UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

INVESTIGATING THE IMPACT OF DATA VISUALIZATION BASED ON REAL TIME CONSTRUCTION PROJECT INFORMATION

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE IN CONSTRUCTION MANAGEMENT

By

RAFAEL ONOFRE VALLES CHACIN
Norman, Oklahoma
2016

INVESTIGATING THE IMPACT OF DATA VISUALIZATION BASED ON REAL TIME CONSTRUCTION PROJECT INFORMATION

A THESIS APPROVED FOR THE
COLLEGE OF ARCHITECTURE

BY

____
Dr. Tamera McCuen, Chair

____
Dr. Somik Ghosh

____
Mr. Malcolm Coetzee

Acknowledgements

I would like to express my sincere gratitude to my master's committee chair, Dr. Tamera McCuen, for encouraging me to explore new and innovative ways to apply skills and knowledge in a field that is meaningful to me and whose guidance and support helped me to make this study possible.

I also would like to express my special gratitude to my wife, Ingrid Chan, for her support, patience and understanding during the stressful hours completing this study, and to my family in Venezuela, for their positive thoughts and comments.

Table of Contents

List of Tables

List of Figures

Abstract

Managing construction projects is a complex operation. Multiple activities are executed at the same time, and numerous variables impact building development. Therefore, construction professionals need access to worksite information to understand the current state of the project; this will allow them to make better analyses and decisions that may improve project performance. Thus, the more information that decision-makers can obtain from the job site, the better they will be able to identify problems and find solutions. Current software applications allow data visualization for project control and monitoring. However, multiple elements are not supported by these systems, which makes them a poor communication mechanism. Data is not collected nor analyzed efficiently, and a thorough study of current project conditions cannot be performed, which may lead to poor results. Consequently, a system capable of classifying, organizing, and storing data was implemented to test the usability of data visualization systems in the interpretation of ongoing project progress based on activity data. The test demonstrated that construction data visualization provided valuable information to construction professionals and that the application of this system enabled the exploration of various data relationships significant for project management.

Chapter 1: Introduction

With the rapid evolution of computer technology, the construction industry is taking advantage of many of the strengths that computer-aided design (CAD) software provide for project development, from design to construction. Therefore, a considerable amount of data is being generated from multiple sources; for example, project design, planning and construction. Additionally, data is produced during construction, and as a consequence, there is constant reproduction of multi-dimensional data. For this reason, this study focused on the implementation of data visualization based on construction activity progress information during the project execution phase.

The architectural, engineering and construction (AEC) industry is currently facing a training crisis, with an aging workforce and little effort to utilize visualization techniques (Goedert et al., 2008). Although, there are significant advancements in the development of new and innovative graphic representations of construction projects. These improvements have helped the construction industry explore new ways to visualize projects before construction begins. However, little attention is focused on the visualization of data. Most of the effort that the computer industry has invested in the construction sector is in the visualization of spaces, structures, and construction sequences (Russel et al., 2009). For example, the AEC industry used to work only with two-dimensional (2D) construction documents because it was the format that was available to provide information. The AEC industry is currently transitioning to three-dimensional (3D) model projects. This new format of construction documents contains a substantial amount of data, which enables the combination of other systems to work together and produce new information. That is the case with four-dimensional

(4D) construction simulations, which integrate 3D models and construction schedules to visualize construction progress over time.

The data presented to project managers is complex, confusing, and configured in a way that requires experience and knowledge to interpret the information (Golparvar-Frad et al., 2009). The data characteristics that project managers receive could lead to poor communication with non-construction team members and the overlook of important situations (Lee and Rojas, 2013). Another problem that project managers face during project execution is that data is generated in multiple formats, such as images, written reports, and job site observations. The fragmentation of data formats hinders the growth of a robust system that allows a consistent structure of information.

Humans can obtain information better and faster if data is presented in a suitable visual format other than textual/numerical or tables (Russel et al., 2009). Additionally, as a creative process, visual analytics is also linked to the human process of understanding, knowledge creation, insight generation, and decision making (Keim et al., 2008). Also, construction teams would be able to test multiple data relationships that may not have been analyzed before.

According to Heer and Shneiderman (2012), visualizations provide a powerful means of making sense of data, although multiple factors need attention before producing data visualizations. The factors that help the creation of data visualizations may vary, from the tools available to gather the information to the time required recording the statuses of the activities and the other elements that impact project execution on the construction site. In the same way, there are substantial limitations that thwart the development of a practical solution that could facilitate data collection and visualization. Additionally, manual input from professionals when preparing construction documents, such as schedules and reports, increases data inconsistencies and errors. The lack of basic formats for data entry prevents the optimal use of the information and its integration with other systems. Therefore, well-structured data is required to take advantage of the information. Russel et al., (2009) observed that construction data is often poorly organized because the lack of proper grouping and sub-grouping which can lead to missed opportunities to associate related data, and more often than not it is incomplete (p. 1047)

These limitations in data structure create gaps between project control information and data analytics. Consequently, the current methods applied to analyzing the current project status are complicated and time-consuming. Chen and Teizer (2010) discussed that data is generated from multiple objects in the construction site and they may be isolated from each other. These objects may include the study of multiple documents, such as construction drawings, project schedules, reports, job site visits, and digital pictures. Chen and Teizer (2010) argued that traditional methods of project information delivery does not allow the project team to be able to make rapid a correct decision. Decision makers require that data be gathered from numerous sources, rigorous reasoning is performed, relationships among data are examined, and other types of analyses are performed that could provide them with a clear understanding of current project status (Song, 2005). It is a problematic process that requires constant project monitoring and a high level of data organization for the efficient use of information.

The reason why data visual analytics is considered so essential for decision-making is that it allows for the exploration of new and innovative alternatives to identify and solve problems, as stated in Cybulski, Keller, Nguyen, and Saundage‘s (2013) study. The implementation of visual analytics methodologies is expected to improve the execution of construction analyses. Therefore, a greater number of issues that are currently unexplored can be discovered and analyzed to increase productivity and performance on the construction site and, in the same way, to support monitoring, perform project diagnostics, and forecast construction conditions. These are considered critical components for the advancement of the industry. We are in the era of information, and data surrounds us in every activity that we perform. It is paramount that we implement data analysis in our day-to-day operations so we can obtain better outcomes for our actions. The ultimate goal of this research was to put data analytics into production, to identify practical uses, and to execute a rapid adoption of the system.

In order to explore the importance of data visualization further and make significant advancements using information contained in the construction schedule, it was necessary to create a structured data composition and provide clear and consistent data across the project duration. Establishing a set of rules would help considerably the optimization in the utilization of data for analyses, not only to follow activity progress but also to learn how multiple elements can influence the performance of project's activities.

As a result, it was possible to identify patterns in data syntax obtained from the project schedule. The lack of a standardized naming convention for the activities' names and the absence of a structure of the information hinder the expansion of the value of the job site data for in-depth analyses. The establishment of a logical syntax structure would enable the information collected from the job site to be transferred among systems. Doing so would provide the opportunity to create efficient data visualization systems. Structured and reliable data would increase the probability of using data for decision making and storing it as historical data for further analysis.

