalism that enables selection of effective visual representations for unique performance monitoring needs of project members is developed. The formalism identifies the data characteristics and user's intent for each query and searches through the *heuristics* table to find the matching visual representations. Applicable visual representations are selected according to different combinations of data characteristics and user's intent variables. The visual representations that are identified as a result of the formalism are proposed to be sorted according to individual preferences of users, and through the use of *heuristics* that compare effectiveness of visual representations as a part of future studies.

The developed framework is validated in terms of (1) its *precision* in terms of retrieving applicable visual representations, (2) its *usability* for construction management domain, and (3) its *usefulness* in terms of the effectiveness of the suggested visual representations. *Precision* is evaluated as percent reduction in the number of visual representations retrieved with the formalism as opposed to all visual representations in the library. *Usability* shows the relevancy of the retrieved *heuristics* to performance monitoring among all *heuristics*. *Usefulness* shows the perspectives of project members on suggested visual representations through the formalism.

5.2 Practical Implications

This PhD research can possibly have practical implications for construction project members who are monitoring projects, and construction management software development teams. This section overviews such practical implications.

5.2.1 Supporting project members for project awareness

It is discussed that with the use of the framework and the formalism, the project members can identify effective visual representations for monitoring the performance of construction projects. Hence, by enabling the project members to generate and visualize their performance metrics in a more effective way, it is envisioned that this research study would likely increase the project awareness of project members, preventing probable information overload.

Additionally, with the identification of the similarity and differences in performance visualization needs of project members, the visual representation selection of project members for various performance metrics can be stored in user repositories. This would possibly enable the novice project members to take advantage of the expertise of senior project members in selecting the right performance metrics to track the project information.

5.2.2 Assisting software developers / other domains

The detailed analysis of commonalities and differences in project performance visualization needs of project members can be an invaluable reference for software developers. Indeed, they can be aware of the needs of project members that would help them improve their solutions based on the identified needs. Similarly, such information can be used by construction companies to form organizational knowledge on the visualization of project information to monitor the performance.

The contextual inquiry method of that is used for the identification of project performance visualization framework can be adopted by other domains to build up on the developed framework and modify it according to needs of their domains.

5.3 Future Research

Several aspects of the framework and formalism could not be covered as a part of this PhD research. The following future research directions are identified as possible extensions of this study.

5.3.1 Extending the Elicitation of Information Requirements

The research described in this thesis used Contextual Design to identify the commonalities and differences in information visualization needs of project members. The scope of the findings is limited to the interviews performed with project members

that live in Pittsburgh/ PA. Another user study with a set of project members from a different location can be performed in order to compare the findings.

During the use of the Contextual Design method, only three steps of the process is used including contextual inquiry, work modelling, and consolidation. The consolidation is followed by work redesign; creating storyboards from consolidated data to understand the ways to improve the users' work, user environment design; structuring the system work model, and defining its parts and their inter-relationship, mock-up and test with users; testing the developed ideas with users through prototypes. The last three steps of contextual inquiry are mostly related to developing a working prototype, which was out of scope for the information requirements identification purpose of this research study. As a future work, such steps of the Contextual Design can be implemented in order to improve the findings.

This effectiveness of visual representations utilized in this research study is dependent on the *heuristics* identified through previous research studies. Hence, the size and quality of the *heuristics* space is dependent on the number and quality of these studies. It is possible to extend the *heuristics* look-up table with the identification of new findings.

5.3.2 Extending the Framework

The research described in this thesis developed a framework that focuses on capturing the information items required to monitor the quantitative performance of construction projects. In this research study, I specifically focused on the performance measurements

that are relevant to the construction phase of projects and limited the performance metrics to quantitative ones. Qualitative performance metrics and other stages of construction projects are kept out of scope. In order to have a comprehensive pool of performance measures, the framework can be extended to include qualitative performance metrics. Additionally, components such as simulation, and historical data can be added to the framework to cover other stages of construction projects such as estimation and planning.

The scope of the physical environment that project personnel interacts with project performance metrics has been defined as their office spaces and their personal computers that they will use to generate and visualize performance metrics to monitor the construction projects. The use of various digital platforms and their effect in generating views are not addressed in this research study. Hence, the framework can be extended with programming components to take into account the use of different platforms (e.g., handheld devices). However, in order to address such a need, the effectiveness and use of visual representations in these environments need to be studied.

This research study did not focus on addressing how comprehensive and/or generic the project data needs to be represented and stored since such issues are large-scale research problems within themselves and were beyond the scope of this research. Hence, the framework can be extended with the use of a project database, rather than the use of project data metadata.

The inoperative components of the framework are derived to help characterize the framework with the user and visualizations related factors. These components are decision, frequency, metadata about visual representations and user characteristics. No existing research study has been found on the effect of decision and frequency of visualization for the selection of visual representations. Hence, these components were not put in use to facilitate performance monitoring. These components can be studied to be used for selection of visual representations within an automated visualization environment.

Additionally, metadata about visual representations is defined as a component of the developed framework. However, since the developed formalism was limited to identification of visual representations and did not capture the generation of visual representations, a comprehensive metadata is not developed. Such metadata should capture attributes for defining the space, representing the visual primitives, graphical objects, data characteristics and the interaction types as shown in Figure 5.1. A number of research studies identified such requirements and developed taxonomies on representing the visual representations. These studies are complimentary to each other in terms of the targeted visual representations, and context of the taxonomies in terms of capturing spatial decompositions of charts, and required attributes to identify how they can be generated and interacted. Hence, the framework can be extended by synthesizing such studies to enable generation of proposed visual representations.

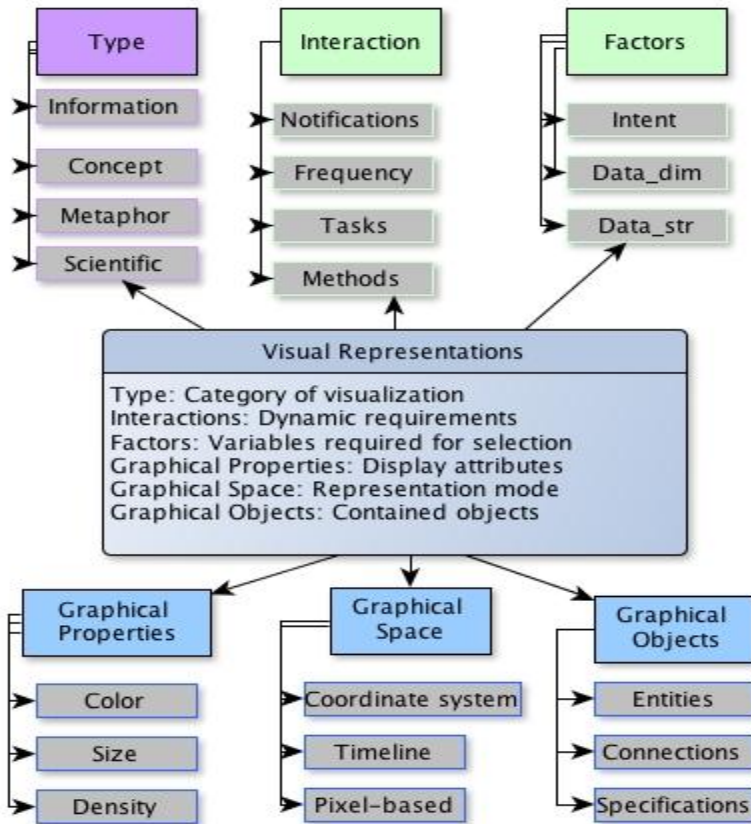


Figure 5.1 Visual Representations Metadata

5.3.3 Extending the Formalism

The variables that are used for selection of visual representations are data characteristics and intent of a user. Information visualization is a widely studied topic with an increasing rate in the number of studies performed. With the identification of the new *heuristics*, the formalism can be extended through adding new types of visual representations, or with different classification of intents, and data characteristics.

Moreover, in addition to data characteristics and intent, other information items are stored within the *heuristics* table that are not currently in use, such as data size and

representation mode. These variables may be useful in the future, if more research in performed on these subjects.

- Data size: Out of 490 *heuristics*, 70 of them had discussions on the required size of the data for a certain visual representation. However, except 10 of them, these studies used vague terms, such as “few” variables, “many” fields and “small” data sets. Since these terms cannot be quantified, this variable has been used in the selection of visual representations. While data size would be an important criterion for selection of visuals, further research is required to reason with this variable within *ProPerVisualWork*. These can be helpful in selecting visuals, but needs further findings.
- Representation mode: 5 out of 35 studies discussed the representation mode of the visual representations. Representation mode defines the arrangement of visuals based on whether they are pixel-oriented, icon-based, or coordinates-based. Though currently these representation modes do not provide any input for the selection of visual representations, with further research in the future on mapping of such representation modes with visual representations and interaction, representation modes can be another variable for the selection of visual representations.

The developed formalism depends on the findings of the existence research studies, and suggests a list of effective visual representations for a given metric query. Further research can be performed to empirically test the effectiveness of visual representations for different combinations of data characteristics and intents. Hence, the formalism can be extended with validated empirical results on effectiveness of visual representations.

It is known that the perception of each individual may vary while comprehending information through different means of visualization. Since a list of effective visual representations is presented to users, their selections can be stored in a user repository based on the queried information.

CHAPTER 6

6 Bibliography

- Abdel-Khalik, A.R. (1973). "The Effect of Aggregating Accounting Reports on the Quality of the Lending Decisions: An Empirical Investigation." *Journal of Accounting Research*, vol 11, pp 104-138.
- Abdi, H., and Williams, L. J. (2010). "Principal Component Analysis." *Wiley Interdisciplinary Reviews: Computational Statistics*, vol 2, pp 433-459.
- Abeid J., and Arditi, D. (2003). "Time-lapsed Digital Photography Applied to Project Management," *Journal of Construction Engineering and Management*, 128(6), pp 530-535.
- Abela, A. (2006). "Chart Suggestions— A thought-starter." From: http://img.labnol.org/di/choosing_a_good_chart2.pdf, <Last Accessed: December 2012>.
- Abdul-Rashid, I., Bassioni, H., and Bawazeer, F. (2007). "Factors Affecting Safety Performance in Large Construction Contractors in Egypt." *Proceedings 23rd Annual ARCOM Conference, 3-5 September 2007 Belfast, UK*.
- Abou-Zeid, A., Russell, J.S., hanna, A.S., Park, S.C. (1995). "Data Flow Model for Communications between Project Participants in a Highway Bridge Project." *Canadian Journal of Civil Engineers*, 22(6), pp 1224-1234.

- Aigner, W., Bertone, A., Miksch, S., Tominski, C., Heidrun, S. (2007). "Towards a Conceptual Framework for Visual Analytics of Time and Time-Oriented Data." *IEEE Winter Simulation Conference*, Piscataway, NJ, USA, 2007, pp 721-729.
- Akinci, B., Fischer, M., Levitt, R., and Carlson, R. (2002). "Formalization and Automation of Time-Space Conflict Analysis." *Journal of Computing in Civil Engineering*, 16(2), pp 124- 134.
- Alarcon, L.F. and Ashley, D.B. (1996). "Modeling Project Performance for Decision Making." *Journal of Construction Engineering and Management*, 122(3), pp 265-273.
- Amar, R., Eagan, J., Stasko, J.T. (2005). "Low-level Components of Analytic Activity in Information Visualization." *IEEE Symposium on Information Visualization*, Minneapolis, MN, USA, pp 111-117.
- Aouad, G., Kirkham, J., Brandon, P., Brown, F., Child, T., Cooper, G., Ford, S., Oxman, R., Young, B. (1995). "The Conceptual Modeling of Construction Management Information." *Automation in Construction*, 3(4), pp 267-282.
- Arayici, Y., and Aouad, G. (2005). "Computer Integrated Construction: An Approach to Requirements Engineering." *Engineering, Construction, Architecture Management*, 12(2), p 194, 2005.
- Atasoy, G. and Akinci, B. (2010). "Evaluation of Toolkits for Visualizing Information about Performance of Earthwork Activities on a Construction Job Site." *Computing in Civil and Building Engineering*, Proceedings of the International Conference, vol 30.
- Azimi, R., Lee, S., AbouRizk, S.M., Alvanchi, A. (2011). "A Framework for An Automated and Integrated Project Monitoring and Control System for Steel Fabrication Projects." *Automation in Construction*, 20(1), pp 88-97.

- Balfanz, D., Schirmer, Grimm, M., Tazari, M. (2003). "Mobile Situation-Awareness within the Project Map." *Computers and Graphics*, 27(6), pp 893-898.
- Ballard, G. and Howell, G. (1994). "Implementing Lean Construction: Improving Downstream Performance." *2nd Conf. on Lean Const.*, Catolica Universidad de Chile, Santiago, Chile, pp 111-125.
- Barbacci, B., Klein, M.H., Longstaff, T.A., Weinstock, C.B. (1995). "Quality Attributes." *Technical Report CMU/SEI-95-TR-021*, 1995.
- Bartke, K. (2005). "2D, 3D, and High-Dimensional Data and Information Visualization." *Seminar on Data and Information Management*.
- Bassioni, H.A., Price, A.D.F. and Hassan, T.M. (2004). "Theoretical Formulation of a Framework for Measuring Business Performance in Construction." in Ruddock, L., Sexton, M., Amaratunga, D., Aouad, G. and Kagioglou, M. (Eds), *4th International Postgraduate Research Conference*, 1-2 April, University of Salford, Salford, pp 419-30.
- Beatham, S., Anumba, C., and Thorpe, T. (2004). "KPIs: A Critical Appraisal of their use in Construction." *Benchmarking Int. J.*, 11(1), pp 93–117.
- Bertini, E., and Santucci, G. (2004). "Quality Metrics for 2D Scatterplot Graphics: Automatically Reducing Visual Clutter." *Proceedings of Smart Graphics'04* Banff, Canada, p 77-89, 2004, Springer Verlag
- Beyer, H. and Holtzblatt, K. (1998). "Contextual Design: Defining Customer-centered Systems." Morgan Kaufmann Publishers, San Fransisco, CA.

- Bohn, J.S., and Teizer, J. (2010). "Benefits and barriers of Construction Project Monitoring Using High-Resolution Automated Cameras." *Journal of Construction Engineering and Management*, 136(6), pp 632-640.
- Boriah, S., Chandola, V., and Kumar, V. (2008). "Similarity Measures for Categorical Data: A Comparative Evaluation." *Proceedings of the 8th SIAM International Conference on Data Mining*, Atlanta, Georgia, USA, Apr., pp 243–254. Society for Industrial and Applied Mathematics (SIAM).
- Botero, L.F., and Alvarez, M.E. (2005). "Last Planner: an Advance in Planning and Controlling Construction Projects. Case Study in the City of Medellin." *Proceedings of the 4th Brazilian Symposium on Construction Management and Economics*, Porto Alegre, 24-26 October, 2005.
- Brown, C.E. and O’Leary, D.E. (2000). "Expert Systems." From: <http://carfield.com.hk/document/ai/Expert+Systems.html>, <Last Accessed: October 2012>.
- Bryde, D. J. (2003). "Project Management Concepts, Methods and Application." *International Journal of Operations and Production Management*, 23(7), pp 775–793.
- Burkhard, R.A., G. Andrienko, N. Andrienko, J. Dykes, A. Koutamanis, W. Kienreich, R. Phaal, A. Blackwell, M. Eppler, J. Huang, M. Meagher, A. Grün, S. Lang, D. Perrin, W. Weber, A. Vande Moere, B. Herr, K. Börner, J.-D. Fekete and D. Brodbeck (2007). "Visualization Summit 2007: Ten research goals for 2010." *Information Visualization*, vol 6, pp 169-188.
- Buzzsaw (Buzzsaw, 2010), www.buzzsaw.com, last accessed <01/20/2013>

- Cadiz, J.J., Venolia, G., Jancke, G., Gupta, A. (2002). "Designing and Deploying an Information Awareness Interface." *CSCW'02*, pp 314-323, November 16-20, 2002, New Orleans, Louisiana, USA.
- Card, S. (2008). "Information Visualization." In: Jacko, J., Sears, A. *Handbook of Human-Computer Interaction*. Lawrence Erlbaum Associates, pp 509-543.
- Chambers, J. M., Cleveland, W. S., Kleiner, B. (1983). "Graphical Methods for Data Analysis." *Wadsworth statistics/probability series*, Duxbury Press, Boston.
- Chan, A.P.C., and Chan, A.P.L. (2004). "Key Performance Indicators for Measuring Construction Success." *Benchmarking: An International Journal*, 11(2), pp 203-321.
- Cheah, C.Y.J., Garvin, M.J., Miller, J.B. (2004). "Empirical Study of Strategic Performance of Global Construction Firms." *Journal of Construction Engineering and Management*, 130(6), pp 808-817.
- Cheng, J., Proverbs, D.G., Oduoza, C. and Fleming, C. (2005). "A Conceptual Model towards the Measurement of Construction Client Satisfaction", 21st Annual ARCOM Conference, SOAS, University of London, vol 2, pp 1053-1062.
- Chengzhi, Q., Chenghu, Z., Tao, P. (2003). "Taxonomy of Visualization Techniques and Systems - Concerns between users and developers are different." *Asia GIS Conference*, 2003.
- Cheung, S. O., Suen, H. C. H. and Cheung K. K. W. (2004). "PPMS: A Web-based Construction Project Performance Monitoring System." *Automation in Construction*, vol 13, pp 361-376.

- Chewning, E.C.Jr., and Harrell, A.M. (1990). "The Effect of Information Load on Decision Makers" cue Utilization Levels and Decision Quality in a Financial Distress Decision Task." *Accounting, Organizations and Society*, 15(6), pp 527-542.
- Chin, G.J., Singhal, M., Nakamura, G., Gurumoorthi, V., Freeman-Cadoret, N. (2009). "Visual Analysis of Dynamic Data Streams." *Information Visualization*, 8 (3) p 18.
- Chiu, C. (2011). "Visualization of Construction Management Data." PhD Thesis, Civil Engineering, University of British Columbia.
- Chiu, C. and Russell, A.D. (2011). "Design of a Construction Management Data Visualization Environment: A Top Down Approach." *Automation in Construction*, vol 20, pp 399-417.
- Chou, J., and Chong, W. K., (2008). "A Web-based Framework of Project Performance and Control System", *Robotics, Automation and Mechatronics, 2008 IEEE Conference on*, pp 803-807, 21-24 Sept. 2008.
- Chuah, M.C., and Roth, S.F. (2002). "Visualizing Common Ground." *Proceedings of Information Visualization*.
- Cleveland, W.S., McGill, R. (1984). "Graphical Perception: Theory, Experimentation and Application." *Journal of the American Statistical Association* 79 (387) p 24.
- Cleveland, W.S., McGill, R. (1987). "Graphical Perception: The Visual Decoding of Quantitative Information on Graphical Displays of Data." *Journal of the Royal Statistical Society*, 150 (3), pp 39.
- Construction Consultants KPIs Working Group (2007). "UK Construction Consultants Key Performance Indicators – Handbook." <http://www.constructingexcellence.org.uk/zones/>

kpizone/resources/data/2007/consultants/consult_handbook07.pdf <last accessed 01/20/2013>

Construction Institute (2005). "Performance Metric Formulas and Definitions." http://construction-institute.org/scriptcontent/toolkit/Metrics_Definitions_04012005.pdf.

Constructware (Constructware, 2011), <http://usa.autodesk.com/adsk/servlet/pc/index?id=6871224&siteID=123112> last accessed <01/20/2013>.

Costello, L., Grinstein, G., Plaisant, C., and Scholtz, J. (2009). "Advancing User-centered Evaluation of Visual Analytic Environments through Contests." *Information Visualization*, 8(3), pp 230-238.

Cox, R.F., Issa, R.R.A., and Ahrens, D. (2003). "Management's Perception of Key Performance Indicators for Construction." *J. of Const. Engineering and Management*, 129(2), pp 142-151.

Dainty, A.R., Cheng, M. and Moore, D.R. (2003). "Redefining Performance Measures for Construction Project Managers: an Empirical Evaluation." *Construction Management and Economics*, vol 21, pp 209-218.

Dawood, N., Sikka, S., Marasini, R. and Dean, J. (2006). "Development of Key Performance Indicators to Establish the Benefits of 4D Planning." *Proceedings of 22nd ARCOM Conference*, Editors: D. Boyd, Birmingham, UK, pp 709-718.

Doolin, D. (2010) "Defining Static vs. Dynamic Information on Your Website." <http://website-in-a-weekend.net/website-design/static-dynamic-information/>, <last visited October 29, 2010>.

- de la Garza, J.M., Howitt, I. (1998) "Wireless Communications and Computing at the Construction Job Site." *Automation in Construction*, 7(4), pp 327-347.
- de Wit, A. (1998) "Measurement of Project Success." *International Journal of Project Management*, 6(3), pp 164-170.
- DETR: Department of the Environment, Transport, and the Regions (2000), KPI Rep. for the Minister for Construction, KPI Working Group, London.
- Deru, M., and Torcellini, P. (2005) "Source Energy and Emission Factors for Energy Use in Buildings." *National Renewable Energy laboratory (NREL) Technical Report NREL/TP-550-38617*, Revised June 2007 Colorado.
- Dull, R.B., and Tegarden, D.P. (1999). "A Comparison of Three Visual Representations of Complex Multidimensional Accounting Information." *Journal of Information Systems* 13(2), pp 117-131.
- Elmqvist, N. and Fekete, J., (2010). "Hierarchical Aggregation for Information Visualization: Overview, Techniques and Design Guidelines." *IEEE Transactions on Visualization and Computer Graphics*, 16 (3), p 15.
- Engelhardt, J. (2002). "The Language of Graphics: A Framework for the Analysis of Syntax and Meaning in Maps, Charts and Diagrams." *ILLC Dissertations*, Amsterdam: University of Amsterdam.
- Eppler M.J., and Burkhard, R.A. (2007). "Visual Representations in Knowledge Management: Framework and Cases", *Journal of Knowledge Management*, 11(4), pp 112-122.
- Fayyad, U., Grinstein, G.G., Wierse, A. (2002). "Information Visualization in Data Mining and Knowledge Discovery." Morgan Kaufmann, 2002.