Also, the use of physical documents to record worksite data is not beneficial for the optimization of data management. There are critical limitations in the data transition from a physical to a digital format; one example is how handwriting may slow down the process for data input, which increases ambiguity and inconsistency. Additionally, manual input may impede automated data transfer. Digital images do not allow for data transfer, which makes it an unsuitable format for data collection. These limitations were the most consistent practices in the way data is currently collected during this study, and they prevented automation in the development of a database.

Chapter 2: Literature Review

Visual Analytics

New technological developments have made it possible to easily store and record information and conveniently manipulate this data for efficient analysis and the creation of applications that allow data visualization. According to Sun, Wu, Liang, and Liu (2013), “visual analytics employs interactive visualizations to integrate users’ knowledge and inference capability into numerical/algorithmic data analysis processes” (p. 852). At the same time, Cybulski, Keller, Nguyen, and Sundae (2013), stated that the interactive visual analytics process "is characterized as collaborative and social by nature as it comprises of data analyst, problem, and visual domain, who share ideas and actions during analytic activities" (p. 20). Equally important, Cybulski et al., (2013) described that creative problem-solving takes place when the analytics task involves data exploration to discover insights or solve problems. Root-Berstain (2004) said that highly creative people share common abilities, including acute skills of observation, imagination, abstracting, pattern recognizing, transforming, and synthesizing. Similarly, in different research, the author emphasized that mental images allowed analysts to transform their cognitive process into perceptual tasks, which are known to be more robust and more efficient than symbolic information processing (Brodbeck, Mezza and Leanne 2009, p. 28).

Visual analytics methodologies are implemented in multiple fields, and Sun et al., (2013) stated that this technique is used to solve real-world problems such as network traffic analysis, engaging education, concepts, sports analysis, database analysis, and biological data analysis. Furthermore, Sun et al., (2013) discussed that when using visual data exploration, the user directly interacts with the visual interface to analyze and explore the data. These statements support the concept that “digital technologies are opening up areas that were once reserved for experts and professionals” (Cybulski et al., 2013, p. 23). Healey and Enns (2012) state that a fundamental goal of visualization is to produce images of data that support visual analysis, exploration, and the discovery of novel insights. Based on Chiu and Russel's (2010) study, there are significant opportunities for the integration of advanced interactive tools and techniques, along with visual analytic tools in support of a diverse range of construction management functions. Additionally, they found that the construction data visualization environment, when integrated with a construction management (CM) information system, provides a CM analytics information technology infrastructure. Project participants, from experts to novices, are enabled to journey quickly through complex CM data presented in easily understood, natural visual forms.

The increasing volume of free digital data available from multiple sources provides the opportunity to improve many aspects of the environment around us. According to Heer and Shneiderman (2012), “to get the most out of data, users must be able to make sense of it: to pursue questions, uncover patterns of interest, and identify errors” (p. 1). In the same way, the authors recognized that visualization provides a powerful means of making sense of data.

Although visualization systems provide a clear image of data, Heer and Shneiderman (2012) discussed that, to be most effective, visual analytics tools must support the fluent and flexible use of visualizations at rates with the pace of human thoughts. In support of the utilization of data visualization, Mansmann et al., (2012) considered that dynamic visual analytics provides the opportunity to circumvent real-time information overload by combining incremental analysis algorithms and visualizations to facilitate data stream analysis and provide situational awareness. Moreover, the authors stated that real-time analysis is a challenging and relevant field, motivated both by the potential to obtain value from up-to-date information and the threat of damage if timely assessments are unavailable.

Project Control Data Visualization

Computer tools have enhanced the way construction projects produce and communicate information by providing applications capable of creating graphic representations of building objects. According to Lee and Rojas (2013), two essential elements that can help managers make better-informed decisions are 1) timeliness and 2) accuracy of the project information. In the same study, the authors discussed that in typical large-scale construction projects, the large volume of data that project managers must analyze to create a comprehensive picture of the current status of the project results in a time-consuming process in which issues are often overlooked. Song, Pollais, Pena-Mora (2005) discussed that to implement project control effectively, project managers must analyze different sets of information to create an accurate picture of project progress. Additionally, they found that to achieve this goal, project managers need to get access to the information promptly.

Cheng and Teizer (2010) studied the use of effective visualization tools to communicate information from the construction site for fast decision making. The study revealed that the current processes of data delivery do not allow contractors to make appropriate decisions due to an insufficient amount of data and the absence of its visual representation. In Rojas and Lee's (2007) article, the authors stated that project control data visualization had been a limitation to the performance of contractors on the job site. Also, the lack of effort to determine the structure of information hinders the progress of identifying the value of control data to contractors. Song et al., (2005) supported this statement in their research when they revealed that the data contained in reports must be accurate, reliable, and structured in a meaningful way that makes sense to the team members. Additionally, they considered that it is important that the information presented be flexible for the building of a model-based interphase.

According to Cybulski et al., (2013), unstructured data might contain relevant information, but it is hard to analyze because it does not possess a simple graphic representation that allows analysis. Previous studies revealed that there is a need for transforming project data into meaningful information for the purposes of improving construction performance and facilitating contractors' decision making (Rojas & Lee, 2007). Rojas and Lee (2008) considered graphic representations of project control data necessary because they are easy to understand, communicate, and disseminate to other project participants. According to research by Abbaszadegan and Grau (2015), the integration of data is a critical component of the automation process. Also, data automation showed that decision makers with accurate and timely information could produce positive improvements in a project's performance and outcomes. Moreover, Hartmann, Fisher and Keim (2012) found that the alignment of building information modeling (BIM) technology and work processes allow the team to gain an in-depth understanding of project operations and the functionality of BIM-based tools.

Similarly, a study by Russell, Chiu, and Korde (2009) found that data visualization has the potential to enhance the understanding of project status by team members, allowing them to make faster and better-informed decisions. The researchers also recognized that humans are conditioned to interpret data better and more quickly when it is presented in a suitable graphical representation rather than in large textual/numerical tables. The relevance of data visualization was reinforced by Golparvar-Fard, Tang, Cho and Siddiqui (2013), who determined that contractors would be able to evaluate the relationships and impacts of some activities on construction performance. Kim, Park, Lim and Kim (2013) found that to obtain the full potential of data, the information should be distributed efficiently to project participants. In the same research study, they demonstrated that allowing team members to access current project information improved the existing practices of project control and management. In Song et al.'s, (2005) study, the researchers found that with multi-dimensional project control data, team participants were able to diagnose project problems and identify their causes. Lee and Rojas (2013) revealed that project managers are receptive to using a data visualization approach for analyzing project performance. Additionally, they found that project managers were able to explore and quickly discover valuable information that supported their decisions. Cybulski et al., (2013) learned that interactive visual analytics (IVA) could be used to find different relationships among data. This allows users to identify problems, such as hidden data patterns. This leads to the generation of new insights, enabling the team to solve problems. The research team provided examples in which this method proved its effectiveness in areas such as business, engineering, health, education, entertainment, industries, and community service.