- Fry, B. (2007). "Visualizing Data: Exploring and Explaining Data with the Processing Environment." O'Reilly Media, Incorporated.
- Fouskas, K.G., and Pateli, A.G. (2002). "Applying Contextual Inquiry for Capturing End-Users Behaviour Requirements for Mobile Exhibition Services." *1st International Conference on Mobile Business*, July 2002.
- Ganah, A.A., Bouchlaghem, N.B., and Anumba, C.J. (2010). "VISCON: Computer Visualization Support for Constructability." *ITCON*, vol 10, pp 69- 89.
- Gershon, N., Eick, S.G., Card, S. (1998). "Information Visualization." *Interactions*, 5(2), pp 9-15
- Glossary:D, Definition of "Dynamic", <http://www.bbn.com/utility/glossary/d>, <last visited October 29, 2010>.
- Golparvar-Fard, M.G., and Pena-Mora, F. (2007). "Application of Visualization Techniques for Construction Progress Monitoring." *Computing in Civil Engineering*, L. Soibelman and B. Akinci, eds., ASCE, Reston, VA., p 27.
- Golparvar-Fard, M., Pena-Mora, F., Arboleda, C.A., and Lee, S. (2009). "Visualization of Construction Progress Monitoring with 4D Simulation Model Overlaid on Time-Lapsed Photographs." *Journal of Computing in Civil Engineering*, 23(6), pp 391-404.
- Grinstein, G.G., Trutschl, M., Cvek, U. (2001). "High-dimensional Visualizations." A. Keim, S. Eick (Eds.), *Workshop on Visual Data Mining*, San Fransisco, KDD, 2001.
- Hajjar D., and AbouRizk, S.M. (2000). "Integrating Document Management with Project and Company Data." *Journal of Computing in Civil Engineering*, 14:1(70).
- Haksever, A.M. (2000). "A Model to Predict the Occurrence of Information Overload of Project Managers." *Proceedings of International Conference on Construction*

- Information Technology 2000: Taking the Construction Industry into the 21st Century*, CIB-W78, 28-30 June, vol1, Reykjavik, Iceland, 415-427.
- Harris, R.L. (2000). "Information Graphics: A Comprehensive Illustrated Reference," Oxford University Press, USA.
- Hartman, T., Gao, J., and Fischer, M. (2008). "Areas of Application for 3D and 4D Models on Construction Projects," *Journal of Construction Engineering and Management*, 134(10), pp 776-785.
- Hays, B.J. (2002) "A Graphical Representation Framework for Enhanced Visualization of Construction Control Processes." Master's Thesis, Civil Engineering Department, Virginia Tech.
- Heer, J., Card, S.K., and Landay, J.A (2005). "Prefuse: a Toolkit for Interactive Information Visualization." In: CHI 2005, Portland, Oregon, USA.
- Heer, J., Kong, N., and Agrawala, M. (2009). "Sizing the Horizon: the Effects of Chart Size and Layering on the Graphical Perception of Time Series Visualization." *Proceedings of the ACM CHI 2009 Conference on Human Factors in Computing Systems*, pp 1303–1312, 2009.
- Helfman, J.I. and Goldberg, J.H. (2007) "Selecting the Best Graph based on Data, Tasks, and User Roles," UPA Conference Patterns: Blueprints for Usability, June 11-15, 2007, Austin, Texas, USA.
- Helfman, J.I. and Goldberg, J.H. (2010) "Graph Selection Method," Patent ID: 7705847.
- Hiremath, H.R., Skibniewski, M.J. (2004) "Object Oriented Modeling of Construction Process by Unified Modeling Language" *Automation in Construction*, 13(4).

- Hoffman, P.E. and Grinstein, G.G, (2002). “A survey of visualizations for high-dimensional data mining.” Fayyad, U., Grinstein, G., Wierse, A., eds. *Information Visualization in Data Mining and Knowledge Discovery*. Ney York: Morgan Kaufmann, pp 47-82.
- Horn, R. E. (2001). “Visual Language and Converging Technologies in the Next 10-15 Years.” *National Science Foundation Conference on Converging Technologies for Improving Human Cognition*, pp 124-132. December 3. www.stanford.edu/~rhorn.
- Huang, D., Tory, M., Staub-French, S., Pottinger, R. (2009). “Visualization Techniques for Schedule Comparison.” *Computer Graphics Forum*, 28(3), pp 951-958. Blackwell Publishing Ltd.
- Huhnt, W. and Richter, S. (2010). “Visualization of Construction Process Models”, *Proceedings of the International Conference on Computing in Civil and Building Engineering*, Nottingham UK, Nottingham University Press.
- Icke I., and Sklar E. (2009). “Visual Analytics: A Multi-faceted Overview.” *Technical Report #TR-2009005*, Dept of Computer Science, The Graduate Center, City University of New York.
- IFC, <http://buildingsmart.com/standards/ifc> <last accessed 06/30/2011>.
- InfoVis (InfoVis, 2005), <http://ivtk.sourceforge.net/> <last accessed 01/20/2013>.
- Jankun-Kelly, T. J. (2003). “Visualizing Visualization A Model and Framework for Visualization Exploration Computer Science.” California, University of California. Doctor of Philosophy p 134.
- Johansson, J. (2008). “Efficient Information Visualization of Multivariate and Time-Varying Data.” PhD Thesis, Linköping University.

- Jung, V. (1999) "MetaViz: Visual Interaction with Geospatial Digital Libraries.", *International Computer Science Institute*, Technical Report, TR-99-017.
- Jung, Y. and S. Woo (2004). "Flexible Work Breakdown Structure for Integrated Cost and Schedule Control." *Journal of Construction Engineering and Management* 130(5): p 10.
- Kasik, D.J., Ebert, D., Lebanon, G., Park, H., Pottenger, W.M. (2009). "Data Transformations and Representations for Computation and Visualization Information." *Visualization*, 8 (4) p 11.
- Kagioglou, M., Cooper, R., and Aouad, G. (2001). "Performance Management in Construction: A Conceptual Framework." *Constuction Management Economics*, vol 19, pp 85–95.
- Kang, L.S., and Paulson, B.C. (1998). "Information Management to Integrate Const and Schedule for Civil Engineering Projects." *Journal of Construction Engineering and Management*, 124(5), pp 381-389.
- Kasik, D.J., Ebert, D., Lebanon, G., Park, H., Pottenger, W.M., (2009). "Data Transformations and Representations for Computation and Visualization." *Information Visualization*, 8 (4) pp 11.
- Keim, D.A, (2002). "Information Visualization and Visual Data Mining." *IEEE Trans. on Vis. and Comp. Graphics*, pp 8, 1-8.
- Keim,D., Andrienko,G., Fekete,J.-D, Kohlhammer,J. (2008). "VisualAnalytics: Definition, Process, and Challenges." In: Kerren,A.,Stasko, J.T., Fekete, J.-D., North, C.J. (eds.) *Information Visualization*, LNCS,vol 4950, Springer, Heidelberg.

- Ko, C., Cheng, M. and Wu, T. (2007). "Evaluating Sub-contractors Performance Using EFNIM." *Automation in Construction*, vol 16, pp 525-530.
- Kog, Y.C., Chua, D.K.H., Loh, P.K. and Jaselskis, E.J. (1999) "Key Determinants for Construction Schedule Performance." *International Journal of Project Management*, 17(6), pp 351-359.
- Korde, T., Li, M. and Russell, A.D. (2005). "State-of-the-art Review of Construction Performance Models and Factors." *Proceed of ASCE Constr. Congr.*, Editors: I. Tommelein, San Diego, CA, USA.
- Kosara, R., Hauser, H., and Gresh, D.L. (2003). "An Interaction View on Information Visualization." *Eurographics*, pp127-133.
- KPI Working Group (2000). "KPI Report for the Minister for Construction." Online at <http://www.berr.gov.uk/files/file16441.pdf>. <Last accessed on October 17, 2008>.
- Kumar, N. and Benbasat, I. (2004). "The Effect of Relationship Encoding, Task Type, and Complexity on Information Representation: An Empirical Evaluation of 2D and 3D Line Graphics." *MIS Quarterly*, 28(2), pp 255-281.
- Kuo, C., Tsai, M., Kang, S., Hsieh, S. (2009) "Construction Dashboard: An Exploratory Information Visualization Tool for Multi-system Construction." *9th International Conference on Construction Applications of Virtual Reality*, November 5-6, 2009.
- Kuo, C. H., Tsai, M. H., and Kang, S. C. (2011). "A Framework of Information Visualization for Multi-System Construction." *Automation in Construction*, 20(3), 247-262.

- Lee, N. and Rojas, E. (2009) "Developing Effective Visual Representations to Monitor Project Performance." *Proceedings of the 2009 Construction Research Congress*, Seattle, Washington, April, pp 826-835.
- Lengler, R., and Eppler, M. J. (2007) "Towards a Periodic Table of Visualization Methods for Management," *IASTED Proceedings of the Conference on Graphics and Visualization in Engineering (GVE 2007)*, Clearwater, Florida, USA.
- Leung, N., Chan, S., and Issa, R.R.A. (2003). "Meta-data-based Collaboration in Construction Project Management." *Proceed. of Towards a Vision for Info. Tech. in Civil Eng.*, Editors: I. Flood, Nashville, Tennessee, pp 1-12.
- Li, J., Martens, J., and van Wijk, J. (2010). "Judging Correlation from Scatterplots and Parallel Coordinate Plots." *Information Visualization*, 9(1), pp 13-30.
- Li, M. and Russell, A.D. (2006). "Diagnosing Construction Performance." *Joint International Conference on Computing and Decision Making in Civil Engineering* Montreal, Canada.
- Li, Q., and North, C. (2003). "Empirical Comparison of Dynamic Query Sliders and Brushing Histograms." *Information Visualization*, pp 147-153. 21 Octobre 2003, Seattle, WA, USA.
- Lin, G. and Shen, Q. (2007), "Measuring the Performance of Value Management Studies in Construction: Critical Review." *Journal of Management in Engineering*, 23(1), pp 2-9.
- Liston, K., Fischer, M. and Kunz, J. (2000). "Designing and Evaluating Visualization Techniques for Construction Planning." *Proceedings of the Eight International*

- Conference on: Computing in Civil and Building Engineering (ICCCBE-VIII)*, Aug 14-16 2000; Stanford, CA, v.2, 1293-1300.
- Liston, K., Fischer, M. and Winograd, T. (2001). "Focused Sharing of Information for Multi-Disciplinary Decision Making by Project Teams." *ITcon*, vol 6, pp 69-82.
- Lohse, G.L., Biolsi, K., Walker, N., Rueter, H.H. (1994). "A Classification of Visual Representations." *CACM* 37, (12), pp 14.
- Mahmoud, S.Y. and Scott, S. (2002). "The Development and Use of Key Performance Indicators by the UK Construction Industry." *18th Annual ARCOM Conference*, Editors: D. Greenwood, Univ. of Northumbria, vol 2, pp 587-594.
- Mackinlay, J., Hanrahan, P., Stolte, C. (2007) "Show me: Automatic Presentation for Visual Analysis." *IEEE Transactions on Visualization and Computer Graphics*, 13(6), pp 1137-1144.
- Mao, W., Zhu, Y. and Ahmad, I. (2007) "Applying Metadata Models to Unstructured Content of Construction Documents: A view based approach." *Automation in Construction*, 16(2), pp 242-252.
- McCullough, B. (1997). "Automating Field Data Collection in Construction Organizations." *4th ASCE Construction Congress* Minneapolis, Minnesota.
- Mehrotra, S., Butts, C., Kalashnikov, D., Venkatasubramanian, N., Rao, R., Chockalingam, G., Eguchi, R., Adams, B., and Huyck, C. (2004). "Project Rescue: Challenges in Responding to the Unexpected." *Proceedings of 16th Annual Symposium on Electronic Imaging Science and Technology*, San Jose, CA, January 2004.

- Memon, Z.A., Majid, M.Z.A, and Mustaffar, M. (2007). "A Systematic Approach for Monitoring and Evaluating the Construction Project Performance." *The Journal of the Institution of Engineers, Malaysia* 67(3), pp 26-32.
- Meyer, M., T. Girba, et al. (2006). "Mondrian: An Agile Information Visualization Framework." *ACM Symposium on Software Visualization*.
- Mohamed, S. (2003). "Scorecard Approach to Benchmarking Organizational Safety Culture in Construction." *Journal of Construction Engineering and Management*, 129(1), pp 80-88.
- MS Project (MS Project, 2008), www.microsoft.com/project/ last accessed <01/20/2013>
- Myatt, G.J. and Johnson, W.P, (2009). "Making Sense of Data II: A Practical Guide to Data Visualization, Advanced Data Mining Methods, and Applications." New Jersey, Wiley.
- Navon R., Sacks, R. (2007). "Assessing Research Issues in Automated Project Performance Control (APPC)." *Automation in Construction*, 16(4), pp 474-484.
- Newby, G.B. (2002). "Empirical Study of a 3D Visualziation for Information Retrieval Tasks." *Journnal of Intelligent Information Systems*, vol 18, pp 31-53.
- O'Shaugnessy, J. (1972). "Inquiry and Decisions." Allen & Unwin, London.
- Pena-Mora, F. and Li, M. (2001). "Dynamic Planning and Control Methodology for Design/Build Fast-Track Construction Projects" *Journal of Construction Engineering and Management*, 127(1) pp 1-17.
- Pfitzner, D., Hobbs, V. and Powers, D. (2000). "A Unified Taxonomic Framework for Information Visualization." J. Weckert (Ed.), *2nd Australian Institute of Computer Ethics*, vol 1, Canberra 2000.

- Piccolo (Piccolo, 2005), www.cs.umd.edu/hcil/jazz/ last accessed <01/20/2013>
- Pike, W.A., Stasko, J., Chang, R., and O'Connell, T.A. (2009). "The Science of Interaction." *Information Visualization*, vol 8, pp 263-274.
- Pirolli, P. and Card, S. (2005). "The Sensemaking Process and Leverage Points for Analyst Technology as Identified Through Cognitive Task Analysis." *Proceedings of Analyst*, pp 2-4.
- Pradhan, A. (2009) "An Approach for Using Data from Multiple Sources to Support Construction Productivity Analyses." PhD thesis, Department of Civil and Environmental Engineering, Carnegie Mellon University.
- Prefuse (Prefuse, 2007), www.prefuse.org last accessed <01/20/2013>
- Prieto-Diaz, R. (1991) "Implementing Faceted Classification for Software Reuse," *Communications of the ACM, Special Issue on Software Engineering*, 34(5), pp 88-97.
- Primavera (Primavera, 2009), <http://www.oracle.com/us/products/applications/primavera/overview/index.html> last accessed <01/20/2013>
- Quass, D., Rajaraman, A., Sagiv, Y., Ullman, J., Widow, J. (1995). "Querying Semistructured Heterogeneous Information." *Lecture Notes in Computer Science, Deductive and Object Oriented Databases*, vol 1013, pp 319-344.
- Quinian, J.R. (1986). "Induction of Decision Trees." *Machine Learning*, 1(1), pp 81-106.
- Reinhardt, J., Akinci, B., and Garrett, J.H.Jr, (2005). "Framework for Providing Customized Data Representations for Effective and Efficient Interaction with Mobile Computing Solutions on Construction Sites." *Journal of Computing in Civil Engineering*, 19(2), pp 109-118.

- Rester, M., Pohl, M., Hinum, K., Mikscg, S., Ohmann, S., Popow, C., and Banovic (2005). "Assessing the Usability of an Interactive Information Visualization Methods as the First Step of A Sustainable Evaluation.", *Empowering Software Quality*, pp 31-44.
- Robertson, G.G., Mackinlay, J.D., and Card, S.k. (1991). "Cone Trees: Animated 3D Visualizations of Hierarchical Information." *Proceedings of the ACM Conference on Human Factors in Computing Systems* (CHI '91); 1991 April 27 - May 2; New Orleans; LA. NY: ACM; 1991; 189-194.
- Robinson H.S., Carrillo, P.M., Anumba C.J. and Al-Ghassani, A.M. (2001). "Perceptions And Barriers In Implementing Knowledge Management Strategies In Large Construction Organisations." *Proc., RICS Foundation Construction and Building Research Conference* (COBRA) 2001 Conference, Glasgow Caledonian University, UK, 3 - 5 September, 2, pp 451-460.
- Robinson, H.S., Anumba, C.J., Carrillo, P.M. and Al-Ghassani, A.M. (2005). "Business Performance Measurement Practices in Construction Engineering Organizations." *Measuring Business Excellence*, 9(1), pp 13-22.
- Rohn, S., Aziz, Z., and Pena-Mora, F. (2010). "An Object-based 3D Walk-through Model for Interior Construction Progress Monitoring." *Automation in Construction*.
- Russell, A.D.; Chiu, C.-Y.; and Korde, T. (2009). "Visual Representation of Construction Management Data." *Automation in Construction* 18(8), pp 1045-1062.
- Russell, A.D., and Udaipurwala, A. (2003). "Construction Schedule Visualization" *Proceedings of the 2002 Information Technology in Civil Engineering Symposium*,

- eds. Thorton, S.E., Padgett, J.A., Schexnayder, C. J., Washington, District of Columbia, November 2-3, 2002, ASCE, Reston, Virginia, 12p.
- Russell S.J., and Norvig P. (2002). "Artificial Intelligence: A Modern Approach." Prentice Hall, Englewood Cliffs, NJ.
- Sanvido, V., Grobler, F., Parfitt, K.; Guvenis, M. and Coyle, M. (1992). "Critical Success Factors for Construction Projects." *Journal of Constr. Engineering and Management*, 118(1), pp 94-111.
- Schonlau, M. (2003). "Visualizing Categorical Data Arising in the Health Sciences using Hammock Plots." *Proceedings of the Section and Statistical Graphics, American Statistical Association*.
- Schreck, T., von Landesberger, T., Bremm, S. (2010). "Techniques for Precision-based Visual Analysis of Projected Data," *Information Visualization*, 9(3), pp 181-193.
- Sedig and Sumner (2006). "Characterizing Interaction with Visual Mathematical Representations." *International Journal of Computers for Mathematical Learning*, 11, pp 1-55.
- Shen, L., Lu, W., Yao, H. and Wu, D. (2005). "A Computer-based Scoring Method for Measuring Environmental Performance of Construction Activities." *Automation in Construction*, 14(3), pp 297-309.
- Shields, M.D. (1980) "Some Effects on Information Load on Search Patterns Used to Analyze Performance Reports." *Accounting, Organizations, and Society*, 5(4), 429-442.

- Shneiderman, B. (1996). "The eyes have it: A Task by Data Type Taxonomy for Information Visualization." *IEEE Computer Society, IEEE Symposium on Visual Languages*, 1996, p. 336.
- Shriniasan, Y.B., and van Wijk, J.J. (2008). "Supporting the Analytical Reasoning Process in Information Visualization." *Proceedings of CHI'08*, pp 1237-1246.
- Snowball, D. (1980). "Some Effects of Accounting Expertise and Information Load: An empirical study." *Accounting, Organizations and Society*, 5(3), pp 323-338.
- Sohail, A., Miles, D.W. and Cotton, A.P. (2002). "Developing Monitoring Indicators for Urban Micro Contracts in South Asia." *International Journal of Project Management*, vol 20, pp 583-591.
- Soibelman, L., and Caldas, C. (2000). "Information Logistics for Construction Design Team Collaboration." *Proceedings of Computing in Civil and Building Engineering*, pp 341-348.
- Song, K., Pollalis, S.P., and Pena-Mora, F. (2005). "Project Dashboard: Concurrent Visual Representation Method of Project Metrics on 3D Building Models." *Proceed. of the ASCE Int. Conference on Computing in Civil Engineering*, Editors: L. Soibelman and P. Feniosky, Cancun, Mexico, pp 147.
- Songer, A.D. and Hays, B. (2003). "A Framework for Multi-dimensional Visualization of Project Control Data." *Proceedings of Construction Research Congress 2003*, Honolulu, Hawaii, USA, 121-130.
- Songer, A.D., Hays, B., and North, C. (2004). "Multidimensional Visualization of Project Control Data." *Construction Innovation: Information, Process, Management*, 4(3), pp 173-190.

- Songer, A.D., Subramanian, P.S., Diekmann, J.E. (2000). "Integrating Work Flow Modeling and 3D CAD for Visualizing Project Performance." *CIB W78*.
- Speier, C., Valacich, J.S., Vessey, I. (1999). "The Influence of Task Interruption on Individual Decision Making: An Information Overload Perspective." *Decision Sciences*, 30(2), pp 337-360.
- Spence, R. (2007) "Information Visualization: Design for Interaction." 2nd Edition, *Prentice Hall, Pearson Education Limited*, Essex, England.
- Srivastav, S., Lannon, S., Alexander, D. K., Jones, P. (2009). "A Review and Comparison of Data Visualization Techniques Used in Building design and Building simulation." *Eleventh International IBPSA Conference*.
- Stasko, J., Gorg, C., and Liu, Z. (2008). "Jigsaw: Supporting Investigative Analysis Through Interactive Visualization." *Information Visualization* vol 7, pp 118-132.
- Staub-French, S., Fischer, M., Kunz, J., and Paulson, B. (2003). "A Generic Feature-driven Activity-based Cost Estimation Process." *Adv. Eng. Inf.*, 17(1), pp 23-29.
- Stewart, R.A., and Mohamed, S. (2001). "Utilizing the Balanced Scorecard for IT-IS Performance Evaluation in Construction." *Construction Innovation: Information, Process, Management*, 1(3), pp 147-163.
- Stolte, C., Tang, D., Hanrahan, P. (2002). "Query, Analysis, and Visualization of Hierarchically Structured Data Using Polaris" *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '02)*, ACM, New York, NY, USA, 112-122.
- Stolte, C., Tang, D., Hanrahan, P. (2003). "Multiscale Visualization Using Data Cubes." *Visualization and Computer Graphics, IEEE Transactions*, 9(2), pp 176- 187.

- Svahanberg, M. Wohlin, C., Lundberg, L., and Mattsson, M. (2002). "A Method for Understanding Quality Attributes in Software Architecture Structures." *Proceedings of the 14th International Conference on Software Engineering and Knowledge Engineering*, pp 819-826, Ischia, Italy, 2002.
- Swarnadhipathi, K., and Boyd, D. (2007). "The Relationship between Theory, Methodology, and Construction Practice in Research into Business Performance of Medium-size Contractors." *Association of Researchers in Construction Management (ARCOM) Doctoral Workshop*, March 05, 2008, University of Wolverhampton, UK.
- Takim R. and Akintoye A. (2002). "Performance Indicators for Successful Construction Project Management." *Proc. of 18th Conf. of the Association of Researchers in Construction Management*, Editors: D.J. Greenwood, University of Northumbria at Newcastle upon Tyne, vol2, pp 545–555.
- Thorpe, T. and Mead, S. (2001). "Project Specific Web Sites: Friend or Foe." *Journal of Construction Engineering and Management*, 127(5), pp 406-413.
- Thorpe, T. and J. Murray (1996). "COMPOSITE: Site Communications Survey, Official report prepared for DOE, Department of Civil and Building Engineering." Loughborough University, United Kingdom.
- Thakur, S. and Rhyne, T.M. (2009). "Data vases: 2D and 3D Plots for Visualizing Multiple Time Series." *Advances in Visual Computing*, 929-938.
- Tidwell, J. (2005) "Designing interfaces," O'Reilly Media, Incorporated.
- Tidwell, J. (2006) *Designing Interfaces*, O'Reilly Media, CA, 2006.
- Tominski, C., Abello, J., Schumann, H. (2004). "Axes-based Visualizations with Radial Layouts." *ACM symposium on Applied Computing*, pp 1242-1247.

- Tory, M., Staub-French, S., (2008). "Qualitative Analysis of Visualization: A Building Design Field Study." *Proceedings of the Conference on Beyond Time and Errors: Novel Evaluation Methods for Information Visualization*, April 5, 2008, Florence, Italy.
- Trickett, S. and Trafton, J. (2006). "Toward a Comprehensive Model of Graph Comprehension: Making the Case for Spatial Cognition." *Diagrammatic representation and Inference*, pp 286-300.
- Turban, E., Aronson, J., Liang, T.P. (2001). "Decision Support Systems and Intelligent Systems." Pearson Prentice Hall.
- Vico (Vico, 2010), www.vicosoftware.com last accessed <01/20/2013>
- Watson, P. (2002). "Developing an Efficient and Effective Control System." *Building Engineering*, 77(3), pp 28-36.
- Weber, M., Alexa, M., Müller, W. (2001). "Visualizing Time-series on Spirals." *Proceedings of the IEEE Symposium on Information Visualization*, p. 7.
- Wehrend, S. and Lewis, C. (1990). "A Problem Oriented Classification of Visualization Techniques." *VIS '90 Proceedings of the 1st conference on Visualization*.
- White, J. V. (1984). "Using Charts and Graphs: 1000 Ideas for Visual Persuasion." Bowker.
- Wickham, H. (2009). "A Layered Grammar of Graphics." *Journal of Computational and Graphical Statistics* 19(1), p 26.
- Wilkinson, L and Wills, G. (2005). "The Grammar of Graphics." Springer. Chicago, IL, USA.

- Wills, G. and Wilkinson, L. (2010). "AutoVis: Automatic Visualization," *Information Visualization*, vol 9, pp 47-69.
- Wills, G. (2012). "Visualizing Time: Designing Graphical Representations for Statistical Data." Springer.
- Wu, I. and Hsieh, S. (2012). "A Framework for Facilitating Multi-dimensional Information Integration, Management, and Visualization in Engineering Projects." *Automation in Construction*, vol 23, pp 71-86.
- XMDV (xmdv, 2010), www.davis.wpi.edu/xmdv/ last accessed <01/20/2013>
- Yang, H., Yeung, J.F.Y., Chan, A.P.C., Chiang, Y.H., and Chan, D.W.M. (2010). "A Critical Review of Performance Measurement in Construction." *Journal of Facilities Management*, 8(4), pp 269-284.
- Yang, J. and Peng, S. (2008). "Development of a Customer Satisfaction Evaluation Model for Construction Project Management." *Building and Environment*, 43(4), pp 458-468.
- Yang, J., Ward, M.O., and Rundensteiner, E.A. (2003). "Hierarchical Exploration of Large Multivariate Data Sets." pg. 201 in *Data Visualization: the state of the art*, Kluwer Academic Publishers, London.
- Ye, J.D.; and Froese, T. (2009). "Visualization Configuration Model for Integrating Presentation of Construction Management Data." *Proceedings of the 2009 Construction Research Congress*, Seattle, Washington, April, 979-987.
- Young, F. W., Valero-Mora, P. M., Friendly, M. (2006). "Visual Statistics: Seeing Data with Dynamic Interactive Graphics." Wiley-Interscience.

- Yu, I., Kim, K., Jung, Y., Chin, S. (2007). "Comparable Performance Measurement System for Construction Companies." *Journal of Management in Engineering*, 23(3), pp 131-139.
- Zantout, H. and Marir F. (1999). "Document Management Systems from Current Capabilities towards Intelligent Information Retrieval: an Overview." *International Journal of Information Management*, 19(6), 471-484.
- Zeb, J.; Chiu, C.-Y. and Russell A.D. (2008). "Designing a Construction Data Visualization Environment." *CSCE 2008 Annual Conference*, June 10-13 2008, Quebec, QC.
- Zhou, M.X. (1999). "Visual Planning: A Practical Approach to Automated Presentation Design" *Proceedings of the International Joint Conference on Artificial Intelligence 99 (IJCAI 99)*, pp 634-641.
- Zhou, M.X., and Feiner, S.K. (1998). "Visual Task Characterization for Automated Visual Discourse Synthesis." *Joint Eurpgraphics/IEEE TCVG Symposium on Visualization*, Los Angeles, CA, USA, pp 392-399.
- Zhu, Y. (2007). "Measuring Effective Data Visualization." *Lecture Notes in Computer Science, Advances in Visual Computing*, vol 4842, pp 652-661.

CHAPTER 7

7 Appendices

7.1 Appendix A: Performance Metric Categories and Related Performance Metrics Identified from Prior Research Studies

Appendix A is composed of eight tables representing the performance metrics identified from previous research studies. The tables A1 through A8 are separated based on the performance categories of time, cost, quality, safety, health and environment, communication of information, productivity, business performance and subjective measures.

Table A.1. Performance Metric Category: Time (Schedule) (Cheung et al. 2004; Korde et al. 2005; Construction Institute 2005; Alarcon and Ashley 1996; KPI Working Group 2000; Dawood et al. 2006; Mahmoud and Scott 2002; Takim and Akintoye 2002; Sinthawananarong 2000; Sohail et al. 2002)

Perf. Metric Category	Performance Metrics
Achievement of critical dates	Planned number of critical dates (Cheung et al. 2004), Actual number achieved on time (Cheung et al. 2004), % achieved on time (Cheung et al. 2004)
Achievement of milestones	Planned number of milestones (Cheung et al. 2004), Actual number achieved on time (Cheung et al. 2004), % achieved on time (Cheung et al. 2004), On time milestone completion (Cox et al. 2003)
Turn-around time for submissions in period	Average (days) (Cheung et al. 2004), Maximum (days) (Cheung et al. 2004)
Schedule growth factors	Project Schedule Growth ((Actual total project duration – initial predicted project duration)/initial predicted project duration) (Construction Institute 2005), Delta Schedule Growth (Schedule Growth), Project Schedule Factor (Actual total project duration/(initial predicted project duration + approved changes)) (Construction Institute 2005), Delta Schedule Factor (1-Schedule Factor) (Construction Institute 2005), Phase duration Factor (Actual Phase duration /actual total project duration) (Construction Institute 2005)
Total Project Duration	Actual total project duration (weeks) (Construction Institute 2005, KPI Working Group 2000)
Construction Phase Duration	Actual construction phase duration (weeks) (Construction Institute 2005)
Effectiveness	Duration of the start-up phase to full production (Alarcon and Ashley 1996)
Time predictability	Design phase (Takim and Akintoye 2002), Construction phase (Takim and Akintoye 2002, KPI Working Group 2000), Design and construction phases (KPI Working Group 2000), client change orders (KPI Working Group 2000), project leader change orders (KPI Working Group 2000)
Rectifying defects	Time to rectify defects (KPI Working Group 2000, Mahmoud and Scott 2002)
SPI	schedule performance index : BCWP/BCWS (Dawood et al. 2006)
Project Progress	Completed gross floor area (m2) divided by the total construction floor area specified on the outset of the model input (Sinthawananarong 2000)
Foreman delay survey	Total number of operative delay hours divided by the total number of the operative working hours acc to the master plan (Sinthawananarong 2000)
Time growth	Final contract duration / initial contract duration (Sohail et al. 2002)
Lead time	Lag time (between the establishment of requirement and the contract commencement) / contract period (Sohail et al. 2002)

Table A.2. Performance Metric Category: Cost (Cheung et al. 2004, Korde et al. 2005; Construction Institute 2005; Alarcon and Ashley 1996; KPI Working Group 2000; Dawood et al.2006; Mahmoud and Scott 2002; Takim and Akintoye 2002; Sinthawanarong 2000; Sohail et al. 2002; Cox et al. 2003); Financial/Profitability/Turnover (Construction Consultants Working Group 2007; Cheng et al. 2005; Takim and Akintoye 2002)

Perf.	Metric
Category	Performance Metrics
Interim payment	This month's certified value vs. accumulated certified (Cheung et al. 2004); Certified interim payment amount/submission amount(%)(Cheung et al. 2004);% of work done (compared with the certificate forecast in the budget estimate) (Cheung et al. 2004); % of the contract sum certified (Cheung et al. 2004)
Variation order	Certified value amount out of submission amount (%)(Cheung et al. 2004);Variations amount vs. original contract sum (Cheung et al. 2004); Estimated total variation order amount vs. variation order amount in the interim payment (Cheung et al. 2004), Certified variation order amount paid by daywork content vs. total certified variation amount (%)(Cheung et al. 2004)
Cost claims	Certified amount of additional claims vs. original contract sum (Cheung et al. 2004), Certified additional cost claims vs. submitted additional cost claims by contractor (Cheung et al. 2004), Certified extension of time out of original contract period (Cheung et al. 2004), Certified extension of time claims vs. extension of time claims by contractor (Cheung et al. 2004), Amount set-off, including liquidated damages from retention money (Cheung et al. 2004), Total duration of delay out of original contract period (Cheung et al. 2004), Average time taken to resolve a claim (days) (Cheung et al. 2004)
Final Forecast	Account
Cost factors	growth
Cost for construction	for
Cost predictability	

	construction phases (KPI Working Group 2000), client change orders (KPI Working Group 2000), project leader change orders (KPI Working Group 2000)
Rectifying defects	Cost to rectify defects (KPI Working Group 2000; Mahmoud and Scott 2002)
CPI	Cost performance index, BCWP/ACWP (Dawood et al. 2006)
Variance of labor	Decrease or increase in the budget in respect of labor cost (Sinthawanarong 2000)
Variance of material	Decrease or increase in the budget in respect of labor cost (Sinthawanarong 2000)
Variance of plant	Decrease or increase in the budget in respect of plant usage expenditures (Sinthawanarong 2000)
Cost growth	Final completion cost/ initial cost (Sohail et al. 2002)
Accuracy and proximity of the estimated	accuracy of preliminary technical estimates (preliminary estimates/engineering estimates) (Sohail et al. 2002), Proximity of the engineer's estimated cost and the initial contract cost (engineer's estimated cost/initial contract cost) (Sohail et al. 2002), Proximity of the engineer's estimated cost and the final contract cost (engineer's estimated cost/ final contract cost) (Sohail et al. 2002)
Cost Compares	Compare the current costs accrued to the budgeted costs in dollars (Cox et al.2003)
Profit index	Profit before tax and interest/value of sales*100 (Construction Consultants Working Group 2007)

Table A.3. Performance Metric Category: Quality (Cheung et al.2004, Korde et al.2005, KPI Working Group 2000, Sinthawananarong 2000, Takim and Akintoye 2002, Cox et al. 2003), Rework/Defects (Construction Institute 2005, Dawood et al. 2006, Takim and Akintoye 2002)

Perf.	Metric
Category	Performance Metrics
Inspections	Number of site inspections conducted (Cheung et at. 2004), Percentage of quality control inspections passed at the first time of the approval (Sinthawananarong 2000)
Non-compliance records	Number of non-compliance records received (Cheung et at. 2004), Number of non-compliance records closed (Cheung et at. 2004), Total number of non-compliance records rectified (Cheung et at. 2004), Average time to close out non-compliance records (Cheung et at. 2004), Cause and effect score calculated from the frequency of non-conformance cases modified by possible causes (%) (Sinthawananarong 2000)
Work	
Rejection	Due to lateness (Cheung et at. 2004), Due to workmanship (Cheung et at. 2004)
Survey (Samples)	
Rejections	Due to lateness (Cheung et at. 2004), Due to workmanship (Cheung et at. 2004)
Defects	1 to 10 scale ranging from defect free to totally defective (KPI Working Group 2000)
Outstanding quality issues	Number of issues at available for use (KPI Working Group 2000), Number of issues at the end of defect rectification period (KPI Working Group 2000)
Quality	
control/rework	The change in the number of man hours and material costs for repairing work in place or rehandling materials (Cox et al.2003)
Total filed rework factor	Total direct cost of field rework / actual construction phase cost (Construction Institute 2005)
Rework index	Number of errors –drawing (Dawood et al.2006), Number of corrections – drawing/design (Dawood et al.2006), Number of claims – quality (Dawood et al.2006), Number of planning clashes (Dawood et al.2006)

Table A.4. Performance Metric Category: Safety, Health and Environment (Cheung et al. 2004, Korde et al. 2005, Construction Institute 2005, KPI Working Group 2000, Construction Consultants Working Group 2007, Takim and Akintoye 2002, Sinthawanarong 2000, Cox et al. 2003)

Perf.	Metric	
Category		Performance Metrics
Statistics		Number of accidents reported (Cheung et at. 2004), Number of man-days lost (Cheung et at. 2004)
Monitoring and Compliance		Safe work practices (Cheung et at. 2004), Tools and machinery (Cheung et at. 2004), Personal protective equipment (Cheung et at. 2004), Fire precautions (Cheung et at. 2004), Electrical safety (Cheung et at. 2004), Housekeeping (Cheung et at. 2004), Hygiene and first aid facilities (Cheung et at. 2004)
Education training and Campaigns		Number of meetings (Cheung et at. 2004), Number of toolbox talks (Cheung et at. 2004), Number of introduction courses (Cheung et at. 2004)
Inspections and Audits		Number of in-house inspections (Cheung et at. 2004), Number of audits conducted (Cheung et at. 2004), Number of government inspections/visit (Cheung et at. 2004)
Complaints and prosecutions		Number of complaints received from staff (Cheung et at. 2004), Average time taken to close out the complaints (days) (Cheung et at. 2004), Number of prosecutions issued by government departments (Cheung et at. 2004), Average time taken to close out the case (days) (Cheung et at. 2004)
Total Recordable Incident Rate (TRIR)		Total number of recordable cases * 200000 / Total site work-hours (Construction Institute 2005)
Dart (LWCIR)		Total number of DART cases * 200000 / Total site work-hours (Construction Institute 2005)
Reportable Accidents (inc fatalities)	(inc fatalities)	Reportable accidents per 100000 hours worked (inc fatalities) (KPI Working Group 2000)
Reportable accidents (non-fatal)	(non-fatal)	Reportable accidents per 100000 hours worked (non-fatal) (KPI Working Group 2000)
Lost time accidents	time	Lost time accidents per 100000 hours worked (KPI Working Group 2000)
Fatalities		Fatalities per 100000 hours worked (KPI Working Group 2000)
Safety index		Number of accidents per 1000 man hours worked(KPI Working Group 2000), Time lost in accidents per 1000 man hours worked(KPI Working Group 2000), Change in

		the number of accident s or safety related problems (Cox et al. 2003)
Safety cost and incident score		Combination of accident occurrence index and total cost of safety (Sintawanarong 2000)
Compliance		Number of site inspections conducted (Cheung et al.2004), Number of complying (Cheung et al.2004), Air quality (Cheung et al.2004), Noise level, Waste (Cheung et al.2004), Discharge water (Cheung et al.2004)
Material control and waste management		Right amount of row materials are ordered (Cheung et al.2004), Surplus materials returned to the store (Cheung et al.2004), Construction and demolition wastes reused and recycled (Cheung et al.2004), Construction and demolition wastes sorted (Cheung et al.2004)
Meetings		Number of meetings held (Cheung et al.2004)
Education training and campaigns		Number of workshops, Number of people participating (Cheung et al.2004)
Inspection and audits		Number of public complaints received (Cheung et al.2004), Average time taken to close out the complaints (days) (Cheung et al.2004), Number of prosecutions issued by the EPD (Cheung et al.2004), Average time taken to close out the case (days) (Cheung et al.2004)

Table A.5. Performance Metric: Communication of Information (Cheung et al. 2004, Dawood et al. 2006), Changes (Construction Institute 2005, KPI Working Group 2000, Mahmoud and Scott 2002)

Perf.	Metric
Category	Performance Metrics
Communication management	Number of meetings held in the month (Cheung et al.2004), Number of change orders approved by the client (Cheung et al.2004), Number of day work orders issued (Cheung et al.2004), Number of RFI raised (Cheung et al.2004), Number of letters/ Engineer Reply Forms (ERF)issued to contractor/consultant (Cheung et al.2004),
	Number of letters/ Contractor Submissions Forms (CSF) received from contractor/ consultant (Cheung et al.2004)
Communication index	Number of meetings per week (Dawood et al.2006), Time spent of meetings per week (Dawood et al.2006), Number of request for information responded (Dawood et al.2006)
Change Factor	Cost
	Total cost of changes / actual total project cost (Construction Institute 2005)
Change orders 1	Number of individual change orders approved originating from client (KPI Working Group 2000)
Change orders 2	Number of individual change orders approved originating from project manager (KPI Working Group 2000)
Change order 3	Number of change orders (Mahmoud and Scott 2002)

Table A.6. Performance Metric Category: Productivity (Cheng et al. 2005, Korde et al. 2005, Construction Consultants Working Group 2007, Dawood et al. 2006, Takim and Akintoye 2002), Work Progress (Abudayyeh and Rasdorf 1993), Lean Thinking (Ballard and Howell 1994, Ballard 1994, Ballard 1997)

Perf.	Metric
Category	Performance Metrics
Productivity index 1	(value of sales – (value of goods purchased + value of goods subcontracted) / average number of UK employees, partners and directors (Construction Consultants Working Group 2007)
Productivity index 2	Tones of concrete used per day/m ³ (Dawood et al.2006), Pieces of steel used/day or week (Dawood et al.2006), Number of piles driven/day (Dawood et al.2006), Number of pile caps fixed/day (Dawood et al.2006)
Units/MH	The number of completed units put in place per individual man-hour of work (Cox et al.2003)
\$/unit	Dollar value associated with putting one complete unit in place including the materials, equipment and labor costs, and waste (Cox et al.2003)
Team performance index	Personnel turnover & productivity (Dawood et al.2006), Timeliness of information from team (Dawood et al.2006)
Weekly Plan Execution	Percent planned activities completed (Ballard and Howell 1994, Ballard 1994), Planned productivity (Ballard and Howell 1994)
Look-ahead Planning	Measuring the extent to which weekly work plan assignments previously appeared on look-ahead schedules-assignments anticipated (Ballard 1997), Measuring the extent to which assignment that appeared on look-ahead schedules appeared on weekly work plans when scheduled-assignments are made (Ballard 1997), Tracking the change over time of the scheduled dates for specific assignments using time/track charts (Ballard 1997)
SPI	schedule performance index : BCWP/BCWS (Abudayyeh and Rasdorf 1993)
CPI	cost performance index : BCWP/ACWP (Abudayyeh and Rasdorf 1993)

Table A.7. Performance Metric: Business Performance (KPI Working Group 2000)

Perf. Metric Category	Performance Metrics
Profitability	Company profit expressed as a percentage of turnover (KPI Working Group 2000)
Productivity	Company value added per employee expressed in pounds (KPI Working Group 2000)
Return on Capital employed	Company profit expressed as a percentage of capital employed (KPI Working Group 2000)
Return on value added	Company profit expressed as a percentage of value added (KPI Working Group 2000)
Interest cover	Company profit before tax and interest expressed as a percentage of interest payable (KPI Working Group 2000)
Return on investment	Compare of estimated return on investment for the client planned and actual as a percentage at the end of the period for rectifying defects (KPI Working Group 2000)
Profit predictability	Change in final gross margin btw actual and the planned (KPI Working Group 2000)
Ratio of value added	Company value added expressed as a percentage of turnover (KPI Working Group 2000)
Repeat business	The value of repeat business expressed as a percentage of turnover (KPI Working Group 2000)
Outstanding money	Amount of final account at available for use – after rectifying defects/at available for use (KPI Working Group 2000)
Time taken to reach final account	Time taken to reach final account from available for use expressed in weeks (KPI Working Group 2000)

Table A.8. Qualitative Performance Metric based on Subjective Measures: Project Members – Subs Performance (Ko et al.2007), Project Managers (Dainty et al.2003), People/Partners (Cheung et al. 2004, Takim and Akintoye 2002) and Client Satisfaction / Value (business benefit) (Alarkon and Ashley 1996, Cheung et al.2004, Cheng et al. 2005, 12Manage Premium 2008, Construction Consultants Working Group 2007, KPI Working Group 2000, Dawood et al. 2006, Takim and Akintoye 2002), Training and Recruitment (Construction Consultants Working Group 2007, Cheng et al. 2005, Takim and Akintoye 2002)

Perf.	Metric
Category	Performance Metrics
Subs Performance Index (filled in likert scale)	Construction technique (Ko et al.2007), Duration control ability (Ko et al.2007), Corporative manner (Ko et al.2007), Material wastage (Ko et al.2007), Service after work completion (Ko et al.2007), collaboration with other sub-contractors (Ko et al.2007), safety and protection, tool usage habits (Ko et al.2007), Working space clearance (Ko et al.2007), Management ability (Ko et al.2007), Subs's personality (Ko et al.2007), Economic condition (Ko et al.2007)
Team B	Team foundation and development (Dainty et al.2003), People management (Dainty et al.2003), Low staff turnover (Dainty et al.2003), Supportiveness (Dainty et al.2003)
Leadership	Monitoring and evaluation (Dainty et al.2003), Forward thinking (Dainty et al.2003), Responsibility (Dainty et al.2003), Delegation (Dainty et al.2003), assertiveness (Dainty et al.2003), flexibility (Dainty et al.2003)
Decision-making	Clear decision making approach (Dainty et al.2003), Recognizing key issues (prioritize issues) (Dainty et al.2003), Safety management and awareness (Dainty et al.2003), Problem solving (Dainty et al.2003), Risk management (Dainty et al.2003), Achieving production targets (Dainty et al.2003), Co-ordination skills (Dainty et al.2003)
Mutuality & approachability	Mutuality and thrust (Dainty et al.2003), Approachability (Dainty et al.2003)
Honesty & integrity	Honesty (Dainty et al.2003), Integrity (Dainty et al.2003), Commitment (Dainty et al.2003), Admitting weaknesses (Dainty et al.2003)
Communication	Communication (Dainty et al.2003), Knowledge transfer (within team) (Dainty et al.2003)
Learning, understanding and application	Rapid understanding of issues (Dainty et al.2003), Learning from mistakes (Dainty et al.2003), Problem identification (Dainty et al.2003), Commercial awareness (Dainty et al.2003), Thoroughness (Dainty et al.2003), Technical expertise (Dainty et al.2003)
Self-efficacy	Task motivation (Dainty et al.2003), Dedication (Dainty et al.2003), Initiative (Dainty et al.2003), Enthusiasm (Dainty et al.2003), Self-discipline (Dainty et al.2003), Time management (Dainty et al.2003)
External relations	Managing client relations (Dainty et al.2003), Presentation skills (Dainty et al.2003), Business acumen (Dainty et al.2003), Understanding organizational objectives (Dainty et al.2003)

	et al.2003)
Perceptive Views	On time (Cheung et al.2004), On cost (Cheung et al.2004), On quality (Cheung et al.2004), On safety (Cheung et al.2004), On environment (Cheung et al.2004), On communication (Cheung et al.2004), On contract relations (Cheung et al.2004), On claims (Cheung et al.2004), On issue resolution (Cheung et al.2004)
Satisfaction of owners needs in global sense	Realization of quantity produced (Alarkon and Ashley 1996), Realization of operational and maintenance costs (Alarkon and Ashley, 1996), Realization of flexibility (Alarkon and Ashley 1996), Finished product (up to the month) (Cheung et al.2004) - 1 to 10 score KPI Working Group 2000)
Product	Overall performance (Cheung et al.2004), Ability to identify and resolve problems (Cheung et al.2004), Leadership skills (Cheung et al.2004), Speed and reliability of service (Cheung et al.2004), Attention to detail (Cheung et al.2004), Overall value for money (Cheung et al.2004)
Project manager	Overall performance (Cheung et al.2004), Speed and reliability of service (Cheung et al.2004), Ability to keep promises (Cheung et al.2004), Resolution of defects (Cheung et al.2004)
Contractor and supplier	Tangibles (appearance of physical facility, equipment, communication) (12Manage Premium 2008), Reliability (12Manage Premium 2008), Responsiveness (12Manage Premium 2008), Competence (12Manage Premium 2008), Courtesy (12Manage Premium 2008), Credibility (12Manage Premium 2008), Feel secure (12Manage Premium 2008), Access (12Manage Premium 2008), Communication (12Manage Premium 2008), Understanding the customer (12Manage Premium 2008), Client score out of 10 (Construction Consultants Working Group 2007, KPI Group 2000)
(Quality of Service	1 to 10 scale on: Training, Profitability (Cheng et at 2005), Use of various management systems (Cheng et at 2005), Productivity (Cheng et at 2005), Repeat business (Cheng et at 2005), Success of fee proposal (Cheng et at 2005), Overall performance (Construction Consultants Working Group 2007), Health and safety awareness (Construction Consultants Working Group 2007), Value for money (Construction Consultants Working Group 2007, Timely delivery (Construction Consultants Working Group 2007, client specified criteria (KPI Working Group 2000)
Client satisfaction level	Number of client change orders (Dawood et al. 2004), Number of client queries (Dawood et al. 2004), Satisfaction questionnaire (Dawood et al. 2004), Number of claims (time/cost) (Dawood et al. 2004)
Satisfaction index	Total number of training days provided/average number of FTE(full time employment)
Recruitment index	UK employees, partners and directors (Construction Consultants Working Group 2007)

7.2 Appendix B: Consolidated Interview Results

Appendix B presents the consolidated interview results in a table format. An initial description of the columns of the interviews spreadsheet is introduced., and it is followed by three table presenting 132 metric queries identified through interviews. Since the interviews spreadsheet table is composed of thirty six columns, they could not fit into the page width. Hence, the table is divided into four sections, each composed of eight columns on average.