4D Scheduling

Construction organizations are increasingly combining three-dimensional models (3D) with construction schedules to produce construction simulations (4D) to visualize project construction in a virtual environment, and companies have realized the benefit of its implementation: a competitive advantage (Wang, 2007). These problems are not easy to visualize in a two-dimensional (2D) schedule, like a Gantt chart, as indicated by Wang (2007) who found project interpretation to be a consistent obstacle in the development of construction schedules. The lack of consistency in the way planners understand and build schedules affects the communication of information to the rest of the team. The researchers also determined that the absence of integration hampers scheduling and planning success because the two activities are usually generated independently.

In a like manner, Mahalingam, Kashyap, and Mahajan (2009) studied the implementation of 4D simulations during the construction stage of the project. The study found that 4D could produce a positive impact in the evaluation of project execution by contrasting assumptions generated during the planning stage with site productivity. In support of this statement, Dang and Bargstädt (2015) commented that the ability to visualize construction processes offers project stakeholders a tool to improve communication, detect time-space conflict, understand a process, and control execution. On the contrary, the authors found challenges in the implementation process and the adoption of 4D. Mahalingam et al., (2009) concluded that the efforts invested by professionals to develop the simulation and update the activities to produce the information needed for analysis might not represent a substantial gain to maintain and repeat the process.

However, Karshenas and Sharma (2010) determined that any advancement in the communication of the schedule could provide meaningful improvements in the effectiveness of project management, construction coordination, and control. The authors proposed the use of building information models (BIM) to extract the activities and build the schedule by selecting 3D elements of the project in a virtual environment based on the assembly composition transfer of information to a scheduling software, which would result in a reduction of the number of inconsistencies and improve planning. Hartmann et al., (2012) argued that current "technology push" implementation to existing work processes should change to align to the functionalities of the BIM-based tools. Additionally, the researchers stated that BIM-based tools require an integrated work environment, and to accomplish this goal, the authors suggested that team members should align their processes to a collaborative way of work. Abbaszadegan and Grau (2015) found in their investigation that automated data analytics systems could enable project managers to control project information and act in a proactive manner so they can make decisions faster.

A related study by Goedert and Meadati (2008) emphasized the need for data integration in construction to facilitate the improvement of building management throughout its life cycle. A significant remark was made in the urgency to extend the use of 4D simulations and collect data to update the plan and make comparisons with actual information. According to Goedert and Meadati's (2008) information collection and quantification are consistent limitations that hinder the progress of the integration of data in the construction management process. Furthermore, Behzadan, Menassa, and Pradhan (2015) found that the current practices of construction management used 4D simulations actively during the planning process but not during project execution due to the complexity of creating, integrating, and maintaining new data. As a result, there are critical limitations to the extended use of construction simulations. According to Hartmann et al., (2012), the use of traditional tabular support tools in a complex manner for large construction projects integrates multiple disciplines but does not allow project participants to visualize and understand quickly the risk implications of the project. Therefore, the complexity of the information's format, limited collaboration, project assessment and risk mitigation limit teams' ability to make decisions on time. Equally important, Golparvar-Fard et al., (2013) mentioned that data reliability represents a critical factor for success in data collection and that current practices showed that project participants spend a significant amount of time documenting the project status and analyzing data.

In Abbaszadegan and Grau’s (2015) study, a gap was identified in a project report cycle, which was evidence that decisions were made with outdated information. On the other hand, Goedert and Meadati (2008) found that the data collected and transferred during a project's construction has the potential to be used for further analysis after the project's completion. In Cheng and Teizer’s (2010) article, the authors observed that real-time data collection becomes a critical activity to inform contractors of the current state of the construction, which allowed them to improve the quality of their decisions. In the same way, the results of Behzadan et al., (2015) revealed that the data collected from the job site contains valuable data for keeping the simulation active and which produced useful information for decision making.

Research Questions

1. Why do we need data visualization systems for construction monitoring?
2. What are the current limitations to implement data visualization during project control and monitoring?
3. Does the project team find project data visualization information valuable?

Chapter 3: Methodology

This chapter provides the multiple methods utilized for data collection at each step of the study and the procedure employed for data collection and its analysis. Each section contains detailed information about the steps required for the execution of this research. Two mechanisms were designed to collect the information. The first method included interviews with construction professionals in the initial and final stages of the research. To gather the information from the professionals, two questionnaires were created to collect two data sets aiming to understand human processes and interaction with data processing and to learn how construction sites generate data. The second method involved the design of an application to keep the records of activity progress in an electronic format. Additionally, a diagram to execute the study was required. The diagram created was based on the findings of the interview session in the initial stage of the research.

Participants

The participants in this study were three professional constructors who held different positions in a construction company. The ages of the participants were between 21-45, and their levels of experience in the control and monitoring construction schedules varied from an intermediate to a high level of expertise with over ten years of experience. The participants had different roles and responsibilities in the control and monitoring of the construction process, including project management, superintendent, and project engineer. Professionals were recruited based on the level of decision-making authority for the project. The reason for this selection process was to gather information from the people who analyzed the current state of the project. The comments provided by these professionals helped to understand their methodologies in collecting and processing the information.

Context

The project selected for this research study was the construction of a new five-story building with one basement level. The primary superstructure support of the building was concrete with a metal structure for the roof. Additionally, the exterior veneer is brick and limestone. The construction area of the project was 142,000 square foot, and the budget was $100 million. The location of the project is on a university campus in the south central region of the United States. High traffic of pedestrians around the project limited accessibility of construction equipment and materials. The building foundations were 100% completed, and the structure was 60% completed at the start of this study.

Procedures

The development of this study required two types of data; qualitative and quantitative. The two data types allowed for analysis of different aspects of the project. The aspects analyzed were 1) current practices and processes performed by the general contractor (GC) and 2) the percent of completion of the construction activities during project execution. The two methods employed to collect data were 1) interview to collect qualitative data and 2) documentation of project progress to collect quantitative data. Human input was critical before proposing any solution. The investigation of current practices and the dynamics of the construction site were important.

The data collection of this study consisted of three segments: 1) interviews with construction professionals in the early stages of the research, 2) daily observations of

GC's operations updating construction activities, and 3) interviews based on the participants' experiences operating a data visualization application. It was important to follow the sequence of data collection procedures because it determined the results produced by the data visualization application, so the participants were able to prove its usability. Based on the information obtained from the participants in this study, it was possible to identify the limitations of construction monitoring. The method employed consisted in creating plans for each step of the process. These plans worked as a road map to visualize the procedure flow and data flow. The Research Guide Diagram in Figure 1 is a visual representation of the study's road map.

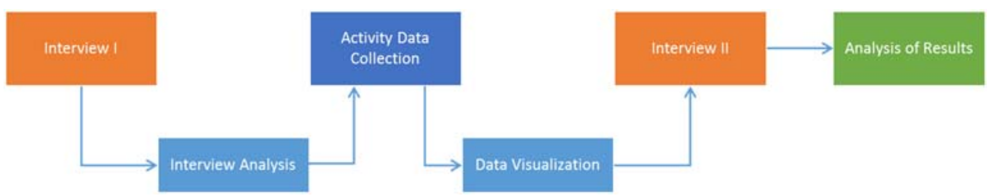


Figure 1. Research Guide Diagram

First Interview Session

After completing the feasibility phase, the next step was to collect data from construction professionals to understand current practices in the control and monitoring processes. The method selected to gather this information consisted in interviews with key construction professionals.