The column labels for the collected data:

- Metric ID: ID number for assigned or each of the performance metrics
- Job: Responsibility of the interviewee
- Project Type: Type of the project that interviewee is working on, H stands for Highway and B stands for Building.
- Experience Level: the years of experience that the interviewee has
- Graph Number: Each graph requested per interviewee is given a number to store the sequence
- Graph Name: the type of the visual representation requested per performance metric
- Performance Metric Category: the category of the performance metric, Productivity is abbreviates as Prod., and Communication as Comm.
- Metric vs. Source of Info: M stands for metric, and S stands for Source
- CoI: The concept of interest of the metric query, Activity abbreviated as acvt.
- Short Description: Short description of the metric query
- Intent: The intent of the interviewee
- Decision Purpose: the use of purpose of the decision to be taken, abbreviations: descriptive: decrp., predictive: predic., explanative: expln., prescriptive: presc., evaluative: eval.
- LoD project: a composing level of detail of metric query

- LoD Component: a composing level of detail of metric query
- LoD Actor: a composing level of detail of metric query
- LoD Resource: a composing level of detail of metric query
- LoD Space: a level of detail of metric query
- LoD Actor: a composing level of detail of metric query
- LoD Activ.: a composing level of detail of metric query
- LoD Documentation: a composing level of detail of metric query
- LoD Overhead: a composing level of detail of metric query
- LoD Time: a composing level of detail of metric query
- LoD Issues: a composing level of detail of metric query
- LoD Attributes: a composing level of detail of metric query
- LoD project: a filtering level of detail of metric query
- LoD Resource: a filtering level of detail of metric query
- LoD Space: a filtering level of detail of metric query
- LoD Activ.: a filtering level of detail of metric query
- LoD Attributes: a filtering level of detail of metric query
- LoD Time: a filtering level of detail of metric query
- Alert: binary to indicate if interviewee is interested in having notifications or not
- Alert Criteria: the criteria set to receive notifications, abbreviations are used
- Decision Detail: the decisions that interviewee would make after monitoring the performance metric
- Interaction: the type of interaction that the interviewee would do with the monitored metric
- Frequency: the frequency that the interviewee would monitor the performance metric. Always with alter is abbreviated as Al&Al, Further Graph as F.Gr, When information as W.Info.

Table B.1: First nine columns of the interviews spreadsheet for all metric queries

ID	job	Prj. type	Exp.	graph number	graph name	Metric Category	Metric vs Source of Info	CoI
1	Area Manager	H	18	1	3D Line Chart	Cost	m	Indirect/Overhead
1	Area Manager	H	18	1	3D Line Chart	Cost	m	Indirect/Overhead
2	Area Manager	H	18	2	Red Button	Cost	m	Activ.
3	Area Manager	H	18	2.1	Table	Cost	m	Activ.
3	Area Manager	H	18	2.1	Table	Cost	m	Activ.
4	Area Manager	H	18	2.2	Spreadsheet	Cost	m	Activ.
5	Area Manager	H	18	3	Gantt Chart	Time	m	Project
6	Area Manager	H	18	4	Spreadsheet	Prod.	m	Resource
7	Area Manager	H	18	5	Table	Comm.	m	Documentation
8	Area Manager	H	18	6	Calendar	Quick Access	source	ToDo's/ Milestones
9	Area Manager	H	18	7	Line Chart	Cost	m	Indirect/Overhead
10	Area Manager	H	18	8	Red Button	Comm.	m	Documentation
11	Area Manager	H	18	9	Table	Quality	m	Project
12	Area Manager	H	18	10	Red Button	Safety	m	Issues and Inspection
13	Area Manager	H	18	11	Red Button	Prod.	m	Resource
14	Area Manager	H	18	12	Table	Cost	m	Resource
15	Head of Estimation Department	B	30	1	Gantt Chart	Time	m	Project
16	Head of Estimation Department	B	30	2	Scatter Chart	Prod.	source	Resource
17	Head of Estimation Department	B	30	3	Gantt Chart	Time	source	Project
17	Head of Estimation Department	B	30	3	Gantt Chart	Time	source	Project
18	Head of Estimation Department	B	30	4	Table	Quality	m	Documentation
19	Head of Estimation Department	B	30	5	Bar Chart	Prod.	m	Resource
20	Head of Estimation Department	B	30	6	Flag	Prod.	m	Resource
21	Head of Estimation Department	B	30	7	Flag	Cost	m	Project
22	Head of Estimation Department	B	30	8	Camera	Quick Access	source	Activ.
23	Project engineer	H	10	1	Number	Cost	m	Project

24	Project engineer	H	10	2	Scatter Chart	Cost	m	Activ.
25	Project engineer	H	10	3	Gantt Chart	Time	m	Activ.
26	Project engineer	H	10	4	Odometer	Cost	m	Project
27	Project engineer	H	10	4.1	Spreadsheet	Cost	m	Activ.
28	Project engineer	H	10	5	Spreadsheet	Comm.	m	Documentation
29	Project engineer	H	10	6	Button/ Link	Quick Access	source	Issues and Inspection
30	Project engineer	H	10	7	Camera	Quick Access	source	Components
31	Project engineer	H	10	8	Flag	Safety	m	Issues and Inspection
32	Superintendent	H	30	1	Scatter Chart	Prod.	m	Resource
33	Superintendent	H	30	2	Line Chart	Prod.	m	Activ.
34	Superintendent	H	30	3	Gantt Chart	Time	m	Activ.
35	Superintendent	H	30	4	Button/ Link	Quick Access	source	Issues and Inspection
36	Superintendent	H	30	5	Button/ Link	Quick Access	source	Documentation
37	Superintendent	H	30	6	2D Drawing	Prod.	source	Space
38	Superintendent	H	30	7	Calculator	Quick Access	source	ToDos/ Milestones
39	Superintendent	H	30	8	Spreadsheet	Prod.	source	Resource
40	District Manager	Heavy Civil	30	1	Clustered Bar Chart	Time	m	Project
41	District Manager	Heavy Civil	30	1.1	Pie Chart	Time	m	Activ.
42	District Manager	Heavy Civil	30	2	Clustered Bar Chart	Cost	m	Project
43	District Manager	Heavy Civil	30	2.1	Pie Chart	Cost	m	Activ.
44	District Manager	Heavy Civil	30	3	Bell Curve	Comm.	m	Documentation
45	District Manager	Heavy Civil	30	3.1	Table	Comm.	m	Documentation
46	District Manager	Heavy Civil	30	4	Bell Curve	Comm.	m	Documentation
47	District Manager	Heavy Civil	30	4.1	Table	Comm.	m	Documentation
48	District Manager	Heavy Civil	30	5	3D Column Chart	Safety	m	Issues and Inspection
49	Superintendent	B	1	1	Table	Quick Access	source	Documentation
50	Superintendent	B	1	2	Gantt Chart	Time	m	Space
51	Superintendent	B	1	2.1	4D BIM	Prod.	source	Space
52	Superintendent	B	1	3	Button/ Link	Quick Access	source	Issues and Inspection
53	Superintendent	B	1	4	Camera	Quick	source	Components

						Access		
54	Superintendent	B	1	5	Spreadsheet	Prod.	source	Resource
55	Superintendent	B	1	6	Table	Comm.	m	Documentation
56	Superintendent	B	1	6.1	Table	Comm.	m	Documentation
57	Superintendent	B	1	7	Table	Comm.	m	Documentation
58	Superintendent	B	1	7.1	Table	Comm.	m	Documentation
59	Superintendent	B	1	8	Button/ Link	Quick Access	source	Documentation
60	Superintendent	B	1	9	Line Chart	Quality	m	Issues and Inspection
61	Superintendent	B	1	9.1	Table	Quality	m	Issues and Inspection
62	Superintendent	B	1	10	Spreadsheet	Safety	m	Issues and Inspection
63	Superintendent	B	1	11	Button/ Link	Quick Access	source	Documentation
64	Superintendent	B	1	12	Calendar	Quick Access	source	ToDos/ Milestones
65	Project engineer	B	8	1	Gantt Chart	Time	m	Resource
66	Project engineer	B	8	1.1	BIM	Prod.	source	Space
67	Project engineer	B	8	2	Spreadsheet	Comm.	m	Documentation
68	Project engineer	B	8	3	Table	Comm.	m	Documentation
69	Project engineer	B	8	4	Bar Chart	Comm.	m	Documentation
70	Project engineer	B	8	5	Flag	Comm.	source	Documentation
71	Project engineer	B	8	6	Calendar	Prod.	m	Resource
72	Project engineer	B	8	7	Calendar	Quick Access	source	ToDos/ Milestones
73	Project engineer	B	8	8	Flag	Safety	m	Issues and Inspection
74	Project manager	B	25	1	Spreadsheet	Cost	m	Resource
75	Project manager	B	25	1.1	Spreadsheet	Cost	m	Activ.
76	Project manager	B	25	2	Number	Cost	m	Indirect/Overhead
77	Project manager	B	25	3	Calendar	Quick Access	source	ToDos/ Milestones
78	Project manager	B	25	4	Gantt Chart	Quick Access	source	Activ.
79	Project manager	B	25	5	Table	Comm.	m	Documentation
80	Project manager	B	25	6	Table	Comm.	m	Documentation
81	Project manager	B	25	7	Table	Comm.	m	Documentation
82	Project engineer	B	2	1	Gantt Chart	Time	m	Activ.

83	Project engineer	B	2	1.1	4D BIM	Time	m	Components
84	Project engineer	B	2	2	Calendar	Quick Access	source	ToDos/ Milestones
85	Project engineer	B	2	3	Table	Comm.	m	Documentation
86	Project engineer	B	2	4	Table	Comm.	m	Documentation
87	Project engineer	B	2	5	Table	Comm.	m	Documentation
88	Project engineer	B	2	6	Line Chart	Cost	m	Resource
88	Project engineer	B	2	6	Line Chart	Cost	m	Resource
89	Project engineer	B	2	7	Button/ Link	Safety	m	Issues and Inspection
90	Project engineer	B	2	8	Spreadsheet	Comm.	source	Documentation
91	Project engineer	B	2	8.1	2D Drawing	Comm.	source	Documentation
92	Project engineer	B	2	9	Table	Comm.	source	Documentation
93	Project manager	B	9	1	Gantt Chart	Time	m	Activ.
94	Project manager	B	9	2	Spreadsheet	Comm.	m	Documentation
95	Project manager	B	9	3	Table	Cost	m	Activ.
96	Project manager	B	9	4	Table	Quick Access	source	Documentation
97	Project manager	B	9	5	Table	Comm.	source	Documentation
98	Project manager	B	9	6	Camera	Quick Access	source	Components
99	Project manager	B	9	7	Button/ Link	Quick Access	source	Issues and Inspection
100	Owner Representative	B	20	1	Scatter Chart	Comm.	m	Documentation
101	Owner Representative	B	20	1.1	Table	Comm.	m	Documentation
102	Owner Representative	B	20	2	Pie Chart	Comm.	m	Documentation
103	Owner Representative	B	20	3	Gantt Chart	Time	m	Activ.
104	Owner Representative	B	20	4	Pie Chart	Comm.	m	Documentation
105	Owner Representative	B	20	5	Button/ Link	Quick Access	source	Components
106	Owner Representative	B	20	6	Pie Chart	Comm.	m	Documentation
107	Owner Representative	B	20	7	Button/ Link	Comm.	source	Documentation
108	Owner Representative	B	20	8	Button/ Link	Comm.	source	Documentation

109	Owner Representative	B	20	9	Button/ Link	Cost	source	Indirect/Overhead
110	Owner Representative	B	20	10	Scatter Chart	Comm.	m	Documentation
111	Owner Representative	B	20	11	Button/ Link	Quality	m	Issues and Inspection
112	Project engineer	H	4	1	Gantt Chart	Time	source	Activ.
113	Project engineer	H	4	2	Spreadsheet	Comm.	m	Documentation
114	Project engineer	H	4	2.1	Table	Comm.	m	Documentation
115	Project engineer	H	4	3	Calendar	Prod.	source	Resource
116	Project engineer	H	4	4	Spreadsheet	Cost	m	Resource
117	Project engineer	H	4	5	Combo Box	Quick Access	source	Documentation
118	Project engineer	H	4	6	Button/ Link	Quick Access	source	Documentation
119	Project engineer	H	4	7	Button/ Link	Quick Access	source	ToDo's/ Milestones
120	Project engineer	H	4	8	Button/ Link	Quick Access	source	Documentation
121	Project engineer	H	8	1	Spreadsheet	Prod.	m	Resource
122	Project engineer	H	8	1.1	Scatter Chart	Prod.	m	Resource
122	Project engineer	H	8	1.1	Scatter Chart	Prod.	m	Resource
123	Project engineer	H	8	2	Spreadsheet	Prod.	m	Activ.
123	Project engineer	H	8	2	Spreadsheet	Prod.	m	Activ.
124	Project engineer	H	8	3	Spreadsheet	Prod.	m	Resource
125	Project engineer	H	8	4	Gantt Chart	Quick Access	source	Activ.
126	Project manager	H	15	1	Gantt Chart	Time	m	Activ.
127	Project manager	H	15	2	Gantt Chart	Quick Access	source	Activ.
128	Project manager	H	15	3	Scatter Chart	Prod.	m	Activ.
129	Project manager	H	15	4	Table	Prod.	source	Resource
130	Project manager	H	15	5	Spreadsheet	Cost	m	Activ.
131	Project manager	H	15	6	Calendar	Quick Access	source	Resource
132	Project manager	H	15	7	Table	Comm.	m	Documentation

Table B.2: 10th to 17th columns of the interviews spreadsheet for all m categories

Metri c ID	Short Description	Intent	use	LoDc- Project	LoDc- Compon ent	LoDc- Actor	LoDc- Resource	LoDc- Space
1	Delta CashFlow	Trend	Descpr.	Multi- Project				
1	Delta CashFlow	Compare	Descpr.	Multi- Project				
2	Delta Cost of Project	Detail	Descpr.	Multi- Project				
3	Delta Cost of Project	Compare	Explan.	Project				
3	Delta Cost of Project	Overview and Detail	Explan.	Project				
4	Delta Cost of Activities	Overview	Explan.					
5	Delta Schedule	Compare	Descpr.	Multi- Project				
6	Over-time % factor	Overview	Descpr.	Multi- Project			Foremen	
7	Status of documents	Overview	Descpr. quick access	Multi- Project				
8	Agenda	Overview						
9	Projected Profit	Trend	Predic.	Multi- Project				
10	EEO %	Overview and Detail	Descpr.				Craft	
11	Cost of Rework	Overview	Descpr.	Multi- Project				
12	Safety related problems	Overview	Descpr.	Project				
13	Delta Material procuremen t and usage	Overview	Descpr.	Multi- Project			Material	
14	Cost of Wasted Materials	Overview and Detail	Descpr.	Multi- Project			Material	
15	% completeness and delta schedule	Overview	Descpr.	Multi- Project				
16	Projected Resource Utilization	Trend	Predic.				Estimator and equipment	
17	Total Project Duration	Overview	Predic.	Multi- Project				
17	Total Project Duration	Trend	Predic.	Multi- Project				
18	Attributes of bids	Overview	eval.	Multi- Project				

19	Projected Resource Utilization	Trend	Predic.	Multi-Project	Estimator	
20	Projected Resource Utilization	Compare	Predic.	Multi-Project	Estimator	
21	Delta Cost of Project	Detail	Descpr.	Multi-Project		
22	Site tracking	Overview	quick access			
23	Delta Cost of Project	Detail	Descpr.	Project		
24	Delta Cost of Activities	Compare	Descpr.			
25	Delta Schedule	Compare	Descpr.			
26	Projected Cost overrun	Compare	Predic.	Project		
27	Projected Cost overrun	Trend	Explan.			
28	Details of documents	Overview	Descpr.	Project		
29	Weather	Overview	quick access			
30	Site tracking	Overview	quick access			
31	Safety related problems	Overview	Descpr.			
32	Delta Units	Trend	Descpr.		Foremen	
33	Units	Trend	Descpr.	units		
34	Delta Schedule	Compare	Predic.			
35	Weather	Overview	quick access			
36	Contact Information	Detail	quick access		all memb.	
37	Space usage	overview	Explan.	utilities	Subs	Work Area
38	Widgets	Overview	quick access			
39	Projected Resource Utilization	Overview	Predic.		Equipment	Site
40	Total Project Duration	Compare	Explan.	project		
41	Causes of schedule overrun	Compose	Explan.			
42	Delta Cost of Project	Compare	Descpr.	project		
43	Causes of	Compose	Explan.			

	cost overrun					
	Turnaround					
44	time of documents	Correlate	eval.	Multi-Project		
	Details of documents	Overview	Explan.	Multi-Project		
45						
	Turnaround					
46	time of documents	Correlate	eval.	Multi-Project		
	Details of documents	Overview	Explan.	Multi-Project		
47						
	Safety related problems	cluster	Descpr.	project		
48						
49	Contact Information	Detail	quick access		Subs	
	Space conflict anchored in Gantt Chart	Overview and Detail	Predic.			Work area
50						
51	Space usage	Overview	Predic.		Subs	Work area
			quick access			
52	Weather	Overview				
	Site tracking	Overview	quick access			
53						
	Material procurement and usage	Overview and Detail	Descpr.		material	
54	Status of documents	Overview and Detail	Descpr.			
55	Details of documents	Overview	prescr.			
56	Status of documents	Overview and Detail	Descpr.			
57	Details of documents	Overview	prescr.			
58	Contract and Specifications	Overview	quick access			
59						
60	Quality Problems	Correlate	Explan.		Subs	
	Quality Problems	Overview	Explan.			
61	inspection	Correlate	Predic.			
62	Contact Information	Detail	quick access		all memb.	
63			quick access			
64	Agenda	Overview				
	Delta Schedule	Compare	Descpr.		Subs	
65	Space usage	Overview	Explan.			Elevatin
66	Details of documents	Overview and Detail	prescr.			
67						

68	Status of documents	Overview and detail	Descpr.		Subs
69	Status of documents	Cluster	prescr.	architect, owner	Subs
70	Action items of meeting minutes	Overview	prescr.		
71	Delta Delivery schedule	Overview and detail	Descpr. quick access		material
72	Agenda Safety related problems	Overview	prescr.		Subs
73	Delta Cost of Project	Compare	Descpr.		Subs
74	Delta Cost of Activities	Compare	Explan.		
75	Contingency	Overview	Descpr.		
76	Achievement of milestones	Overview	prescr.		
77	Schedule of Activities	Overview	Descpr.		
78	Details of documents	Overview	prescr.		
79	Details of documents	Overview	prescr.		
80	Details of documents	Overview	Descpr.		
81	Delta Schedule 4D Activ. Progress	Compare	Descpr.	components	
82	Achievement of milestones	Overview	Descpr.		
83	Details of documents	Overview	prescr.		
84	Details of documents	Overview and Detail	prescr.		
85	Details of documents	Overview	Descpr.		Subs
86	Anticipated delta contract value	Compare	Descpr.		Subs
87	Anticipated delta contract value	Trend	Descpr.		Subs

89	Safety related problems	Overview	Descpr.		
90	Details of documents	Overview	Descpr.	Subs	Floor, room
91	2D punch list	Overview	Descpr.	Subs	Floor
92	Activ. To-Do	Overview	prescr.		
93	Delta Schedule	Compare	Descpr.		
94	Status of documents	Overview	Descpr.	Subs	
95	Delta Cost of Project	Compare	Descpr.		
96	Contact Information	Detail	quick access	all memb.	
97	Activ. To-Do	Overview	prescr.		
98	Site tracking	Overview	quick access		
99	Weather	Overview	quick access		
100	Turnaround time of documents	trend	Descpr.		
101	Status of documents	Overview and Detail	Descpr.		
102	Turnaround time of documents	distribution	Descpr.		
103	Delta Schedule	Compare	Descpr.		
104	Comments on drawing and reports	distribution	Descpr.		
105	Site tracking	Overview	quick access		
106	Turnaround time of documents	distribution	Descpr.		
107	Action items of meeting minutes	Overview	Descpr.		
108	Comments on drawing and reports	Overview	Descpr.		
109	payments	Overview	Descpr.		
110	Status of documents	trend	Descpr.		
111	Addressing inspection requirements	Overview	Descpr.		

112	Lookahead schedule	Overview	Descpr.	crew
113	Status of documents	Overview	Descpr.	
114	Status of documents	Overview	Explan.	
115	Material procurement and usage	Overview and Detail	prescr.	material
116	Delta Cost of Resources	Compare	Descpr.	labor and equipment
117	Contact Information	Detail	quick access	
118	Emails	Detail	quick access	
119	Widgets	Detail	quick access	
120	Contract and Specifications	Overview	quick access	
121	Delta Resource Utilization	Overview	Explan.	Labor and Equipment
122	Delta Resource Utilization	Compare	Predic.	Labor and Equipment
122	Delta Resource Utilization	Trend	Predic.	Labor and Equipment
123	Delta Units	Compare	Descpr.	
123	Delta Units	Overview and Detail	Descpr.	
124	Delta Material procurement and usage	Overview and Detail	Descpr.	material
125	Schedule of Activities	Overview	prescr.	
126	Delta Schedule	Compare	prescr.	
127	Lookahead schedule	Overview	prescr.	Subs
128	Units/MH	trend	prescr.	
129	Real time Resource utilization	Overview	prescr.	Equipment
130	Delta Cost of Activities	Compare	prescr.	
131	Delivery schedule	Overview	prescr.	Material and Equipment
132	Status of documents	Overview	prescr.	

Table B.3: 18th to 25rd columns of the interviews spreadsheet for all m queries

ID	LoDc- Activ.	LoDc-Doc	LoDc- Overhead	LoDc- Time	LoDc- Issues and Insp.	LoDc- Attribute	LoDf- Project	LoDf- Resource
1			Cashflow	Monthly		Estimated vs. Actual	Project	
1			Cashflow	Monthly		Estimated vs. Actual	Project	
2	Activ.			Weekly		Estimated vs. Actual	Project	Labor, Equipment, Material
3	Activ.			Weekly		Estimated vs. Actual		Labor, Equipment, Material
3	Activ.			Weekly		Estimated vs. Actual		Labor, Equipment, Material
4	Activ.			Weekly		Estimated vs. Actual		Labor and Equipment
5	CP Activ.			Whole Project		Planned vs. Actual Schedule, % complete	Project	
6				Weekly		Overtime factor	Project	
7		CO		Whole Project		Payment Status / Approval Status / Cost / DaysOpen	Project	
8								
9			Projected Profit	Monthly			Project	
10		EEO Report		Monthly		% ratio		
11				Whole Project		Cost of rework	Project	
12				Monthly	Safety Issues	Safety related problems		
13				Whole Project		Delay	Project	
14				Whole Project		wasted % complete, duration, type	Project	
15						Number		
16				Monthly		Dolar value, End Date, %	Current and	
17								

			complete	<i>Target</i>
			Dolar value,	<i>Current</i>
17			End Date, %	<i>and</i>
			complete	<i>Target</i>
			Number, %	
18	Bid		complete and	<i>Project</i>
			Dolar value	
19		Whole		
		Project	MH required	
			Number /	
20			Available vs	
21			Required	
22			Overbudget	
23		Whole	Actual vs.	
		Project	budget	
24	Activ.		Estimated vs.	
	CP	Weekly	Actual	
25	Activ.	Weekly	Planned vs.	
			Actual	
26			cost overrun	
			threshold	
27	Activ.	Whole	projected	
		Project	cost overrun	
28	CO		pending	
29		Daily		
30				
			Safety	
31			Issues	open
32				units of work
33	Activ.			
	Type			
			Planned vs.	
34	Activ.	Daily	Actual	<i>Subs</i>
35		Daily		
36	contact #			
37				
38				
39		Daily		
			estimated	
40			and projected	
			finish date	
41	Activ.		schedule	
			overrun	
42			Actual vs.	
43	Activ.		budgeted	
			cost overrun	
44	Submittal		turnaround	
45	Submittal		time	<i>project</i>
			details	<i>project</i>
46	RFI		turnaround	
47	RFI		time	<i>project</i>
48		Safety	details	
			# and days of	

			Issues	accidents	
49	Contact #				
50	Activ.		daily		
51	Activ.				
52			daily		
53					
54					
				<i>due and lateness status</i>	
55	RFI				
56	RFI			details	<i>Subs</i>
				due date and lateness status	
57	Submittal			details	
58	Submittal				
59	Activ. Contract				
60			Biweekly	Quality Issues	
61				Quality Issues	problem
62				Inspection	dates
63	Contact #				
64					
65			Daily		
66	Activ.				
				detail and status	
67	RFI				
				detail and status	<i>self- performed and subs</i>
68	RCO			number, etail and status	
69	Submittal				
	Meeting minutes			Action items	
70				delivery schedule	
71	Activ.		Daily		
72			calendar		
				remedial action and status	<i>Subs</i>
73			Safety Issues		
74					
				Estimated vs. Actual	
75	Activ.		Monthly		
76		Contingency	Whole Project calendar		
77					
78	Activ.		Monthly	schedule status and due dates	
79	RFI			require modifications	
80	Submittal			detsild	
81	CO Schedule			Planned vs. Actual	
82	Activ. Versions		Monthly		

83	Activ.					
84				calendar		
85		RFI			who is the party to reply and date	
86		Submittal			require modifications due date and schedule effect	
87		CO				
88						
88						
89				safety issues	remedial action and status	Subs
90		Punchlist				
91		Punchlist			corrections	
92	Activ.	action items			action items	
93	Activ.	action items		Daily	Planned vs. Actual status and date	
94		Submittal			budgeted, actual, projected cost	
95	Activ. Group		Profit	Whole Project		
96		Contract				
97	Activ.	action items			action items	
98		action items				
99						
100		RFI and Submittal			status	
101		RFI and Submittal				
102		Submittal			rate of submission Planned vs. Actual	Subs
103	Activ.			Weekly		
104		Design and Drawings			number of comments	
105						
106		Punchlist		Weekly	completion %	
107		Meeting minutes			Action items	
108		Field Report			comments	
109			Payment	Monthly		
110		CO		Weekly	status rate	
111				Daily	Inspection results	

			and Testing	
112	Activ.	3 weeks		lookahead
	submittal, RFI, transmittal and letters			
113				details
	submittal, RFI, transmittal and letters			
114				details
115				to be produced
116				
117	Contact #			
118				
119				
	Contract and Drawings			
120				man hrs. equip hrs
121				
122	Activ.	daily		
122	Activ.	daily		
123	Activ.			cost variance
123	Activ.			cost variance
				units (cy or sqft etc)
124	Activ.			schedule
125	Activ. CP	Weekly		Planned vs. Actual
126	Activ.			
127	Activ.	3 weeks		lookahead
				quantity/man hr.
128	Activ.	daily		
129		Minutely		
130	Activ.			
131				
	submittal, RFI, transmittal and letters			
132				

Table B.4: Last columns of the interviews spreadsheet for all m queries

ID	LoDf-Space	LoDf-Activ.	LoDf-Attribute	LoDf-Time	Alert	alert criteria	Interact	Freq.
1			\$ value threshold		flag	treshold	Highlight	Al&Al
1			\$ value threshold		flag	treshold	Highlight	Al&Al
2		all performed	Cost / threshold over \$	latest week	on	> (-\$5000)	Alert and Link	Al&Al
3		all performed	Cost / threshold over \$	latest week	flag	> (-\$5000)	Highlight and Link	F.Gr
3		all performed	Cost / threshold over \$	latest week	flag	> (-\$5000)	Highlight and Link	F.Gr
4		all performed		latest week			Sort	F.Gr
5		Activ.	Act:Delay / threshold over x days	Monthly	flag	treshold	Highlight	Al&Al
6			factor / thresholdl over x%				Navigate and sort	Pr.Time
7			Cost / threshold over \$ and Delay / threshold over x days			treshold	Highlight	Al&Al
8								always
9								always
10			Availability / over x %	latest month	on	treshold	Alert	Wh.Info
11								always
12			Availability / one or more		on	if happens over a treshhold day	Alert	Wh.Info
13			Delay over x days		on		Alert	Wh.Info
14			Availability / one or more		flag	treshold	Highlight	Al&Al
15							Sort	always
16							Link	always
17							Extra Info	always
17							Extra Info	always
18			Type of Project					always
19				Monthly				always
20					on		Alert	Wh.Info
21			Cost / over x \$		on		Alert	Al&Al

22							always
23						Retreive value	always
24		latest week					always
25		latest week					always
26							always
27						Sort	F.Gr
28	effect cost and time					Navigate and sort	always
29		Weekly				Retreive value	always
30							always
31		latest week	on	problem		Alert	Wh.Info
32	daily x and hours y for the duration of the activ. daily hours for duration of the activ.						always
33							always
34			highlight	due		Filter and highlight	Al&Al
35		# of days				Retreive value	always
36						Search	Wh.Info
37		Current				Link	always
38							always
39	eqp and job						always
40		Monthly	flag	time growth over 20%		Highlight and Link	Al&Al
41	overrun causes					Target	F.Gr
42			flag	cost overrun > 10%		Highlight and Link	Al&Al
43	overrun causes					Target	F.Gr
44	over ? Days		flag	days> 15		Highlight and Link	Al&Al
45	over ? Days						F.Gr
46	over ? Days		flag	days> 7		Highlight	Al&Al
47	over ? Days						F.Gr
48	without accidents	Days				Extra Info	always
49							always
50		Upcoming 2 weeks	flag	conflict		Highlight, Plan, Link	Al&Al
51			highlight	zoom		highlight, zoom and	F.Gr

					plan	
					Retreive	
52			# of days		value	always
53						always
54				highlight	delay	Highlight
						Al&Al
55				flag	delay	Highlight
56						and Link
57				flag	delay	Al&Al
58						F.Gr
59						Al&Al
						F.Gr
						always
			Latest 2			
60			months		link	always
61						F.Gr
	require					
62	inspection	dates		flag	before	Highlight
63				activ.		and Link
64						Pr.Time
						Wh.Info
						Wh.Info
			current			
65			month		Sort and	
					Link	always
					highlight,	
66			Current	flag	zoom and	
				delay	plan	Al&Al
	written by					
67	her / subs			flag	Sort and	
68				overdue	highlight	Wh.Info
				overdue	Highlight	Wh.Info
69				flag	Highlight	
				due	and Link	Wh.Info
70				on	Alert and	
				due	Link	Wh.Info
			current			
71			month	flag	Highlight	
72			whole	problem	and Link	Al&Al
						Wh.Info
73				on	Alert	Wh.Info
			latest			
74			month		Link	always
			latest			
75			month			F.Gr
					Retreive	
76					value	always
			Upcoming			
77			month			Wh.Info
			current			
78			month			always
		critical and				
		still open -				
79		affecting				Wh.Info
		schedule				
80		outstanding			Extra Info	Wh.Info
		affecting				
81		schedule			Link	Wh.Info
82					Extra Info	always
		% complete			Animate	
83	Activ.	daily			and	always

						retrieval value	
84			meeting times, deliveries, milestones			Link	Wh.Info
85			outstanding			Extra Info	Wh.Info
86			outstanding	flag	open doc	Highlight	Wh.Info
87			affecting schedule			Extra Info	Wh.Info
88							always
88							always
89						Extra Info	Wh.Info
90						Link	Wh.Info
91						Navigate and sort	Wh.Info
92							Wh.Info
93				Upcoming month		Sort	always
94			open			Cluster and Sort	Wh.Info
95					if overbudget - any change	Highlight	always
96				flag			always
97					pop up in the day		always
98							always
99				# of days		Retrieve value	always
100			submitted, answered and outstanding RFI and submitted, to be submitted, outstanding submittals	latest month		Link	Wh.Info
101			open submitted and to be submitted		flag	outstanding (over 2 weeks)	Highlight Wh.Info
102					flag	outstanding (over 2 weeks)	Highlight always
103	Area	critical		latest week		Navigate and sort	always
104			rejected, accepted, commented comments			Link	Wh.Info
105							always
106							always

107	last three					always
108		Latest month				Pr.Time
109		Latest month				Pr.Time
110		Latest month				always
111		Current			Link	Wh.Info
112	Work Location				Cluster Sort and highlight	always
113		outstanding	flag	outstanding		Wh.Info
114		outstanding				Wh.Info
115		upcoming week	flag	delay	Highlight	always
116		current activities	latest week	flag	treshold	Highlight
117						always
118						Wh.Info
119						always
120					Search	always
121		activities performed the previous day	Latest day		Extra Info	always
122			latest week			always
122			latest week			always
123			latest week	highlight	difference over 10%	Sort and highlight
123			latest week	highlight	difference over 10%	Sort and highlight
124			latest week	highlight	schedule variance over 10%	Highlight
125			latest week			always
126			upcoming milestone			always
127					Cluster	always
128		heavy eathwork activities	whole			Pr.Time
129			Latest day			always
130			Latest week		Extra Info	always
131			Latest month			always
132		outstanding				always

7.3 Appendix C: Results of the PCA analysis of m queries

Appendix C presents the variables used on the PCA analysis and the coefficients identified as a result of PCA performed for the validation of RQ 1

Table C.1: Twenty-six original variables and their column rankings

Column #-Variable			
1 Interviewee ID	8 Frequency	15 LoD _c Documentation	22 LoD _f Space
2 Visual Representation	9 LoD _c project	16 LoD _c Overhead	23 LoD _f Activity
3 Performance Metric Category	10 LoD _c Component	17 LoD _c Time	24 LoD _f Attribute
4 CoI	11 LoD _c Actor	18 LoD _c Issues	25 LoD _f Time
5 Intent	12 LoD _c Resource	19 LoD _c Attribute	26 LoD _f Interaction
6 Alert	13 LoD _c Space	20 LoD _f Project	
7 Decision Usage	14 LoD _c Activity	21 LoD _f Resource	

Table C.2: First half of the Coefficient Matrix of PCA Analysis

Coefficient Matrix												
0.08	0.04	0.08	0.34	0.64	-0.49	-0.25	0.20	0.13	-0.22	-0.10	-0.14	0.01
-0.02	0.00	0.94	-0.12	0.00	-0.08	-0.10	-0.15	-0.14	0.13	0.12	0.07	0.06
0.01	0.00	0.06	0.06	0.01	0.09	0.30	0.47	-0.17	-0.24	0.09	0.19	0.66
-0.01	0.03	0.12	0.04	0.03	0.09	0.38	0.42	-0.44	-0.06	0.03	-0.43	-0.48
0.01	0.06	0.26	-0.05	-0.01	0.27	0.36	0.04	0.71	-0.36	-0.24	-0.08	-0.12
0.00	-0.01	-0.01	-0.02	-0.04	-0.02	0.02	-0.02	-0.07	0.05	-0.06	-0.23	0.07
0.00	0.00	0.00	0.04	0.10	-0.01	0.11	0.42	0.40	0.76	0.23	0.00	0.00
0.01	0.02	0.00	0.07	0.07	-0.07	0.13	0.15	-0.07	-0.11	0.10	0.81	-0.41
0.00	0.00	0.00	-0.04	-0.06	0.03	-0.02	-0.05	-0.02	0.04	0.01	0.01	0.12
0.00	0.01	0.01	-0.01	0.00	-0.01	0.00	0.00	-0.01	0.03	0.00	-0.01	-0.12
0.00	0.00	0.00	0.01	0.00	0.00	0.02	0.01	0.03	0.00	-0.01	0.02	-0.01
0.00	-0.01	-0.08	-0.06	0.25	-0.39	0.71	-0.47	-0.05	0.15	-0.01	0.01	0.10
0.00	0.00	0.02	0.02	0.01	-0.01	0.06	-0.07	-0.07	0.02	0.00	-0.05	-0.14
0.00	-0.01	-0.01	-0.02	0.01	-0.01	-0.10	-0.04	0.08	-0.01	0.00	-0.06	-0.15
0.07	0.15	0.07	0.88	-0.08	0.31	0.08	-0.26	-0.04	0.09	0.09	0.00	0.03
0.00	-0.01	0.00	-0.01	0.00	-0.03	-0.08	-0.04	0.06	-0.05	-0.01	0.06	-0.14
-0.01	-0.04	-0.07	-0.10	0.17	0.11	-0.01	-0.13	0.13	-0.28	0.90	-0.14	-0.04
0.00	-0.01	0.01	-0.02	0.00	0.02	0.04	0.07	-0.05	-0.08	0.01	-0.02	0.09
0.97	-0.21	0.01	-0.06	-0.06	0.01	0.01	0.00	-0.01	0.01	0.01	0.00	-0.02
0.00	0.00	-0.01	-0.01	-0.03	0.02	0.00	-0.02	0.03	0.01	0.00	-0.02	0.05
0.00	-0.02	-0.01	0.00	-0.02	-0.02	0.01	0.01	-0.02	0.02	0.02	-0.04	0.17
0.00	0.00	0.00	0.00	0.00	-0.01	0.00	-0.01	0.01	-0.02	0.03	-0.01	-0.01
0.00	0.00	-0.01	-0.01	-0.01	-0.01	-0.04	0.02	-0.01	0.00	0.02	-0.06	0.03
0.19	0.96	-0.04	-0.18	0.02	-0.01	-0.03	-0.01	-0.03	0.03	0.04	0.00	0.02
0.01	-0.05	-0.03	-0.19	0.68	0.63	-0.05	-0.12	-0.16	0.14	-0.17	0.08	0.04
0.00	0.01	-0.01	-0.02	-0.01	0.00	0.01	-0.04	-0.05	-0.03	-0.02	-0.05	-0.05

Table C.3: Second half of the Coefficient Matrix of PCA Analysis

-0.01	0.03	-0.04	-0.02	0.01	-0.09	0.04	-0.01	0.02	0.01	0.01	0.01	-0.01
0.01	0.01	0.00	0.00	0.01	0.03	-0.01	0.01	-0.01	0.01	-0.02	0.00	0.00
-0.04	0.07	0.24	0.02	0.01	0.12	0.12	-0.01	-0.11	-0.03	0.04	0.00	0.02
-0.03	-0.06	-0.12	0.07	-0.02	0.03	0.08	-0.07	0.05	0.02	-0.07	0.01	-0.01
0.05	0.00	-0.03	-0.03	0.00	-0.07	0.01	-0.02	0.01	-0.02	0.02	-0.03	-0.01
0.61	0.69	0.06	-0.12	-0.12	-0.09	-0.14	-0.13	0.00	0.04	0.07	0.00	0.01
0.00	0.03	0.03	0.04	0.05	-0.03	-0.01	0.04	0.01	-0.02	-0.02	-0.02	0.02
0.24	0.10	-0.06	-0.09	0.02	-0.10	0.00	-0.04	0.04	0.05	-0.03	-0.02	-0.01
-0.11	0.14	-0.20	0.01	0.26	-0.40	0.56	-0.36	0.43	-0.20	0.09	0.03	-0.03
0.05	-0.08	0.03	-0.04	0.18	0.12	0.09	0.02	-0.18	0.07	0.93	0.11	0.02
0.01	0.03	-0.03	-0.02	0.03	0.00	-0.03	0.02	-0.10	-0.21	-0.13	0.96	0.02
-0.02	-0.02	0.00	0.04	0.05	0.10	0.01	-0.03	0.01	0.01	-0.01	-0.01	-0.01
-0.09	-0.11	0.75	-0.11	-0.18	-0.57	-0.01	0.05	0.00	-0.03	0.05	0.03	0.02
0.22	-0.06	0.46	-0.22	0.21	0.57	0.39	0.01	0.30	0.06	-0.16	0.03	0.01
0.03	0.01	0.00	0.02	0.04	0.04	-0.01	0.00	0.01	0.00	0.00	0.00	0.00
0.03	0.17	0.24	0.91	0.01	0.10	0.09	-0.12	-0.10	-0.06	-0.01	0.00	0.02
0.02	0.05	-0.02	-0.02	-0.03	-0.03	-0.01	0.00	-0.01	-0.01	0.02	0.01	-0.03
0.06	-0.04	0.05	0.18	0.47	-0.11	-0.51	0.27	0.58	0.13	0.07	0.07	0.06
-0.01	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	0.00
-0.07	0.05	-0.04	0.04	0.11	-0.16	0.21	-0.03	-0.14	0.92	-0.12	0.17	-0.02
0.64	-0.54	-0.16	0.19	-0.26	-0.18	0.23	0.21	0.10	0.02	0.01	0.03	0.01
0.02	-0.06	-0.02	-0.04	0.03	-0.01	-0.02	-0.17	-0.05	0.01	-0.03	-0.02	0.98
0.27	-0.31	0.09	-0.06	0.53	-0.10	-0.21	-0.53	-0.38	-0.08	-0.15	-0.07	-0.16
0.01	0.00	0.02	0.01	-0.01	0.01	0.00	0.00	0.01	0.00	0.00	0.01	0.00
0.03	0.00	0.01	0.02	0.00	0.01	0.00	-0.01	-0.01	0.01	0.00	0.01	0.01
0.06	0.18	-0.05	-0.02	0.46	-0.18	0.28	0.64	-0.37	-0.19	-0.15	-0.13	0.08

7.4 Appendix D: Heuristics and Visual Representations

Appendix D is composed of three tables. Table D.1 presents 490 of the *heuristics* identified from literature with the column headings of data dimension, data variability, data structure, intent, visual representations and interactions. Table D.2 presents 163 visual representations grouped according to their higher level categories. Table D.3 presents the constraints/Compare *heuristics*.