Thus, an interview session was conducted in an office in the GC's job trailer on the work site of the building project. Before the session started, a brief description of the study was provided to the participants so they could obtain a general knowledge of the context of this research. The information allowed them to provide accurate information regarding the particular subject studied. After the introduction to the study, the audio recorder of a smartphone was set between the investigator and the participant and turned on during the session. Then questions from the questionnaire were read from the computer screen to the participants. Due to the nature of the interview system selected and the structure of the questions, the subjects had all the time they needed to provide their answers. Conversely, it was programmed to take around 30-45 minutes with each participant. To preserve the objectivity, each participant in the study had the same set of questions and procedures. The objective was to obtain accurate information, given the strong knowledge and experience these professionals possessed of the construction field.

Second Interview Session

For this phase of the study, meetings with each participant were performed to document their observations using the data visualization application to visualize data obtained from the WWP sessions, the weather conditions, and the crew management databases. It was critical to document their experiences because it was possible to identify weaknesses in data communication. Additionally, during this phase, application's responsiveness was tested to monitor the interaction between the user and the system. Application's responsiveness was tested by allowing the participants to explore and manipulate the application so that buttons and data plots were tested to verify their performance, producing the information requested. These methods evaluated if the user was able to query the information needed and obtain the clear graphical information that would allow them to interpret and analyze the multiple data points collected during the data collection phase.

Access to the application was provided for participants to use the computer where the application was created so they could navigate the information using the different buttons available within the "Activity Data Visualization" screen. These sessions were performed individually so personal assessments of the operation of the system could be executed additionally to document if the application addressed the participant's data query. The duration of the sessions was 30 minutes. During this time, the user had the opportunity to navigate the information and ask questions regarding the system's performance. Each session started with a brief introduction of the system, a quick overview of its capabilities, and a short training period to learn how to operate the application.

Data Processing

This procedure consisted of data transferring and processing of the formation to a database. The implementation began by exporting the data from the scheduling software, which contained the information of the activities observed. The data assigned to each of the activities included planned start, planned finish, estimated duration, percent complete, and status. The data was then transformed using the "Activity Builder" tool, which was created to organize the activity names of the schedule in a predetermined syntax that allowed data organization and classification. Then, the information was transferred to the data visualization application, where the schedule data was transformed from "Raw Data" to structured data so it could be loaded into the activity database. Figure 2 represents the flow of this procedure.

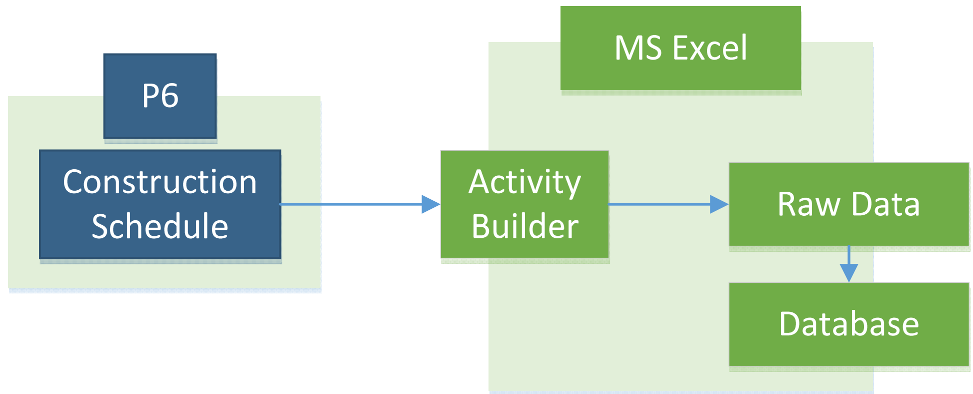


Figure 2. Schedule Data Import

In the same way, three data sets were recorded daily, and they were 1) activity data update, 2) subcontractor crew report, and 3) climate conditions. Each data set consisted of specific pieces of information stored in the application's database. As shown in Figure 3, the updated activity data set was collected during the GC's weekly work plan (WWP) daily sessions. The information documented during the sessions was collected using a data collection interface to update the activities' statuses and to load the information onto the database.

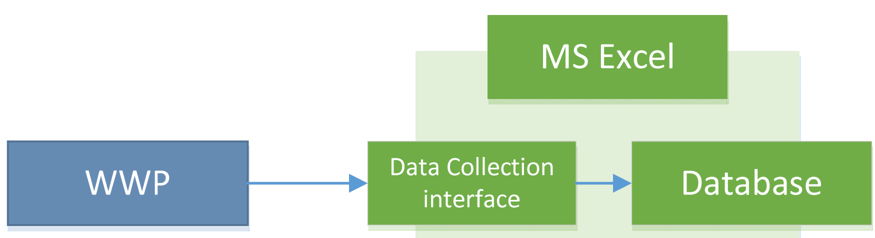


Figure 3. Activity Data Update

Similarly, data from the crew report was obtained by imputing the information manually into the database using a data collection interface created to facilitate data entry. The information from the report was saved and stored in the database. As shown in Figure 4 below.

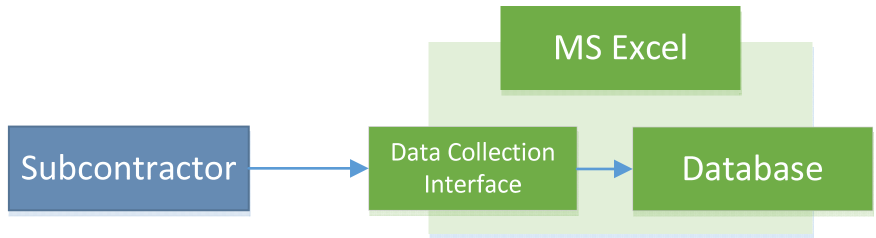


Figure 4. Crew Report Data Flow

Also, information documenting the climate conditions of the geographic location of the construction project was recorded and saved in the database using a data collection interface. Figure 5 displays the procedure executed for documentation of climate conditions.

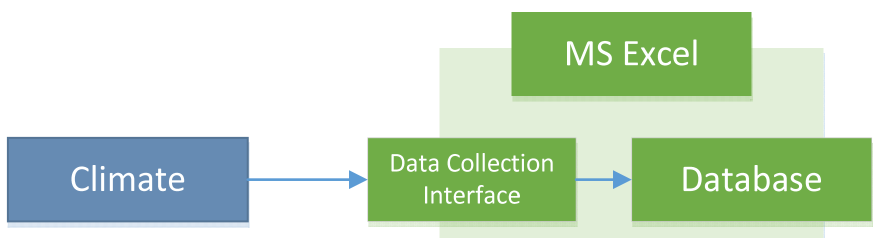


Figure 5. Climate Conditions Data Flow

In the initial phase of this study, just a few data sets were created, such as activity progress, planned start, planned finish, and climate conditions (temperature). After a series of meetings with the participants, the number of sets increased to adapt to their needs. These new datasets were, week or month progress chart, activity complexity, climate conditions (humidity, precipitation), subcontractor mark and man power. After creating the datasets, the visual representation of each set was organized in one place to create a dashboard, so the participants were able to visualize all the information in one single screen. In the final step, the participants accessed the information and performed their analysis based on the information displayed on the screen. Figure 6 displays this procedure.