Table D.1: List of 490 Heuristics

Heuristic ID	Dimension	Variable	Structure	Intent	Visual
1	N/A	N/A	temporal	overview	2D line graphs
2	2D	both	temporal	N/A	2D line graphs
3	3D	N/A	N/A	trend	3D area graph
4	3D	N/A	temporal	N/A	3D bar graph
5	N/A	N/A	temporal	overview	3D kiviart tube:
6	3D	N/A	temporal	N/A	3D line
7	3D	N/A	N/A	N/A	3D line
8	3D	both	N/A	N/A	3D line graph
9	3D	both	N/A	N/A	2D line graphs
10	3D	both	N/A	N/A	2D line graphs
11	3D	both	N/A	N/A	3D line graph
12	N/A	N/A	temporal	overview	3D line graph
13	3D	N/A	N/A	N/A	3D scatter graph
14	N/A	N/A	temporal	overview	3D time wheel -
15	N/A	N/A	hierarchical	O & D	stacked line graphs
16	N/A	N/A	relational	O & D	affinity diagram
17	N/A	quantitative	N/A	Compare	affinity diagram
18	N/A	N/A	N/A	overview	aligned bar graph
19	multiD	N/A	N/A	Compare	andrew curves
20	multiD	N/A	N/A	Compare	andrew curves
21	1D	N/A	N/A	N/A	andrew curves
22	3D	N/A	temporal	O & D	animation
23	N/A	N/A	hierarchical	N/A	animation
24	N/A	N/A	relational	overview	area graph
25	N/A	N/A	N/A	overview	area graph
26	N/A	quantitative	temporal	distributi	on
27	multiD	N/A	N/A	bar graph	
28	N/A	N/A	N/A	N/A	bar graph
29	N/A	N/A	hierarchical	Compare	bar graph
30	N/A	N/A	relational	overview	bar graph
31	1D	quantitative	N/A	overview	bar graph
32	2D	quantitative	N/A	N/A	bar graph
33	3D	quantitative	N/A	N/A	bar graph

34	N/A	N/A	N/A	Compare	bar graph
35	2D	N/A	temporal	N/A	bar graph
36	2D	N/A	N/A	N/A	bar graph
37	N/A	N/A	N/A	Compare	bar graph
38	N/A	N/A	hierarchical	overview	bcg matrix
39	N/A	N/A	relational	overview	bcg matrix
40	N/A	N/A	N/A	detail	box plot
41	N/A	N/A	N/A	N/A	box plot
42	1D	quantitative	N/A	N/A	box plot
43	2D	quantitative	N/A	N/A	box plot
44	3D	quantitative	N/A	N/A	box plot
45	N/A	quantitative	N/A	distributi on	box plot
46	multiD	quantitative	temporal	detail	box plot
47	3D	N/A	N/A	correlate	bubble graph
48	N/A	N/A	N/A	overview	chernoff faces
49	multiD	N/A	N/A	N/A	chernoff faces
50	N/A	N/A	N/A	N/A	chernoff faces
51	N/A	N/A	hierarchical	O & D	chernoff faces
52	N/A	N/A	relational	O & D	chernoff faces
53	1D	quantitative	N/A	N/A	chernoff faces
54	2D	quantitative	N/A	N/A	chernoff faces
55	3D	quantitative	N/A	N/A	chernoff faces
56	multiD	N/A	N/A	N/A	chernoff faces
57	multiD	N/A	N/A	trend	circle segments
58	N/A	N/A	N/A	correlate	circle segments
59	N/A	N/A	temporal	Compare	circular area graph
60	N/A	N/A	hierarchical	correlate	cognitive mapping
61	N/A	N/A	relational	correlate	cognitive mapping
62	1D	N/A	N/A	O & D	color map
63	2D	N/A	N/A	O & D	color map
64	1D	N/A	N/A	cluster	color map
65	2D	N/A	N/A	cluster	color map
66	3D	N/A	N/A	cluster	color map
67	N/A	N/A	N/A	Compare	column graph
68	N/A	N/A	N/A	trend	column graph
69	N/A	N/A	N/A	Compare	column graph
70	N/A	N/A	N/A	Compare	column graph
71	N/A	N/A	temporal	Compare	column graph
72	1D	N/A	N/A	trend	column histogram
73	N/A	N/A	N/A	Compare	column graph combined with curves
74	N/A	N/A	N/A	trend	column graph combined with curves
75	2D	N/A	N/A	correlate	combined pie charts
76	N/A	N/A	hierarchical	O & D	concept map
77	N/A	N/A	relational	O & D	concept map
78	N/A	N/A	hierarchical	O & D	cone tree
79	N/A	N/A	temporal	correlate	cpm critical path method
80	1D	N/A	N/A	O & D	curve density plot

81	1D	N/A	N/A	cluster	curve density plot
82	1D	N/A	N/A	trend	curve density plot
83	N/A	N/A	temporal	overview	cycle diagram
84	N/A	categorical	N/A	overview	daisy maps
85	N/A	N/A	relational	N/A	data model
86	multiD	quantitative	temporal	Compare	data vases
87	multiD	quantitative	hierarchical, relational, spatial, tabular, temporal, spatio-temporal	Compare	data vases
88	multiD	quantitative	temporal	trend	data vases
89	multiD	quantitative	hierarchical, relational, spatial, tabular, temporal, spatio-temporal	trend	data vases
90	multiD	quantitative	temporal	detail	data vases
91	multiD	quantitative	hierarchical, relational, spatial, tabular, temporal, spatio-temporal	detail	data vases decision discovery diagram
92	N/A	N/A	temporal	detail	diagram
93	N/A	N/A	temporal	O & D	decision tree
94	N/A	N/A	relational	N/A	decision tree
95	N/A	N/A	hierarchical	O & D	disk tree
96	N/A	N/A	N/A	Compare	divided bar graph
97	N/A	text	N/A	O & D	document lens
98	N/A	quantitative	spatial	N/A	dot plot entity relationship diagram
99	N/A	N/A	hierarchical	correlate	entity relationship diagram
100	N/A	N/A	relational	correlate	diagram
101	1D	quantitative	N/A	N/A	fan
102	2D	quantitative	N/A	N/A	fan
103	3D	quantitative	N/A	N/A	fan
104	N/A	N/A	temporal	overview	feedback diagram
105	N/A	N/A	temporal	O & D	flowchart
106	N/A	N/A	relational	N/A	flowchart
107	N/A	N/A	N/A	overview	flowchart
108	N/A	quantitative	spatial	N/A	flow maps
109	N/A	quantitative	N/A	N/A	gantt chart
110	N/A	categorical	temporal	N/A	gantt chart
111	N/A	N/A	temporal	overview	gantt chart
112	N/A	N/A	temporal	N/A	gantt chart
113	N/A	N/A	spatial	overview	geo-spatial map
114	multiD	N/A	hierarchical	correlate	geographic map
115	2D	N/A	N/A	O & D	geographic map
116	2D	N/A	N/A	Compare	geographic map
117	multiD	N/A	N/A	N/A	geographic map
118	2D	N/A	N/A	N/A	geographic map
119	N/A	quantitative	temporal	N/A	glyphs
120	N/A	quantitative	N/A	N/A	graphical table grouped column graph
121	N/A	N/A	N/A	Compare	graph
122	N/A	N/A	N/A	overview	heat map
123	N/A	N/A	N/A	distributi on	histogram
124	1D	N/A	N/A	trend	histogram
125	N/A	N/A	hierarchical	overview	histogram

126	N/A	N/A	relational	overview	histogram
127	1D	quantitative	N/A	N/A	histogram
128	2D	quantitative	N/A	N/A	histogram
129	3D	quantitative	N/A	N/A	histogram
130	2D	N/A	N/A	N/A	histogram
131	N/A	quantitative	N/A	distributi on	histogram
132	N/A	N/A	hierarchical	O & D	hyperbolic tree
133	N/A	N/A	relational	O & D	hyperbolic tree
134	N/A	N/A	hierarchical	O & D	hyperbolic view
135	N/A	N/A	hierarchical	trend	information cube
136	2D	N/A	N/A	trend	isogram graph
137	N/A	quantitative	spatial	N/A	isopleths
138	multiD	quantitative	N/A	overview	kite diagram
139	multiD	quantitative	N/A	trend	kite diagram
140	N/A	N/A	hierarchical	overview	knowledge map
141	N/A	N/A	relational	overview	knowledge map
142	multiD	N/A	N/A	N/A	landscape
143	N/A	N/A	temporal	N/A	line graph
144	1D	N/A	N/A	O & D	line graph
145	1D	N/A	N/A	Compare	line graph
146	N/A	N/A	temporal	Compare	line graph
147	N/A	N/A	temporal	Compare	line graph
148	3D	N/A	N/A	N/A	line graph
149	3D	N/A	N/A	N/A	line graph
150	N/A	N/A	hierarchical	overview	line graph
151	N/A	N/A	relational	overview	line graph
152	1D	quantitative	N/A	N/A	line graph
153	2D	quantitative	N/A	N/A	line graph
154	3D	quantitative	N/A	N/A	line graph
155	2D	N/A	temporal	N/A	line graph
156	2D	N/A	N/A	N/A	line graph
157	N/A	N/A	N/A	trend	line graph
158	1D	N/A	N/A	trend	line histogram
159	N/A	N/A	hierarchical	O & D	magic eye view
160	N/A	N/A	spatial	overview	map
161	N/A	N/A	hierarchical	O & D	mindmap
162	N/A	N/A	relational	O & D	mindmap
163	multiD	N/A	N/A	Compare	multiple box plot
164	N/A	quantitative	N/A	correlate	multiple scatter graphs
165	multiD	N/A	N/A	N/A	needle icons
166	N/A	N/A	N/A	trend	netmap
167	N/A	text	N/A	correlate	netmap
168	N/A	categorical	N/A	correlate	network diagram
169	multiD	text	relational	correlate	network diagram
170	N/A	quantitative	N/A	N/A	analytical table
171	N/A	N/A	hierarchical	overview	organizational chart
172	N/A	N/A	relational	overview	organizational chart
173	N/A	N/A	relational	N/A	organizational chart
174	N/A	N/A	N/A	overview	organizational

					chart
175	N/A	quantitative	N/A	correlate	parallel coordinates
176	N/A	N/A	hierarchical	N/A	parallel coordinates
177	multiD	N/A	N/A	cluster	parallel coordinates
178	multiD	N/A	N/A	N/A	parallel coordinates
179	multiD	N/A	N/A	N/A	parallel coordinates
180	multiD	N/A	N/A	N/A	parallel coordinates
181	N/A	N/A	hierarchical	O & D	parallel coordinates
182	N/A	N/A	relational	O & D	parallel coordinates
183	multiD	N/A	N/A	N/A	parallel coordinates
184	N/A	N/A	hierarchical	overview	pareto chart
185	N/A	N/A	relational	overview	pareto chart
186	N/A	N/A	temporal	overview	pert chart
187	N/A	N/A	relational	N/A	pert chart
188	N/A	N/A	N/A	distributi	on
189	1D	N/A	N/A	Compare	pie chart
190	N/A	N/A	N/A	compose	pie chart
191	N/A	N/A	hierarchical	overview	pie chart
192	N/A	N/A	relational	overview	pie chart
193	1D	quantitative	N/A	N/A	pie chart
194	2D	quantitative	N/A	N/A	pie chart
195	3D	quantitative	N/A	N/A	pie chart
196	N/A	N/A	N/A	Compare	pie chart
197	1D	N/A	N/A	distributi	on
198	2D	both	temporal	N/A	pie chart
199	multiD	N/A	N/A	N/A	point plot
200	1D	quantitative	N/A	N/A	point plot
201	2D	quantitative	N/A	N/A	polar coordinate
202	3D	quantitative	N/A	N/A	polar coordinate
203	N/A	quantitative	N/A	distributi	on
204	2D	N/A	N/A	Compare	quantile plot
205	3D	N/A	N/A	Compare	quantile-quantile plot
206	multiD	N/A	N/A	Compare	quantile-quantile plot
207	N/A	N/A	hierarchical	overview	plot
208	N/A	N/A	relational	overview	radar graph
209	multiD	N/A	N/A	N/A	radar graph
210	N/A	N/A	hierarchical	overview	relationalTable
211	N/A	N/A	relational	overview	sankey diagram
212	N/A	N/A	N/A	trend	sankey diagram
					scatterplot

213	multiD	N/A	N/A	trend	scatterplot
214	multiD	categorical or both	N/A	correlate	scatterplot
215	multiD	categorical or both	N/A	correlate	scatterplot
216	N/A	quantitative	N/A	correlate	scatterplot
217	N/A	categorical or both	N/A	Compare	scatterplot
218	multiD	quantitative	N/A	correlate	scatterplot
219	N/A	N/A	hierarchical	N/A	scatterplot
220	2D	N/A	N/A	O & D	scatterplot
221	3D	N/A	N/A	O & D	scatterplot
222	2D	N/A	N/A	Compare	scatterplot
223	3D	N/A	N/A	Compare	scatterplot
224	multiD	N/A	N/A	N/A	scatterplot
225	2D	N/A	N/A	correlate	scatterplot
226	2D	N/A	N/A	trend	scatterplot
227	2D	N/A	N/A	N/A	scatterplot
228	2D	N/A	N/A	N/A	scatterplot
229	N/A	N/A	hierarchical	overview	scatterplot
230	N/A	N/A	relational	overview	scatterplot
231	1D	quantitative	N/A	N/A	scatterplot
232	2D	quantitative	N/A	N/A	scatterplot
233	3D	quantitative	N/A	N/A	scatterplot
234	2D	N/A	N/A	N/A	scatterplot
235	N/A	quantitative	N/A	distributi on	scatterplot
236	2D	quantitative	N/A	correlate	scatterplot
237	2D	both	temporal	N/A	scatterplot
238	multiD	categorical or quantitative	spatial	N/A	scatterplot map
239	multiD	N/A	N/A	O & D	scatterplot matrix
240	multiD	N/A	N/A	correlate	scatterplot matrix
241	multiD	N/A	N/A	N/A	scatterplot matrix
242	multiD	N/A	N/A	N/A	scatterplot matrix
243	3D	quantitative	N/A	correlate	scatterplot matrix
244	multiD	quantitative	N/A	correlate	scatterplot matrix
245	N/A	N/A	N/A	correlate	schematic diagram
246	multiD	N/A	N/A	trend	self-organizing maps
247	multiD	N/A	N/A	cluster	self-organizing maps
248	multiD	both	temporal	cluster	self-organizing maps
249	N/A	N/A	hierarchical	overview	semantic network
250	N/A	N/A	relational	overview	semantic network
251	N/A	categorical	temporal	N/A	sequence chart
252	multiD	quantitative	spatio-temporal	N/A	space-time cube
253	N/A	N/A	spatial	N/A	geo-spatial map
254	N/A	N/A	hierarchical	overview	spectrogram
255	N/A	N/A	relational	overview	spectrogram
256	N/A	categorical	temporal	correlate	spiral graph
257	N/A	quantitative	temporal	correlate	spiral graph
258	N/A	categorical	temporal	correlate	spiral graph

259	N/A	categorical	temporal	trend	spiral graph
260	N/A	quantitative	temporal	trend	spiral graph
261	N/A	categorical	temporal	trend	spiral graph
262	N/A	N/A	temporal	compose	stacked 100% area graph
263	N/A	N/A	temporal	compose	stacked 100% column graph
264	N/A	N/A	N/A	compose	stacked 100% column graph
265	N/A	N/A	temporal	compose	stacked area graph
266	N/A	N/A	temporal	Compare	stacked area graph
267	multiD	categorical	N/A	N/A	stacked bar graph
268	1D	quantitative	N/A	N/A	stacked bar graph
269	2D	quantitative	N/A	N/A	stacked bar graph
270	3D	quantitative	N/A	N/A	stacked bar graph
271	N/A	N/A	temporal	compose	stacked column graph
272	N/A	N/A	temporal	Compare	stacked column graph
273	multiD	N/A	N/A	O & D	star glyphs
274	multiD	N/A	N/A	Compare	star glyphs
275	1D	quantitative	N/A	N/A	star glyphs
276	2D	quantitative	N/A	N/A	star glyphs
277	3D	quantitative	N/A	N/A	star glyphs
278	multiD	N/A	N/A	N/A	star glyphs
279	multiD	N/A	N/A	correlate	star glyphs
280	multiD	N/A	N/A	N/A	star icons
281	N/A	N/A	N/A	Compare	step chart
282	N/A	N/A	N/A	correlate	step chart
283	N/A	N/A	temporal	overview	system dyn/ loop diagram
284	N/A	N/A	hierarchical	overview	supply-demand chain
285	N/A	N/A	relational	overview	supply-demand chain
286	N/A	N/A	N/A	trend	surface graph
287	2D	N/A	N/A	N/A	Survey plot
288	N/A	N/A	hierarchical	overview	table
289	N/A	N/A	relational	overview	table
290	N/A	N/A	N/A	Compare	table
291	N/A	N/A	N/A	overview	table
292	multiD	N/A	N/A	O & D	table lens
293	multiD	N/A	N/A	N/A	table lens
294	N/A	N/A	N/A	Compare	table with colors
295	N/A	text	N/A	overview	tag clouds
296	multiD	N/A	temporal	N/A	temporal parallel coordinates
297	N/A	N/A	N/A	detail	text table
298	1D	N/A	temporal	N/A	themeRiver
299	N/A	text	N/A	N/A	themeRiver
300	multiD	quantitative	temporal	N/A	themeRiver
301	N/A	text	N/A	Compare	tileBars
302	N/A	text	N/A	trend	tileBars
303	N/A	N/A	N/A	overview	time and activ.

					charts
304	N/A	N/A	N/A	overview	timeline
305	N/A	quantitative	temporal	N/A	timeline
306	N/A	N/A	temporal	overview	timeline
307	multiD	N/A	temporal	correlate	timeWheel
308	multiD	N/A	temporal	trend	timeWheel
309	N/A	N/A	spatial	overview	topological map
310	N/A	N/A	hierarchical	overview	tree
311	N/A	N/A	relational	overview	tree
312	N/A	N/A	hierarchical	trend	treemap
313	multiD	N/A	hierarchical	N/A	treemap
314	N/A	N/A	hierarchical	O & D	treemap
315	N/A	N/A	relational	O & D	treemap
316	multiD	text	hierarchical	N/A	treemap
317	3D	N/A	N/A	N/A	triple scatterplot
318	N/A	N/A	hierarchical	overview	tukey box plot
319	N/A	N/A	relational	overview	tukey box plot
320	1D	N/A	N/A	trend	value bar
321	N/A	N/A	hierarchical	overview	venn diagram
322	N/A	N/A	relational	overview	venn diagram
323	N/A	N/A	N/A	overview	verbal step by step diagrams
324	multiD	both	hierarchical	N/A	vertically stacked line charts
325	N/A	N/A	N/A	compose	waterfall chart
326	2D	quantitative	N/A	Compare or trend	analytical table
327	3D	quantitative	N/A	Compare or trend	analytical table
328	multiD	quantitative	N/A	Compare or trend	analytical table
329	2D	quantitative	N/A	distributi on	analytical table
330	3D	quantitative	N/A	distributi on	analytical table
331	multiD	quantitative	N/A	distributi on	analytical table
332	N/A	quantitative	N/A	Compare	proportional area graph
333	2D	categorical or quantitative	N/A	compose or trend	area graph
334	3D	categorical or quantitative	N/A	compose or trend	area graph
335	multiD	categorical or quantitative	N/A	compose or trend	area graph
336	N/A	both	N/A	Compare	bar graph
337	N/A	quantitative	temporal	Compare	bar graph
338	1D, 2D, 3D, multiD	categorical or quantitative	N/A	Compare	grouped bar graph
339	1D, 2D, 3D, multiD	quantitative	N/A	distributi on	box plot
340	3D	categorical or	N/A	Compare	bubble graph

		quantitative			
		categorical or			
341	multiD	quantitative	N/A	Compare	bubble graph
342	N/A	text	N/A	detail	bullet text chart
343	multiD	N/A	N/A	Compare	chernoff faces
344	N/A	quantitative	N/A	correlate or trend	area graph
345	N/A	quantitative	N/A	correlate or trend	bar graph
346	N/A	quantitative	N/A	correlate or trend	pie chart
347	N/A	quantitative	N/A	correlate or trend	column graph
348	N/A	quantitative	N/A	correlate or trend	line graph
349	N/A	quantitative	N/A	correlate or trend	nomogram
350	N/A	quantitative	N/A	correlate or trend	polar graph
351	N/A	quantitative	N/A	correlate or trend	radar graph
352	N/A	quantitative	N/A	correlate or trend	scatterplot
353	N/A	quantitative	temporal	correlate or trend	area graph
354	N/A	quantitative	temporal	correlate or trend	candlestick chart
355	N/A	quantitative	temporal	correlate or trend	cash flow graph
356	N/A	quantitative	temporal	correlate or trend	column graph
357	N/A	quantitative	temporal	correlate or trend	line graph
358	N/A	quantitative	temporal	correlate or trend	control chart
359	N/A	quantitative	temporal	correlate or trend	run chart
360	N/A	quantitative	N/A	correlate or trend	pie chart
361	N/A	quantitative	N/A	correlate or trend	venn diagram
362	N/A	quantitative	N/A	correlate or trend	analytical table
363	N/A	quantitative	temporal	correlate or trend	analytical table
364	N/A	quantitative	temporal	correlate or trend	spreadsheets
365	N/A	quantitative	N/A	distributi on	pie chart
366	N/A	both	N/A	distributi on	box plot
367	N/A	both	N/A	distributi on	histogram
368	N/A	both	N/A	distributi	quantile-quantile

				on	plot
369	N/A	both	N/A	distributi on	pareto chart
370	N/A	both	N/A	distributi on	dot plot
371	N/A	both	N/A	distributi on	geographic map
372	N/A	both	N/A	distributi on	topological map
373	N/A	both	N/A	distributi on	network diagram
374	N/A	both	N/A	distributi on	analytical table
375	N/A	both	N/A	distributi on	spreadsheets
376	N/A	categorical	temporal	correlate	cpm critical path method
377	N/A	categorical	temporal	correlate	pert chart
378	N/A	categorical	temporal	correlate	timeline
379	N/A	categorical	temporal	overview	calendar
380	N/A	categorical	temporal	correlate	gantt chart
381	N/A	categorical	N/A	correlate	network diagram
382	N/A	categorical	N/A	correlate	flowchart
383	N/A	categorical	N/A	correlate	venn diagram
384	N/A	categorical	N/A	correlate	graphical table
385	multiD	N/A	N/A	Compare	radar graph
386	N/A	quantitative	temporal	Compare	column graph
387	1D	both	N/A	Compare	column graph
388	2D	both	N/A	Compare	stacked column graph
389	3D	both	N/A	Compare	stacked column graph
390	multiD	both	N/A	Compare	stacked column graph
391	2D	quantitative	N/A	correlate	scatterplot
392	2D	quantitative	N/A	correlate	scatterplot matrix
393	3D	quantitative	N/A	correlate	scatterplot matrix
394	multiD	quantitative	N/A	correlate	scatterplot matrix
395	N/A	categorical	temporal	correlate	cpm critical path method
396	N/A	quantitative	spatial	overview	shaded map
397	1D, 2D, 3D, multiD	N/A	N/A	distributi on	point plot
398	1D, 2D, 3D, multiD	quantitative	N/A	distributi on	histogram
399	N/A	N/A	N/A	overview	decision tree
400	N/A	categorical or text	N/A	correlate	flowchart
401	N/A	categorical or text	temporal	correlate	flowchart
402	N/A	categorical or text	relational, spatial, tabular, temporal, spatio-temporal	overview	flowchart

403	N/A	categorical or text	relational, spatial, tabular, temporal, spatio-temporal	detail	flowchart
404	N/A	N/A	spatial	compose	geologic map
405	2D	quantitative	N/A	N/A	line graph
406	2D	both	N/A	N/A	line graph
407	2D	quantitative	temporal	N/A	line graph
408	1D	categorical	N/A	overview	histogram
409	1D	quantitative	N/A	overview	histogram
410	1D	categorical	N/A	overview	histogram
411	2D	N/A	N/A	trend	scatterplot
412	2D	N/A	N/A	detail	scatterplot
413	3D	N/A	N/A	N/A	scatterplot matrix
414	multiD	N/A	N/A	N/A	scatterplot matrix
415	multiD	N/A	N/A	N/A	chernoff faces
416	1D	quantitative	N/A	distributi on	dot plot
417	1D	N/A	N/A	distributi on	box plot
418	1D	quantitative	N/A	distributi on	histogram
419	2D	quantitative	N/A	correlate	scatterplot
420	2D	quantitative	N/A	correlate	parallel coordinates
421	3D	quantitative	N/A	correlate	parallel coordinates
422	multiD	quantitative	N/A	correlate	parallel coordinates
423	2D	quantitative	N/A	correlate	scatterplot matrix
424	3D	quantitative	N/A	correlate	scatterplot matrix
425	multiD	quantitative	N/A	correlate	scatterplot matrix
426	2D	quantitative	N/A	correlate	quantile plot
427	N/A	N/A	temporal	detail	linear planing chart
428	2D	N/A	temporal	overview	bar graph
429	N/A	N/A	temporal	overview	network diagram
430	N/A	N/A	temporal	Compare	linear planing chart
431	multiD	N/A	spatio-temporal	overview	3D building model
432	multiD	N/A	spatio-temporal	Compare	3D building model
433	N/A	N/A	temporal	Compare	TbarView
434	N/A	N/A	temporal	Compare	multiple Gantt charts overlayed
435	multiD	N/A	spatio-temporal	overview	3D building model
436	multiD	N/A	spatial	overview	map
437	3D	N/A	temporal	Compare	gantt chart
438	3D	N/A	temporal	correlate	gantt chart
439	N/A	N/A	temporal	overview	gantt chart
440	multiD	N/A	spatial	overview	3D building model
441	N/A	N/A	hierarchical	overview	tree
442	N/A	N/A	relational	overview	network diagram
443	multiD	both	N/A	distributi on	3D bar graph
444	multiD	both	N/A	distributi on	3D area graph

445	multiD	both	N/A	compose	combined scatter and pie
446	multiD	both	N/A	trend	combined scatter and pie
447	1D	N/A	N/A	N/A	histogram
448	1D	N/A	N/A	N/A	pie chart
449	1D	N/A	N/A	N/A	box plot
450	2D	N/A	N/A	N/A	multiple histograms
451	2D	N/A	N/A	N/A	scatterplot
452	3D	N/A	N/A	N/A	scatterplot matrix
453	3D	N/A	N/A	N/A	multiD glyphs
454	multiD	N/A	N/A	overview	multiD glyphs
455	multiD	quantitative	N/A	Compare	chernoff faces
456	multiD	N/A	N/A	Compare	chernoff faces
457	1D	N/A	temporal	N/A	timeline
458	multiD	N/A	temporal	trend	combined scatter and bar
459	N/A	N/A	hierarchical	correlate	tree
460	N/A	N/A	hierarchical	overview	treemap
461	N/A	N/A	relational	overview	circular network
462	multiD	quantitative	N/A	N/A	table lens
463	multiD	categorical or quantitative	N/A	correlate	scatterplot matrix
464	multiD	categorical or quantitative	N/A	N/A	2D scatterplot
465	multiD	categorical or quantitative	N/A	N/A	3D scatter plot
466	multiD	categorical or quantitative	N/A	cluster	self-organizing maps
467	multiD	categorical or quantitative	N/A	N/A	line graph
468	multiD	quantitative	N/A	N/A	polar graph
469	multiD	quantitative	relational	N/A	radviz
470	multiD	quantitative	N/A	correlate	survey plot
471	multiD	categorical or quantitative	N/A	overview	heat map
472	multiD	categorical or quantitative	N/A	overview	Height maps
473	multiD	categorical	N/A	N/A	Dimensional stacking grid
474	multiD	categorical or quantitative	N/A	correlate	circle segments
475	multiD	quantitative	N/A	N/A	Sammon plots: multiD scaling
476	multiD	quantitative	N/A	distributi on	PolyViz
477	multiD	quantitative	N/A	N/A	scatterplot with project pursuit
478	multiD	N/A	N/A	N/A	chernoff faces
479	multiD	N/A	N/A	N/A	star glyphs
480	multiD	quantitative	N/A	N/A	parallel coordinates
481	N/A	N/A	spatial	N/A	map

482	N/A	N/A	tabular	N/A	multicolumn lists
483	N/A	N/A	tabular	N/A	sortable table
484	N/A	N/A	tabular	N/A	spreadsheets
485	N/A	N/A	tabular	N/A	multi-y plot
486	N/A	N/A	hierarchical	N/A	cascade list
487	N/A	N/A	hierarchical	N/A	directed graph
488	multiD	N/A	hierarchical	N/A	treemap
489	N/A	N/A	relational	N/A	directed graph
490	N/A	N/A	relational	N/A	flowchart
491	multiD	both	N/A	N/A	surface graph

Table D.2: List of Visual Representations

Higher Category	Visual Rep	Higher Category	Visual Rep
andrew			
curves	andrew curves	area graph	3D area graph
animation	animation		area graph
boxplot	multiple box plot		circular area graph
	tukey box plot		stacked 100% area graph
	box plot		stacked area graph
daisy maps	daisy maps		surface graph
data vases	data vases		proportional area graph
E-R diagram	entity relationship diagram	bar graph	tileBars
flow chart	affinity diagram		waterfall chart
	decision discovery diagram		3D bar graph
	feedback diagram		alligned bar graph
	flow chart		bar graph
	organizational chart		divided bar graph
	sankey diagram		stacked bar graph
	schematic diagram		grouped bar graph
	sequence chart	bubble graph	bubble graph
	ststem dyn/ loop diagram	column graph	stacked column
	supply-demand chain		column graph
	time and activ. charts		grouped column graph
	verbal step by step diagrams		stacked 100% column graph
	cascade list		stacked column graph
heat map	heat map	point graphic	column histogram
histogram	histogram		curve density plot
	multiple histograms		dot plot
icons	chernoff faces		histogram
	needle icons		histogram
	star glyphs		line histogram
	star icons		point plot
	multiD glyphs		step chart
map	color map	Arrow diagram	cpm critical path method
	flow maps		gantt chart
	geo-spatial map		pert chart
	heat map		multiple Gantt charts overlayed
	Height maps	quantile plot	quantile plot

	isogram graph		quantile-quantile plot
	isopleths	venn diagram	venn diagram
	landscape	bullet text chart	bullet text chart
	map	nomogram	nomogram
	netmap	linear planning chart	linear planing chart
	space-time cube	combined	combined scatter and pie
	topological map		combined scatter and bar
	geographic map		scatterplot with project pursuit
	shaded map		multi-y plot
	geologic map	combined	3D kiviak tube: line graph with stacked glyphs
network diagram	cognitive mapping		column graph combined with curves
	concept map	circle segment	circle segments
	data model	polyViz	PolyViz
	document lens	. Graph	pareto chart
	knowledge map		polar graph
	magic eye view		candlestick chart
	mindmap		cash flow graph
	netmap		control chart
	network diagram		run chart
	radviz		calendar
	semantic network	parallel coordinates	parallel coordinates
	themeRiver		temporal parallel coordinates
	circular network	pareto chart	pareto chart
	directed graph	pie chart	combined pie charts
line graph	3D line		pie chart
	3D time wheel - stacked line graphs	scatterplot	bcg matrix
	circle segments		multiple scatter graphs
	cycle diagram		scatterplot
	fan		scatterplot map
	kite diagram		scatterplot matrix
	polar coordinate		Survey plot
	spiral graph		triple scatterplot
	themeRiver		value bar
	timeline		3D scatter graph
	timeWheel		2D scatterplot
	vertically stacked line charts		3D scatter plot
	2D line graphs		Sammon plots: multiD scaling
	3D line graph	self-organizing maps	self-organizing maps
	line graph		tag clouds
	polar graph	spectrogram	spectrogram
	TbarView	tree	cone tree
	radar graph		decision tree
	line graphs		disk tree
table	graphical table		hyperbolic tree
	information cube		hyperbolic view
	multicolumn lists		tree
	relationalTable		treemap
	sortable table		

spreadsheets
table
table with colors
text table
analytical table
table lens
Dimensional stacking grid

Table D.3: The “constraints/Compares” heuristics

statement	visual representation / mark type related
it is possible to display only two or three variables on a single scatterplot	scatterplot
it is possible to map upto 18 data features onto chernoff faces	chernoff faces
Line charts are best practice for temporal data.	line charts
for two data fields if the following combinations are satisfied the mark type and view type should be selected:	
categorical and categorical	cross-tab text
quantitative (measure) and categorical	bar view bar
quantitative (dependent) and categorical (date)	line view line
quantitative (dependent) and quantitative (dependent)	scatterplot shape
quantitative (independent) and categorical	gantt view gantt
quantitative (independent) and quantitative (dependent)	line view line
quantitative (independent) and quantitative (independent)	scatterplot shape
gantt charts are effective in showing duration with respect to a quantitative independent	gantt chart
compare two values, particularly with categorical fields on the shape shelf, effective is scatter plot	scatterplot
aligned bars are better default than side-by-side vards because aligned is easier to compare visually	bar chart
a dependent measure (such as profit or number of network packets) varies with two independent hierarchical ordinal dimen- sions, one derived from continuous data (such as time). This type of data can be effectively visualized using a vertically stacked small multiple of line charts (e.g., a single column of charts) [Ex3].	vertically stacked small multiple of line charts

geographically- varying dependent measures that can be summarized at multiple geographic levels of detail (such as county or state). Thus, the data contains an ordinal dimension hierarchy that characterizes the geographic levels of detail, two independent spatial dimensions (e.g., latitude and longitude), and some number of dependent measures.	map	color of area, glyphs
visual ranking for quantitative data	any	all of the marks
visual ranking for ordinal data	any	all of the marks
visual ranking for nominal data	any	all of the marks
Bar chart is better than divided bar chart and pie chart – empirical (convey relative magnitude more accurate)	bar chart and pie chart	
Dot chart is better than divided bar chart and pie chart – empirical	dot chart and pie chart	
Grouped dot chart is better than grouped bar chart – empirical	dot chart and pie chart	
Grouped bar chart is better than divided bar chart – empirical	bar charts	
Framed rectangle charts are better than statistical maps with shading – empirical	framed rectangle charts and statistical maps	
More quickly with the normal tree and for shallow browsing	3D cone tree versus normal tree	
Smaller than 500 nodes, a main effect of browsers n task correctness was found, but more than 3000 no effect	Tree maps versus SunBurst	
	ordered tree maps	
the strip was the fastest, followed by the pivot, and then squarified	SpaceTree with Windows Explorer and hyperbolic view visualization	
People found 3D cool but preferred 2D	SpaceTree with Windows Explorer and hyperbolic view visualization	
According to being fast: First two tasks: explorer, space and hyperbolic, for the third space, hyperbolic and explorer	bar chart, table, line chart	
each faster for corresponding task	excel chart	
Excel charts of 16 related timeseries (Australian yields) over 1100 days. The 3D chart is not useful. The 2D chart does provide some value, but suffers from occlusion.	s-scatter	
	excel chart	
Excel chart viewing 30 timeseries each with 2500+ intervals. Broad patterns are visible, but there is a lot of occlusion in some areas	s-scatter	
Excel Visualizer viewing 16 timeseries over 1100 days (Australian yields) OK	visualizer	
dynamic treeView and dynamic treeMap have similar accuracy and user satisfaction levels	dynamic treeView and dynamic treeMap	

dynamic treeView is faster than dynamic treeMap	dynamic treeView and dynamic treeMap
dynamic treeView with attribute ordering is faster than and has higher user satisfaction than with alphabetical ordering	dynamic attribute ordering and alphabetical ordering
dynamic spiral timeline is faster, more accurate, and higher user satisfaction compared to linear timeline	dynamic spiral timeline and linear timeline
dynamic multi-resolution network has higher user satisfaction than dynamic graphView	dynamic network and graphView
dynamic graphView has higher user performance than spreadsheets	dynamic graphView and spreadsheet
increased accuracy of predictions are acquired using multidimensional representations of multidimensional data (3D chart rather than 2D chart)	3D vs 2D
static visual representations (e.g., R1 and R2) may take less time to support decision making than dynamic visual representations	any
the combined time and accuracy results suggest that the static 3D (R2) visual representation may be more effective than the static 2D (R1) or 3D rotatable (R3) visual representations.	any
as the complexity of the graph increases, the time taken to answer questions based on the graphs increases	2D line chart
as complexity increases, the time taken to answer questions using 2D graphs increases more rapidly than for 3D graphs	2D / 3D line charts
when provided with support of dynamic labeling, 3D graphs outperform 2D graphs under all experimental conditions	any 9 or 25 variables
bar chart is faster than pie chart and divided bar chart for Compare task	bar, pie, divided bar
bar charts are more accurate than pie charts for Compare of parts	bar pie
while they are good for both of the purposes, line charts emphasize trend as a result of change rather than the quantities, where as surface charts emphasize the quantities rather than the trends	line and surface charts

7.5 Appendix E: Survey

Survey for Construction Project Monitoring

The purpose of this survey is to observe individual perspectives of different project members in selecting visualization techniques for monitoring the performance of construction project. The survey would take around 10 min. The survey starts with 5 questions designed to acquire information regarding your job description and experience. In the second section of the survey, we present 6 questions within which you will be requested to rate 2 visual representations, in terms of their effectiveness in conveying the given information. Finally in the last section, you will be asked 6 question in order to acquire more insight about your perspective on the way project information is/ should be visualized and interacted.

* Required

Background Information

What is your name? *

What is your responsibility within the company *

Currently, on what type of project are you working on? *

- ☐ Residential Building
- ☐ Commercial Building
- ☐ Industrial
- ☐ Heavy Civil
- ☐ Other:

For how long you have been working in construction industry. (years of experience) *

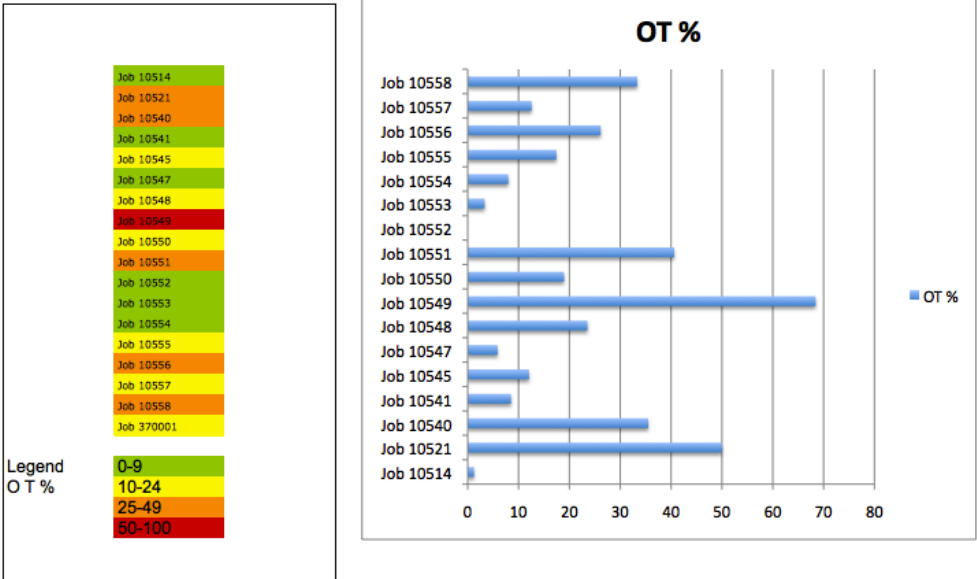
What is the frequency that you monitor the performance of the project that you are working on? in other words, how frequently you would like to have access to the performance information. You can select more than one box considering different performance metrics.

- ☐ Always
- ☐ Always with notifications for the prominent information
- ☐ As soon as the information is available
- ☐ Based on a predetermined time frame, for instance once a day, once a week etc.
- ☐ Other:

Questions Regarding Selection of Visual Representations

Scenario: You are working on a self-performed commercial building construction and you need to monitor different performance aspects of the project. Given the following performance metric queries, please answer each of the questions based on the clarity and effectiveness of the visual representation in conveying the requested information.

Question 1: You are responsible to oversee the productivity of multiple projects. You want to get an overview of the multi-project weekly overtime factor. Please indicate whether the presented visual representations are effective to represent such information



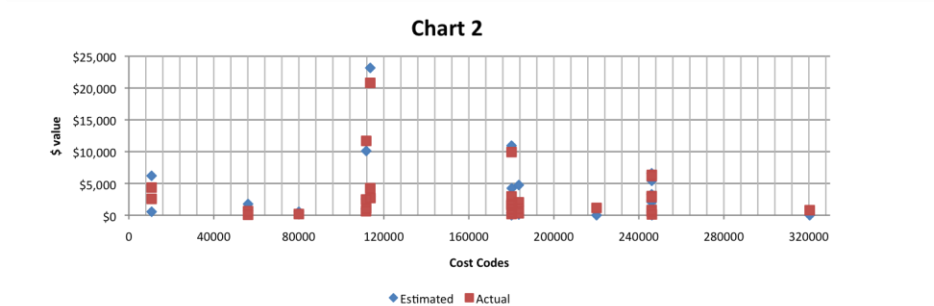
or not.

Definitely effective Effective Neither Ineff. Eff. Nor Not effective Definitely effective not

Chart 1

Chart 2

Question2: You are responsible to oversee the cost of a single project. You want to compare the actual vs. budgeted cost of the activities that are performed the previous week. Please indicate whether the presented visual representations are effective to represent such information or not.

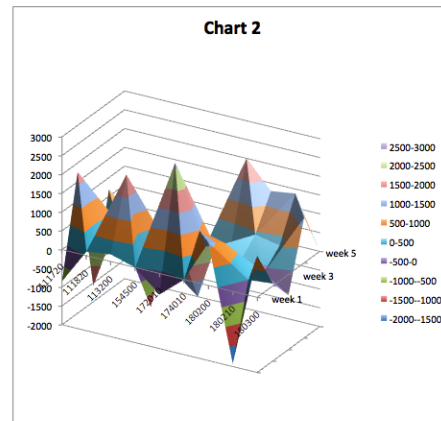
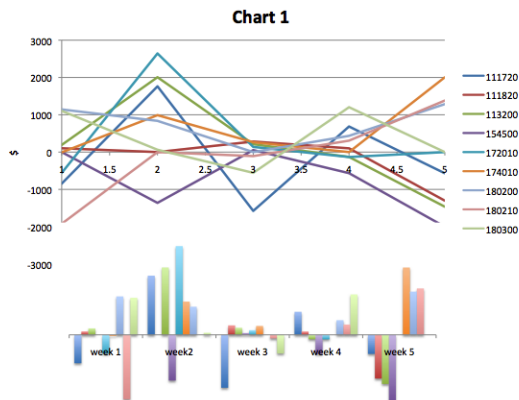


Definitely effective Effective Neither Ineff. Eff. Nor Not effective Definitely effective not

Chart 1

Chart 2

Question 3: You are responsible to oversee the cost of a single project. You want to analyze the trend in cost overrun for each of the activities that are performed to date. Please indicate whether the presented visual representations are effective to represent such information or not.



Definitely effective Effective Neither Ineff. Eff. Nor Not effective Definitely effective not

Chart 1

Chart 2

Question 4: You are responsible to oversee the productivity of a single project. You want overview the space usage for utilities, work areas and etc. Please indicate whether the presented visual representations are effective to represent such information or not.

Chart1

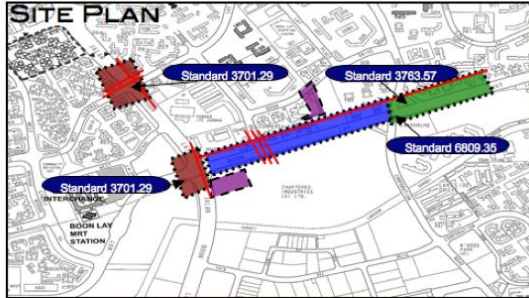


Chart 2

Utilities, Subs Code	Units	station #	add kms.	end kms.
10730 cy	0.3	142	668	
40065 ls	0	0	1044	
41014 lf	16	221	532	
41032 tn	21.4	355	84	
53104 ea	100	229	860	
53969 ea	27.7	573	404	
53970 vf	0	0	1341	
53977 vf	0	0	529	
53979 ea	7.1	115	431	
54262 sf	0.4	288	85	
54265 cy	1.5	324	181	
54266 ea	9.2	212	905	
82154 cy	35.7	54	301	
82164 cy	34.4	150	257	
82194 cy	0	0	48	
85710 lf	18.3	4413	3691	
85722 ea	25	1044	745	
111720 lb	0.5	458	1254	
140270 sy	0	0	170	
170030 ls	0	0	276	
170245 sy	60.3	604	761	
180010 cy	6.1	2749	3290	
180105 sf	20.8	8998	7312	
180116 sf	62.1	9184	6229	
180140 sf	76.7	18242	8513	
180207 sf	3.3	1028	1198	
180305 lb	0	0	180	
220036 lf	0	0	123	
246045 cy	11.3	4144	3424	

Definitely effective Effective Neither Ineff. Eff. Nor Not effective Definitely not effective

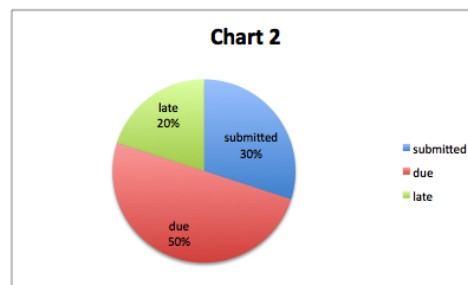
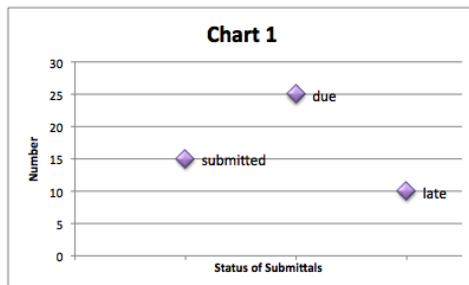
Chart

1

Chart

2

Question 5: You are responsible to oversee the communication of documentations for a single project. You want to analyze the distribution of the submittals based on their submission rate and status. Please indicate whether the presented visual representation are effective to represent such information or not.