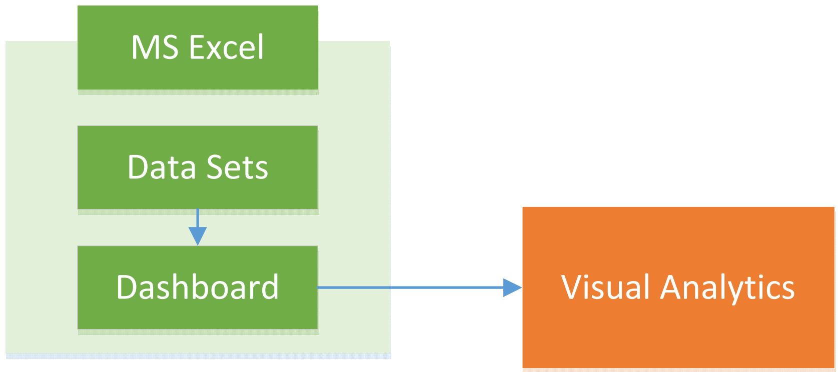


Figure 6. Data Visualization and Analytics Diagram

Visual Analytics Process

The next procedure for the implementation of visual analytics was to provide multiple datasets that represented the values of the information from the activities observed. After generating the dataset that produce data visual representations the GC were provided with the data visualization application to operate the data for real hands-on experience, which allowed the documentation of their comments as they used the application. After the documentation of their comments, a thorough analysis was done to learn how this information can influence the interpretation of the current status of the project based on the project's activity history. Based on the results of the analysis, the information in the application was modified to address the requirements of the participant. Figure 7 represents the entire process as described above.

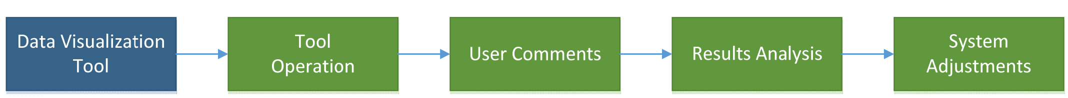


Figure 7. Visual Analytics Implementation Process

Research Design

In the initial stages of the research design process, a critical step needed to be tested before creating the structure of the study to proof the feasibility of continuing with the research. This preparatory step was the examination of the information exported from the scheduling software, which in this case was Oracle Primavera (P6). The execution of this process allowed for the analysis of the data syntax of the schedule content and its integration with the software selected to execute the data visualization plots, Microsoft Excel (MS Excel). The results of the analysis showed that there were multiple inconsistencies in the activity name syntax, which prevented the utilization of the information. Therefore, a computer software application capable of organizing the text automatically had to be designed for data transferring and transformation.

Once the preliminary step was completed and tested, it was possible to start the development of the research. The research design consisted of five fundamental steps, each of them containing a series of necessary actions to move forward in the process. The first step was "map content," the second "create capture data structure," the third "collect data," the fourth "test data," and the fifth step was "prepare research analysis". These steps are shown un Figure 10.

Feasibility Test

An issue found during data formatting prevented the organization and classification from the information imported from P6; the problem consisted in data inconsistencies across the schedule activity names. Two construction schedules were reviewed to verify if there was a pattern in the formation of the plan. For example, some of the disparities were how the activities were named, the delimiters used to organize and identify the activity, and typing mistakes (capitalization, non-capitalization, and spelling errors). Therefore, a naming protocol was proposed to test syntax structure and the usability of activity names to organize and store the information.

Due to the limited access to P6, it was not possible to test the syntax structure proposed for import-export to verify its integration across systems. Although in substitution, a different software was used to perform testing, the software utilized was Microsoft Project (MS Project) which is a scheduling software with similar capabilities as P6.

The syntax testing process was initiated by transforming the activity names from the schedule data imported from P6 to the proposed naming protocol, using the "Activity Builder" tool. The tool is only available in MS Excel; thus the activities were created on MS Excel then transferred to MS Project, and the information was finally sent to the data visualization application also created in MS Excel.

The data transfer from one system to another was successful. The text structure remained as established, allowing the visualization application to recognize the pattern and organize and classify the information into multiple categories, as well as store the data in the database. A diagram of the data integration process can be found in figure 8.

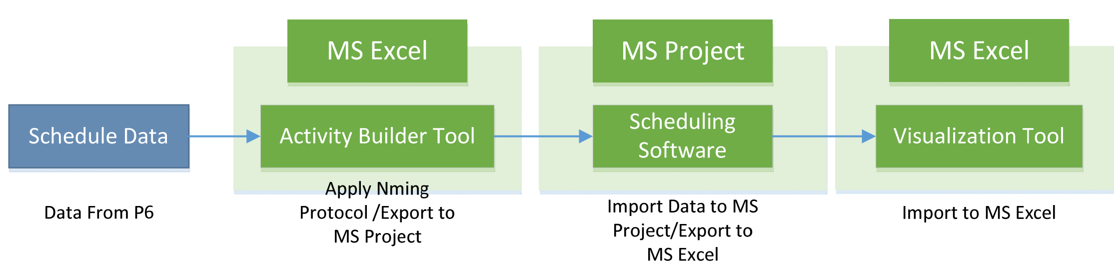


Figure 8. Data Integration Testing

Step 1: Map Content

In this step on the research design, a complete analysis of the expected outcome was planned, thus making it possible to understand what pieces of information were required to complete the research. Based on the results of the investigation, the methodology for obtaining the first part of the information was determined. Thus, two questionnaires with open-ended questions were created to document the general contractor's (GC) current practices for project control and the limitations of data collection and analysis.

Select Participants

According to the study in the previous step and the questions formulated, a group of three professionals who are essential in the decision-making process in the construction project were selected.

Perform Interviews

Once the participants agreed to be part of the study, interview sessions were performed to understand how each member collects, processes, and analyzes construction data. This stage was influential in the development of the study because the participants provided valuable insights that allowed the identification of gaps in the data processing.

Step 2: Create Data Capture Structure

Due to the lack of computer application available to collect construction activity progress data, an application was created to satisfy the need for data management and processing. Based on the analysis previously performed for the expected outcome, the application designed addressed many of the characteristics necessary for this investigation.

Validate Structure

This was the first step to developing the data visualization application. The method employed was creating a diagram to plan the data flow from its collection until its visualizations. To begging the process for collecting construction data, it was required to ask, “what data is needed to recording so the application can provide enough visual information to the participants?”. This question allowed the formation of a data pipeline diagram that shows the three stages of the information collected. The first stage was “data collection”, in which the different actions performed in the construction site were recorded. The second stage was “process data”, in this stage the data had to be organized and classified for an effective use of the information. Therefore, multiple tables helped to distribute the data within the application. And the last stage of the information collected was “event visualization”. Within this last stage, all the data points collected are visualized through the implementation of an application that provided the participants with visual data representation. These visualizations provided by the application intend to display the events on the construction site. The diagram used to represent the data pipeline is shown in Figure 9.

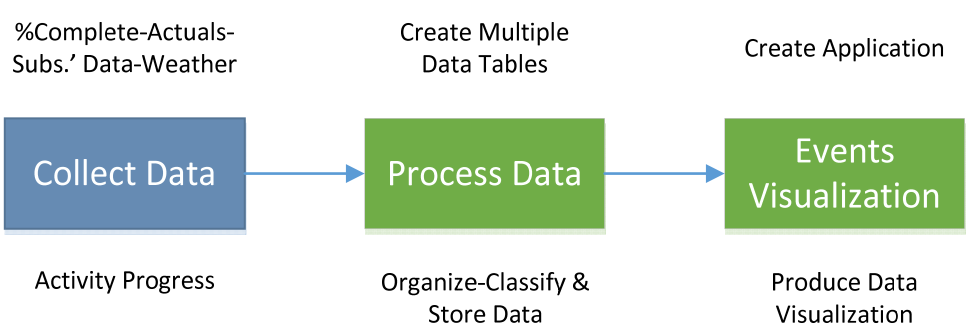


Figure 9. Data Pipeline Diagram

MS Excel VBA

As previously mentioned in this section, the software selected for data transformation and data visualization was MS Excel. Within this program, there is a developer window that allows users to create applications using visual basic for applications (VBA).

This step consisted of writing the computer codes that would allow the application to perform multiple actions automatically, by creating different interfaces that allowed the automation of data entry and classification into the program (MS Excel).

Test

After completing the plan and coding, a testing process was executed to review the output information. This step also involved the identification of code bugs and investigation of the failure to produce the expected results.

Reject/Approve

In this stage, the work took two routes, depending on the results generated in the previous testing phase. Route one was the rejected option; if the outcome was not successful, then the work was transferred to the "validate structure" stage for corrections. Route two was the approved option; if the outcome produced the expected results, then the work can continue to the next step (collect data) for data production.

Step 3: Collect Data

This step consisted of daily work site visits for recording activity progress data by implementing the application designed.

MS Excel

Using the application created in MS Excel, the information collected daily by the GC from the subcontractors was transferred to electronic format. The application allowed the organization, classification, and storage of the data collected.

Data Plots

Within the same application created for data collection, multiple data sets were adjusted to establish data relationship and generate visual data representation. These adjustments facilitated the use of plots for real-time data updates.

Step 4: Test Data

In this step, the information produced based on the data collected from the job site visits provided proof of usability by the construction professionals who participated in the study. Their insights were significant in this phase of the investigation, as issues, concerns, and deficiencies were identified and revised.

Present Results

Demonstration sessions with the participants of this study allowed the display of the information produced by the application created. Additionally, the participants were able to manipulate the information so they could have the experience of navigating the information, which contributed for improved feedback.

Data Usability

For this step, the participants evaluated the information produced by the application. The goal was to gather their observations and add value to the information presented. These comments helped tremendously to confirm the usability of the information based on the assumptions that initiated this investigation.

Step 5: Prepare Research Analysis

In this step, all the information gathered from step one through step four was analyzed and summarized to document the findings of the study and its conclusions. This last step helped to understand the challenges associated for the implementation of data visualization in the construction industry at the same time the benefits of its utilization.

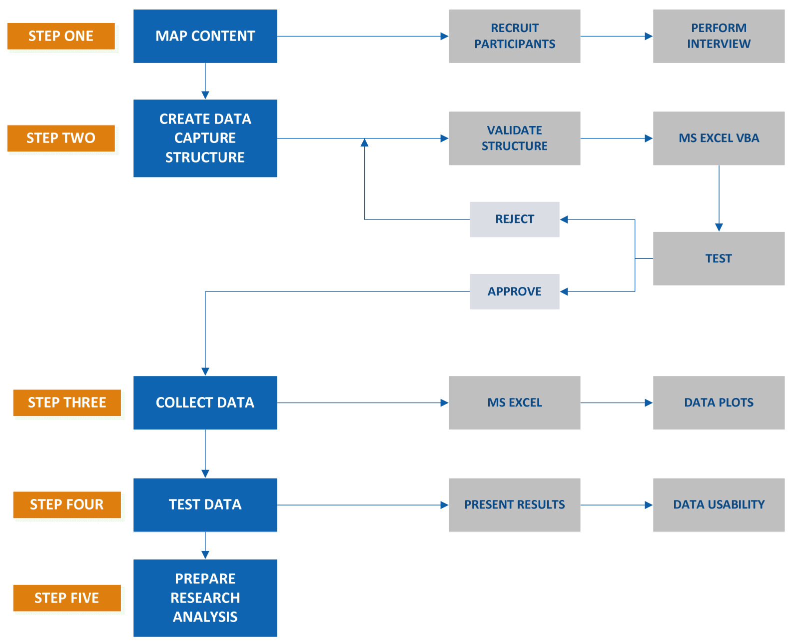


Figure 10. Research Design Diagram

Instruments

Several aspects around the construction work site were not clear to the researcher. Therefore, a questionnaire was formulated to collect information and clarify the methods currently being utilized for project control and monitoring in this particular project. The questionnaire consisted of two separate sections; the first included eleven questions that were asked directly to the participant before data collection started, and the second set was used after gathering and processing the data from the job site, creating a visual representation of the information. The interview format selected was a semi-structured interview due to its conversational nature, which allowed the participants to feel comfortable talking about their roles and experience in the field. Additionally, this type of interview provided an opportunity for the participants to express their thoughts on whether the schedules provided them with valuable information for the execution of the construction project.

The first questionnaire allowed a deeper understanding of how construction schedules were developed and maintained throughout the project construction. Questions were created in such a way that the participants were able to provide information about project control and analysis from their perspective and responsibility on the project. An open-ended structure for the questions was chosen so that the participants would have the opportunity to explain their responses in depth. For the follow-up questions in the second part, the participants were provided with the results of the information collected during the activity progress for data navigation. An audio recorder was used during the interview sessions to record the participants' answers. The recorder allowed for storage of their answers in a digital format for further analysis.

Questionnaires

The questions administered to the participants in the first interview session asked them to provide details about how they perform their activities for analyzing the current state of the construction progress. Both questionnaires can be found in Appendices A and B.

One question during the first interview was "Describe the process you use to gather information about the current status of the project each time you visit the job site." Responses to this question provided valuable information that described the diversity of methodologies implemented by each participant. Additionally, this information helped in the identification of gaps in the process of collecting and storing data. Another question that provided significant information was, "What are the challenges and your concerns when documenting project activities' statuses?" The answers varied with each participant, but they aligned in the context that construction projects are collaborative efforts where information and communication are paramount.

Detailed information shared by the construction professionals during the first interview provided a clear view of the different processes that each of them needed to execute when analyzing project progress. This step was critical for the development of this study.