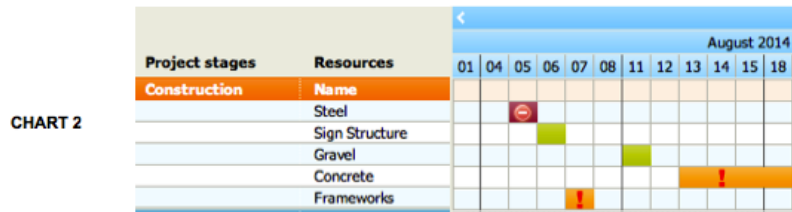
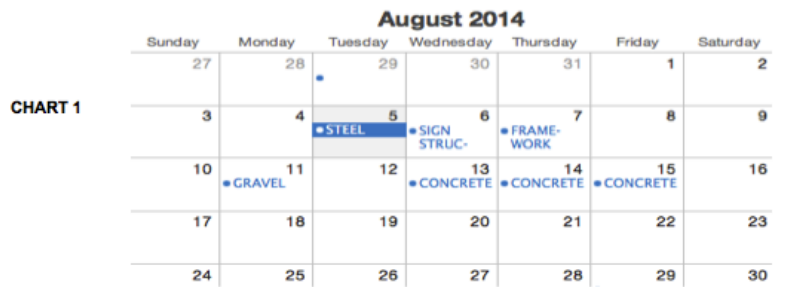
Definitely effective Effective Neither Eff. Nor Ineff. Not effective Definitely effective not effective

Chart 1

Chart 2

Question 6: You are responsible to oversee the productivity of a single project. You want overview the weekly material procurement schedule. Please indicate whether the

presented visual representation are effective to represent such information or not.



Definitely effective Effective Neither Ineff. Eff. Nor Not effective Definitely effective not

Chart 1

Chart 2

What information items are required to define what to monitor

This section is composed of questions to understand what factors did you consider while selecting the visual representations, and if you think further suggestions could be used to improve such visualization or not.

What are the factors that you considered for the selection of visual representations in the previous questions. Since you were not requested to performed data reading tasks for the previous questions, please answer these questions assuming that if you were to read data (cost or rate) for a given day or activ. would these visual representations satisfy your need?

- ☐ Simplicity/ Visual Clutter
- ☐ Ease / Difficulty in Read Data
- ☐ Accuracy of the data
- ☐ Compliment with the intent (make a Compare, or get an overview...)
- ☐ Compliment with the data type (temporal, numerical...)
- ☐ Other:

Would it be more effective if notifications were enabled for the visual representations? For instance, highlighting the outliers, or alert when data is available.

Would it be more effective if interactions with the visual representations were enabled?
For instance, sorting, linking, filtering and etc.

Would you select a different type of visual representation if the level of the detail of the performance queries were different. For instance if you were to track the total cost of the the project rather than each material and equipment cost of activities, would you make a different selection?

- ☐ No, I would select the same visual representation no matter the level of detail
- ☐ I do not know
- ☐ Maybe, I need to compare the visual representations
- ☐ Sure, I would select a different one.

Are you happy with the current practice of construction data monitoring using existing software tools (Excel etc.) with standard visual representations? Would you find it useful if had access to a visualization environment where visual representations are suggested to you based on your queries?

- ☐ I do not use/ value any visualization technique
- ☐ Tables, reports and Gantt Charts are satisfactory for my needs
- ☐ I generate visualization techniques based on the type of the information I monitor, and satisfied with it
- ☐ I am using a software tool / dashboard to meet my needs
- ☐ I do not need but would like to try to visualization environment where visualization techniques are suggested based on the information I need.
- ☐ A visualization environment where visualization techniques are suggested based on the type of the information I am querying would increase my project awareness and / or save time.
- ☐ Other:

Each performance metric can be visualized to make different types of decisions. Their usage such as being Descpr. (describe the performance), Explan. (explain the reasons), prescr. (reveal action items for other parties), Predic. (predict future conditions) may vary based on the type of the information tracked. Would it improve your decision making process during the survey, if the decisions to be made were to be provided to you?

- ☐ Yes
- ☐ Not sure
- ☐ No
- ☐ Other:

7.6 Appendix F: Synthetic Tests

Appendix F is composed of 3 tables. Table F.1 presents the test cases according to their data dim, data var, data structure and intents. Table F.2 presents the number of retrieved *heuristics* for each of the test cases listed according to four facets of combinations of data characteristics and intent. Table F.3 presents the number of retrieved visual representations for each of the test cases listed according to four facets of combinations of data characteristics and intent. Table F.4 presents the identification frequency of each of the visual representations

Table F.1: Test Cases

Test ID	Dimension	Variable	Structure	Intent
1	2D	both	temporal or relational	Trend
2	2D	both	temporal	Compare
3	1D	quantitative	relational or tabular	Detail
4	multiD	both	hierarchical or relational or tabular	Compare
5	multiD	both	hierarchical or relational or tabular	Overview and Detail
6	2D	both	hierarchical or relational or tabular	Overview
7	multiD	both	temporal or hierarchical	Compare
8	multiD	quantitative	hierarchical or relational or tabular	Overview
9	multiD	categorical	hierarchical or relational or tabular	Overview
10	multiD	both	temporal	Overview
11	multiD	quantitative	temporal	Trend
12	multiD	both	hierarchical or relational or tabular	Overview and Detail
13	3D	both	hierarchical or relational or tabular	Overview
14	1D	text	relational or tabular	Overview
15	2D	both	hierarchical or relational or tabular	Overview
16	2D	both	hierarchical or relational or tabular	Overview and Detail
17	multiD	both	temporal	Overview
18	multiD	quantitative	temporal	Trend
19	multiD	both	temporal	Overview
20	multiD	both	temporal	Trend
21	3D	both	relational or tabular	Overview
22	1D	quantitative	temporal	Trend
23	2D	quantitative	temporal	Compare
24	1D	quantitative	relational or tabular	Detail
25	3D	categorical	spatio-temporal	Overview
26	1D	quantitative	relational or tabular	Detail
27	2D	both	relational or tabular	Compare
28	2D	both	temporal or relational	Compare

29	2D	quantitative	relational or tabular	Compare
30	multiD	both	temporal or relational	Trend
31	multiD	both	relational or tabular	Overview
32	1D	both	spatio-temporal	Overview
33	3D	categorical	spatio-temporal	Overview
34	1D	text	relational or tabular	Overview
35	multiD	both	temporal	Trend
36	multiD	both	temporal	Trend
37	multiD	both	temporal or relational	Compare
38	1D	both	spatio-temporal	Overview
39	1D	text	relational or tabular	Detail
40	multiD	categorical	spatial	overview
41	1D	both	temporal	Overview
42	multiD	both	spatio-temporal	Overview
43	multiD	quantitative	temporal	Compare
44	1D	categorical	Relational	Compose
45	2D	both	relational or tabular	Compare
46	1D	categorical	relational or tabular	Compose
47	2D	quantitative	relational or tabular	Correlate
48	multiD	both	relational or tabular	Overview
49	2D	quantitative	relational or tabular	Correlate
50	multiD	both	relational or tabular	Overview
51	multiD	both	Relational	Cluster
52	2D	text	relational or tabular	Detail
53	multiD	both	spatio-temporal	Overview and Detail
54	3D	both	spatio-temporal	Overview
55	1D	both	spatio-temporal	Overview
56	3D	categorical	spatio-temporal	Overview
57	multiD	both	relational or tabular	Overview and Detail
58	multiD	both	relational or tabular	Overview and Detail
59	multiD	both	relational or tabular	Overview
60	multiD	both	relational or tabular	Overview and Detail
61	multiD	both	relational or tabular	Overview
62	1D	text	relational or tabular	Overview
63	2D	both	relational or tabular	Correlate
64	multiD	both	relational or tabular	Overview
65	multiD	both	relational or tabular	Correlate
66	1D	text	relational or tabular	Detail
67	multiD	both	temporal	Overview
68	multiD	both	temporal	Compare
69	3D	both	spatial	Overview
70	multiD	both	relational or tabular	Overview and Detail
71	multiD	both	relational or tabular	Overview and Detail
72	2D	both	relational or tabular	Cluster
73	2D	both	tabular	Overview
74	1D	categorical	temporal	Overview and Detail
75	multiD	both	temporal	Overview
76	multiD	both or text	relational or tabular	Overview
77	multiD	both	relational or tabular	Compare
78	multiD	both	hierarchical or relational or tabular	Compare
79	1D	quantitative	relational or tabular	Overview
80	multiD	both	temporal	Overview
81	multiD	both	temporal	Overview

82	multiD	both	relational or tabular	Overview
83	multiD	both	relational or tabular	Overview
84	multiD	both	relational or tabular	Overview
85	multiD	both	temporal	Compare
86	3D	both	spatio-temporal	Overview
87	multiD	both	temporal	Overview
88	multiD	both	relational or tabular	Overview
89	multiD	both	relational or tabular	Overview and Detail
90	multiD	both	relational or tabular	Overview
91	2D	both	temporal	Compare
92	2D	both	temporal or relational	Trend
93	multiD	both or text	relational or tabular	Overview
94	multiD	both	relational or tabular or spatial	Overview
95	2D	categorical or text	spatial	Overview
96	1D	text	tabular	Overview
97	2D	both	temporal	Compare
98	multiD	both	relational or tabular	Overview
99	multiD	both	hierarchical or relational or tabular	Compare
100	2D	text	relational or tabular	Detail
101	1D	text	tabular	Overview
102	3D	categorical	spatio-temporal	Overview
103	1D	both	spatio-temporal	Overview
104	3D	both	temporal or relational	Trend
105	multiD	both	relational or tabular	Overview and Detail
106	1D	quantitative	relational or tabular	Distribution
107	multiD	both	temporal	Compare
108	1D	categorical	relational or tabular	Distribution
109	3D	categorical	spatio-temporal	Overview
110	1D	quantitative	relational or tabular	Distribution
111	1D	text	tabular	Overview
112	1D	text	relational or tabular	Overview
113	multiD	both	relational or tabular	Overview
114	multiD	both	relational or tabular	Trend
115	1D	text	relational or tabular	Overview
116	multiD	both	spatio-temporal	Overview
117	multiD	both	relational or tabular	Overview
118	multiD	both	relational or tabular	Overview
119	1D	categorical	temporal	Overview and Detail
120	multiD	both	hierarchical or relational or tabular	Compare
121	2D	text	relational or tabular	Detail
122	1D	text	tabular	Detail
123	1D	text	tabular	Detail
124	multiD	both	spatial	Overview
125	multiD	both	relational or tabular	Overview
126	multiD	both	relational or tabular	Compare
127	multiD	both	temporal or relational	Trend
128	multiD	both	relational or tabular	Compare
129	multiD	both	relational or tabular	Overview and Detail
130	multiD	both	relational or tabular	Overview and Detail
131	1D	both	temporal	Overview
132	2D	both	temporal	Compare
133	multiD	both	temporal	Overview
134	multiD	quantitative	temporal	Trend

135	multiD	both	relational or tabular	Overview
136	multiD	both	hierarchical or relational or tabular	Compare
137	multiD	both	temporal	Overview
138	multiD	both	relational or tabular	Overview

Table F.2: The number of retrieved heuristics for each of the test cases listed according to four facets of combinations of data characteristics and intent.

Test Case ID	4s	3s	2s	1s
1	0	0	4	23
2	0	4	17	25
3	0	0	11	12
4	0	1	10	40
5	0	0	4	27
6	0	0	1	24
7	0	2	12	40
8	0	3	9	40
9	0	3	7	37
10	0	2	15	35
11	2	17	27	34
12	0	0	4	27
13	0	0	4	22
14	0	1	1	18
15	0	0	1	24
16	0	0	4	14
17	0	2	15	35
18	2	17	27	34
19	0	2	15	35
20	0	3	5	31
21	0	0	4	22
22	0	10	33	14
23	0	5	31	28
24	0	0	11	12
25	0	1	1	18
26	0	0	11	12
27	0	1	6	27
28	0	1	6	30
29	0	2	17	30
30	0	1	4	36
31	0	0	2	37
32	0	0	0	13
33	0	1	1	18
34	0	1	1	18
35	0	3	5	31
36	0	3	5	31
37	0	1	10	43
38	0	0	0	13
39	0	1	1	10
40	0	6	10	35
41	0	0	14	15

42	0	2	2	33
43	2	9	30	41
44	0	0	0	6
45	0	1	6	27
46	0	0	0	10
47	0	7	29	20
48	0	0	2	37
49	0	7	29	20
50	0	0	2	37
51	0	0	3	23
52	0	1	2	17
53	0	0	4	23
54	0	0	4	18
55	0	0	0	13
56	0	1	1	18
57	0	0	4	27
58	0	0	4	27
59	0	0	2	37
60	0	0	4	27
61	0	0	2	37
62	0	1	1	18
63	0	0	3	17
64	0	0	2	37
65	0	2	3	30
66	0	1	1	10
67	0	2	15	35
68	0	2	20	38
69	0	0	7	20
70	0	0	4	27
71	0	0	4	27
72	0	0	1	14
73	0	0	1	24
74	0	0	10	5
75	0	2	15	35
76	0	1	2	38
77	0	1	10	40
78	0	1	10	40
79	0	1	11	20
80	0	2	15	35
81	0	2	15	35
82	0	0	2	37
83	0	0	2	37
84	0	0	2	37
85	0	2	20	38
86	0	0	4	18
87	0	2	15	35
88	0	0	2	37
89	0	0	4	27
90	0	0	2	37
91	0	4	17	25
92	0	0	4	23
93	0	1	2	38
94	0	2	5	35

95	0	1	4	23
96	0	1	1	18
97	0	4	17	25
98	0	0	2	37
99	0	1	10	40
100	0	1	2	17
101	0	1	1	18
102	0	1	1	18
103	0	0	0	13
104	0	0	5	21
105	0	0	4	27
106	0	4	19	13
107	0	2	20	38
108	0	0	3	10
109	0	1	1	18
110	0	4	19	13
111	0	1	1	18
112	0	1	1	18
113	0	0	2	37
114	0	1	4	33
115	0	1	1	18
116	0	2	2	33
117	0	0	2	37
118	0	0	2	37
119	0	0	10	5
120	0	1	10	40
121	0	1	2	17
122	0	1	1	10
123	0	1	1	10
124	0	2	5	35
125	0	0	2	37
126	0	1	10	40
127	0	1	4	36
128	0	1	10	40
129	0	0	4	27
130	0	0	4	27
131	0	0	14	15
132	0	4	17	25
133	0	2	15	35
134	2	17	27	34
135	0	0	2	37
136	0	1	10	40
137	0	2	15	35
138	0	0	2	37

Table F.3: The number of retrieved visual representations for each of the test cases listed according to four facets of combinations of data characteristics and intent.

Test Case ID	4s	3s	2s	1s
1	0	0	3	17
2	0	4	12	16
3	0	0	11	11
4	0	1	9	26

5	0	0	4	18
6	0	0	1	21
7	0	2	10	24
8	0	3	9	29
9	0	3	7	26
10	0	1	14	24
11	1	15	22	22
12	0	0	4	18
13	0	0	2	20
14	0	1	1	18
15	0	0	1	21
16	0	0	4	11
17	0	1	14	24
18	1	15	22	22
19	0	1	14	24
20	0	3	5	20
21	0	0	2	20
22	0	10	25	12
23	0	5	24	18
24	0	0	11	11
25	0	1	1	16
26	0	0	11	11
27	0	1	5	19
28	0	1	5	21
29	0	2	15	22
30	0	1	4	25
31	0	0	2	26
32	0	0	0	13
33	0	1	1	16
34	0	1	1	18
35	0	3	5	20
36	0	3	5	20
37	0	1	9	28
38	0	0	0	13
39	0	1	1	9
40	0	6	10	24
41	0	0	12	15
42	0	1	2	22
43	1	9	27	26
44	0	0	0	5
45	0	1	5	19
46	0	0	0	9
47	0	4	21	17
48	0	0	2	26
49	0	4	21	17
50	0	0	2	26
51	0	0	3	14
52	0	1	2	14
53	0	0	4	14
54	0	0	2	16
55	0	0	0	13
56	0	1	1	16
57	0	0	4	18

58	0	0	4	18
59	0	0	2	26
60	0	0	4	18
61	0	0	2	26
62	0	1	1	18
63	0	0	3	14
64	0	0	2	26
65	0	1	3	21
66	0	1	1	9
67	0	1	14	24
68	0	2	18	24
69	0	0	5	18
70	0	0	4	18
71	0	0	4	18
72	0	0	1	11
73	0	0	1	21
74	0	0	10	5
75	0	1	14	24
76	0	1	2	27
77	0	1	9	26
78	0	1	9	26
79	0	1	11	20
80	0	1	14	24
81	0	1	14	24
82	0	0	2	26
83	0	0	2	26
84	0	0	2	26
85	0	2	18	24
86	0	0	2	16
87	0	1	14	24
88	0	0	2	26
89	0	0	4	18
90	0	0	2	26
91	0	4	12	16
92	0	0	3	17
93	0	1	2	27
94	0	2	5	24
95	0	1	4	20
96	0	1	1	18
97	0	4	12	16
98	0	0	2	26
99	0	1	9	26
100	0	1	2	14
101	0	1	1	18
102	0	1	1	16
103	0	0	0	13
104	0	0	3	17
105	0	0	4	18
106	0	3	13	12
107	0	2	18	24
108	0	0	3	9
109	0	1	1	16
110	0	3	13	12

111	0	1	1	18
112	0	1	1	18
113	0	0	2	26
114	0	1	4	23
115	0	1	1	18
116	0	1	2	22
117	0	0	2	26
118	0	0	2	26
119	0	0	10	5
120	0	1	9	26
121	0	1	2	14
122	0	1	1	9
123	0	1	1	9
124	0	2	5	24
125	0	0	2	26
126	0	1	9	26
127	0	1	4	25
128	0	1	9	26
129	0	0	4	18
130	0	0	4	18
131	0	0	12	15
132	0	4	12	16
133	0	1	14	24
134	1	15	22	22
135	0	0	2	26
136	0	1	9	26
137	0	1	14	24
138	0	0	2	26

Table F.4: The identification frequency of each of the visual representations

'2D line graphs'	28
'2D scatterplot'	7
'3D area graph'	1
'3D bar graph'	0
'3D building model'	31
'3D kiviatt tube: line graph with stacked glyphs'	12
'3D line'	11
'3D line graph'	24
'3D scatter graph'	11
'3D scatter plot'	7
'3D time wheel - stacked line graphs'	12
'Dimensional stacking grid'	2
'Height maps'	3
'PolyViz'	0
'Sammon plots: multiD scaling'	5
'Survey plot'	23
'TbarView'	9
'affinity diagram'	0
'alligned bar graph'	3
'analytical table'	32
'andrew curves'	154
'animation'	2

'area graph'	13
'bar graph'	219
'bcg matrix'	0
'box plot'	58
'bubble graph'	1
'bullet text chart'	7
'calendar'	0
'candlestick chart'	4
'cascade list'	1
'cash flow graph'	4
'chernoff faces'	402
'circle segments'	13
'circular area graph'	9
'circular network'	0
'cognitive mapping'	0
'color map'	3
'column graph'	103
'column graph combined with curves'	36
'column histogram'	1
'combined pie charts'	3
'combined scatter and bar'	6
'combined scatter and pie'	6
'concept map'	0
'cone tree'	0
'control chart'	4
'cpm critical path method'	0
'curve density plot '	3
'cycle diagram'	12
'daisy maps'	7
'data model'	7
'data vases'	8
'decision discovery diagram'	0
'decision tree'	75
'directed graph'	8
'disk tree'	0
'divided bar graph'	23
'document lens'	0
'dot plot'	2
'entity relationship diagram'	0
'fan'	11
'feedback diagram'	12
'flow maps'	0
'flowchart'	107
'gantt chart'	72
'geo-spatial map'	10
'geographic map'	107
'geologic map'	0
'glyphs'	6
'graphical table'	16
'grouped bar graph'	3
'grouped column graph'	23
'heat map'	69
'histogram'	3

'histogram '	72
'hyperbolic tree'	0
'hyperbolic view'	0
'information cube'	0
'isogram graph'	2
'isopleths'	0
'kite diagram'	4
'knowledge map'	0
'landscape'	74
'line graph'	161
'line histogram'	1
'linear planing chart'	9
'magic eye view'	0
'map'	13
'mindmap'	0
'multi-y plot'	79
'multiD glyphs'	48
'multicolumn lists'	79
'multiple Gantt charts overlayed'	9
'multiple box plot'	14
'multiple histograms'	23
'multiple scatter graphs'	2
'needle icons'	74
'netmap'	13
'network diagram'	12
'nomogram'	6
'organizational chart'	73
'parallel coordinates'	307
'pareto chart'	0
'pert chart'	19
'pie chart'	86
'point plot'	81
'polar coordinate'	11
'polar graph'	11
'proportional area graph'	3
'quantile plot'	4
'quantile-quantile plot'	23
'radar graph'	20
'radviz'	0
'relationalTable'	74
'run chart'	4
'sankey diagram'	0
'scatterplot'	262
'scatterplot map'	1
'scatterplot matrix'	261
'scatterplot with project pursuit'	5
'schematic diagram'	4
'self-organizing maps'	10
'semantic network'	0
'sequence chart'	2
'shaded map'	0
'sortable table'	79
'space-time cube'	4

'spectrogram'	0
'spiral graph'	4
'spreadsheets'	83
'stacked 100% area graph'	0
'stacked 100% column graph'	2
'stacked area graph'	9
'stacked bar graph'	13
'stacked column graph'	29
'star glyphs'	186
'star icons'	74
'step chart'	27
'ststem dyn/ loop diagram'	12
'supply-demand chain'	0
'surface graph'	78
'survey plot'	0
'table'	89
'table lens'	91
'table with colors'	23
'tag clouds'	11
'temporal parallel coordinates'	20
'text table'	10
'themeRiver'	27
'tileBars'	0
'time and activ. charts'	66
'timeWheel'	6
'timeline'	89
'topological map'	5
'tree'	0
'treemap'	2
'triple scatterplot'	11
'tukey box plot'	0
'value bar'	1
'venn diagram'	6
'verbal step by step diagrams'	66
'vertically stacked line charts'	1
'waterfall chart'	2