For the second round of interviews, five questions were asked as shown in Appendix B. Administration of these questions was in the final stages of the investigation. These follow-up questions played a key role in the study and in verifying the implementation of data visualization as impacting the project control and monitoring during construction. It is necessary to mention that to perform this session, the participants needed to have previous experience evaluating data visualization content for accurate responses. For this reason, access to the application with activity data and visual information gave the participants a correct view of the results of the data collected.

The information needed from the participants was the identification of the value of this information for their daily activities and the limitation of data visualization. Additionally, the participants were expected to report their concerns about the difficulties in interpreting the information and data query. By questioning the capabilities of data visualization to provide the information required to perform an assessment of the current state of the project, it was possible to expose the limitations of this data visualization application.

Data Visualization Application

To achieve an accurate and up-to-date construction schedule, it was essential to have a permanent collection of information about the activity status from the job site. In the same way, there are multiple factors that influence the performance of construction trades on the job site, such as weather conditions and crew sizes. Current construction software are not capable to document many of the variables that influence construction job sites. The reason why this information is missing from the current scheduling software is that this kind of data is not supported by the system at this moment.

Consequently, an application was designed to satisfy the need for a tool to collect construction data. The purpose of the application as a result of this study was to collect, update, organize, store and visualize the information associated with the activities under execution in a construction project as other factors that may influence project performance. The application acts as a data repository that classifies, organizes, and stores multiple data sets. The application consists of five screens in total, four screens for data management and one screen for data visualization. Screen one is for data import, screen two is for schedule data and activity progress data, screen three is for climate conditions, and screen four is for crew management (crew report). Screen five displays multiple plots with a graphic representation of the information, which is called "Activity Data Visualization." In this screen, data can be filtered and visualized, so the user can navigate and customize the charts for those that they consider applicable for their assessment.

Activity Progress Data Collection

For one month, data documenting activity progress was collected through observations of daily meetings that the GC and key subcontractors conducted on the job site. Then, the information was exported and restructured from P6 and transferred into MS Excel. Subsequently, the data was arranged in multiple ways to try to improve its capability to be analyzed. By breaking the string of information into various parts, the data was classified in multiple categories. The application of this method enabled the opportunity to transfer each piece of information to different columns, which made it easier to manipulate and maintain. The intention of applying this approach was to increase the value of each component of the series data so graphic charts may provide proper information according to each of the categories. Thanks to the integration between the software, the process of transferring the data from one to another allowed the data to maintain its structure.

In this particular case, the GC implemented lean construction principles for their monitoring operations. Their current practices included the use of the Weekly Work Plan (WWP) approach, which is part of the Last Planner System®, which at the same time is one of the processes that the Lean Construction Institute (LCI) implements to plan and control project progress. Therefore, special permission was granted to observe the daily sessions held at the construction site to update the information programmed on the WWP board. The tasks documented in this study were those associated with the execution of the concrete structure of the project, and included columns, decking, concrete, and rebar.

The procedure implemented for data collection in this phase of the research was through observations of the WWP daily sessions. The sessions were held every morning, five days a week, in the construction trailer on the work site and had a duration of approximately 30 minutes, except on Fridays when an hour was required to summarize the production of the week. The method employed for data entering was typing the information posted on the board employing the application designed for data entry into the computer system created for data management and storage.

After the data was stored in the application, it can be selected and edited with current information obtained from the construction site. Each morning during the sessions, the application was activated to update the statuses of activities. There were days that it was not possible to observe the meeting due to commitments not related to the study. In these cases, a digital picture of the board was taken after the meeting so the information could be updated and transferred into the system.

The data collected to update the statuses of the project activities consisted of actual start, actual finish, percent complete, status, subcontractor, and activity delay cause. The reason for recording these elements was that the data posted on the board followed the lean construction principles applied by the general contractor. To increase the accuracy of data entry for activity update, each of these elements were recorded using an application designed for data update, which reduced the possibility of inconsistencies. This procedure was repeated to document the statuses of the following activities: columns, reinforcing, concrete, and decking.

Screen 1: Data Import Screen

In this screen, the activity and the data associated with it from P6 were located and imported into the application. After importing the content from the schedule, the data contained in the “Activity Name” column was restructured. The use of delimiters to structure the information was important because the application was programmed to separate the data into individual segments for organization and classification.

Then, the “restructured” information was transferred to the “Activity Data” screen. The visualization application was programmed to select particular columns and then transfer the information into the database. The columns used in this research were the activity name and the information assigned by the "activity builder” tool (operation, activity name, category, area, floor, and division Nº), planned start, planned finished, and estimated duration. The rest of the columns from P6 were deleted. The contents of the P6 columns deleted were activity ID, actual start, actual finish, and percent complete. The reason why part of the information was deleted was that the data visualization application was capable of updating the statuses of the activities and maintaining the record of its progress internally. A figure with the data imported from the scheduling software is presented in Figure 11, additionally an enlarged figure can also be found in Appendix C.

LEV3 AREA C						18	1/19/2016	2/11/2016	100	18	42388	2/11/2016	
Installing	Deck	Structure	Area C	Level 03	3	5	1/19/2016	1/25/2016	100	5	42388	1/25/2016	Finished
Installing	Reinforcing	Structure	Area C	Level 03	3	5	1/26/2016	2/1/2016	100	5	42395.33333	2/1/2016	Finished
Installing	Rough MEP	Structure	Area C	Level 03	3	3	1/28/2016	2/1/2016	100	3	42397.33333	2/1/2016	Finished
Pouring	Concrete	Structure	Area C	Level 03	3	1	2/2/2016	2/2/2016	100	1	42402.33333	2/2/2016	Finished
Installing	Columns	Structure	Area C	Level 03	3	4	2/8/2016	2/11/2016	100	4	42408.33333	2/11/2016	Finished
LEV3 AREA B						30	12/18/2015	2/1/2016	100	30	42356	2/1/2016	
Installing	Deck	Structure	Area B	Level 03	3	18	12/18/2015	1/15/2016	100	18	42356	1/15/2016	Finished
Installing	Reinforcing	Structure	Area B	Level 03	3	19	12/21/2015	1/19/2016	100	19	42359	1/19/2016	Finished
Installing	Rough MEP	Structure	Area B	Level 03	3	19	12/21/2015	1/19/2016	100	19	42359	1/19/2016	Finished
Pouring	Concrete	Structure	Area B	Level 03	3	1	1/20/2016	1/20/2016	100	1	42389	1/20/2016	Finished
Installing	Columns	Structure	Area B	Level 03	3	4	1/27/2016	2/1/2016	100	4	42396.33333	2/1/2016	Finished
LEV3 AREA A						20	1/11/2016	2/5/2016	100	20	42380.33333	2/5/2016	
Installing	Deck	Structure	Area A	Level 03	3	6	1/11/2016	1/18/2016	100	6	42380.33333	1/18/2016	Finished
Installing	Reinforcing	Structure	Area A	Level 03	3	4	1/20/2016	1/25/2016	100	4	42389	1/25/2016	Finished
Installing	Rough MEP	Structure	Area A	Level 03	3	6	1/18/2016	1/25/2016	100	6	42387	1/25/2016	Finished
Pouring	Concrete	Structure	Area A	Level 03	3	1	1/26/2016	1/26/2016	100	1	42395.33333	1/26/2016	Finished
Installing	Columns	Structure	Area A	Level 03	3	4	2/2/2016	2/5/2016	100	4	42402.33333	2/5/2016	Finished
LEVEL 4						30	2/2/2016	3/14/2016	30	9	42402.33333	3/14/2016	
LEV4 AREA A						17	2/10/2016	3/3/2016	17.65	3	42410.33333	3/3/2016	
Installing	Deck	Structure	Area A	Level 04	3	2	2/10/2016	2/12/2016	100	2	42410.33333	2/12/2016	Finished
Installing	Reinforcing	Structure	Area A	Level 04	3	9	2/15/2016	2/25/2016	0	0	42415.33333	2/15/2016	Started
Installing	Rough MEP	Structure	Area A	Level 04	3	8	2/16/2016	2/25/2016	0	0	42416.33333	2/25/2016	Planned
Pouring	Concrete	Structure	Area A	Level 04	3	1	2/26/2016	2/26/2016	0	0	42426.33333	2/26/2016	Planned
Installing	Columns	Structure	Area A	Level 04	3	5	2/26/2016	3/3/2016	0	0	42426.33333	3/3/2016	Planned
LEV4 AREA C						17	2/15/2016	3/8/2016	0	0	42415.33333	3/8/2016	
Installing	Deck	Structure	Area C	Level 04	3	6	2/15/2016	2/22/2016	0	0	42415.33333	2/22/2016	Planned
Installing	Reinforcing	Structure	Area C	Level 04	3	6	2/22/2016	2/29/2016	0	0	42422.33333	2/29/2016	Planned
Installing	Rough MEP	Structure	Area C	Level 04	3	5	2/23/2016	2/29/2016	0	0	42423.33333	2/29/2016	Planned
Pouring	Concrete	Structure	Area C	Level 04	3	1	3/1/2016	3/1/2016	0	0	42430.33333	3/1/2016	Planned
Installing	Columns	Structure	Area C	Level 04	3	4	3/3/2016	3/8/2016	0	0	42432.33333	3/8/2016	Planned
LEV4 AREA D						16	2/22/2016	3/14/2016	0	0	42422.33333	3/14/2016	
Installing	Deck	Structure	Area D	Level 04	3	6	2/22/2016	2/29/2016	0	0	42422.33333	2/29/2016	Planned
Installing	Reinforcing	Structure	Area D	Level 04	3	5	3/1/2016	3/7/2016	0	0	42430.33333	3/7/2016	Planned
Installing	Rough MEP	Structure	Area D	Level 04	3	4	3/2/2016	3/7/2016	0	0	42431.33333	3/7/2016	Planned
Pouring	Concrete	Structure	Area D	Level 04	3	1	3/8/2016	3/8/2016	0	0	42437.33333	3/8/2016	Planned
Installing	Columns	Structure	Area D	Level 04	3	4	3/9/2016	3/14/2016	0	0	42438.33333	3/14/2016	Planned
LEV4 AREA B						20	2/2/2016	2/29/2016	45	9	42402.33333	2/29/2016	
Installing	Deck	Structure	Area B	Level 04	3	6	2/2/2016	2/9/2016	100	6	42402.33333	2/9/2016	Finished

Figure 11. Data Import Screen

Screen 2: Activity Data Screen

The primary information comes from the construction schedule, created in this case in P6. The information was transferred from P6 to MS Excel and was transformed from raw data to clean data. The operation to process the data was automated, so little intervention by the user was required. The transformation consisted on the separation of the information into the different components previously established in the activity's name.

The activity data contained on this screen was organized in the respective columns, and the information on the current status of the activity was updated daily. To update the information about the activity, multiple columns were assigned to document the progress, including percentage complete, actual start, actual finish, actual duration, status and record date. The record date represents the date on which the activity was updated; this data was recorded using a data collection interface (climate condition and crew management) on each screen. The crew data collected is shown in Figure 12. An enlarged figure can also be found in Appendix D.

Action	Activity	Category	Work Zone	Floor	Division N°	Estimated Duration	Planned Start	Planned Finish	% Complete	Actual Duration	Actual Start	Actual Finish	Status	Recod Date
Inst.	Columns	Structure	Area D	Level 4	3	5	3/10/2016	3/16/2016	50%	0	3/10/2016		Started	3/14/2016
Drop	DeckDrop	Structure	Area C	Level 4	3	5	3/12/2016	3/17/2016	75%	0	3/12/2016		Started	3/14/2016
Inst.	Decking	Structure	Area A	Level 5	3	4	3/9/2016	3/15/2016	80%	4	3/9/2016		Started	3/14/2016
Inst.	Decking	Structure	Area C	Level 5	3	5	3/16/2016	3/22/2016	0%	0	3/14/2016		Started	3/14/2016
Inst.	Reinforcing	Structure	Area B	Level 5	3	4	3/8/2016	3/15/2016	80%	0	3/8/2016		Started	3/14/2016
Inst.	Columns	Structure	Area D	Level 4	3	5	3/10/2016	3/16/2016	75%	0	3/10/2016		Started	3/15/2016
Inst.	Decking	Structure	Area A	Level 5	3	4	3/9/2016	3/15/2016	100%	4	3/9/2016	3/15/2016	Finished	3/15/2016
Inst.	Decking	Structure	Area C	Level 5	3	5	3/16/2016	3/22/2016	20%	0	3/14/2016		Started	3/15/2016
Inst.	Reinforcing	Structure	Area B	Level 5	3	4	3/8/2016	3/15/2016	90%	0	3/8/2016	3/15/2016	Started	3/15/2016
Inst.	Reinforcing	Structure	Area A	Level 5	3	4	3/16/2016	3/21/2016	0%	0	3/15/2016		Started	3/15/2016
Inst.	Columns	Structure	Area D	Level 4	3	5	3/10/2016	3/16/2016	100%	5	3/10/2016	3/16/2016	Finished	3/16/2016
Pouring	Concrete	Structure	Area B	Level 5	3	1	3/16/2016	3/16/2016	100%	1	3/16/2016	3/16/2016	Finished	3/16/2016
Drop	DeckDrop	Structure	Area C	Level 4	3	5	3/12/2016	3/17/2016	100%	4	3/12/2016	3/16/2016	Finished	3/16/2016
Drop	DeckDrop	Structure	Area D	Level 4	3	5	3/16/2016	3/22/2016	0%	0	3/16/2016		Started	3/16/2016
Inst.	Decking	Structure	Area C	Level 5	3	5	3/16/2016	3/22/2016	40%	0	3/14/2016		Started	3/16/2016
Inst.	Reinforcing	Structure	Area A	Level 5	3	4	3/16/2016	3/21/2016	20%	0	3/15/2016		Started	3/16/2016
Inst.	Reinforcing	Structure	Area B	Level 5	3	4	3/8/2016	3/15/2016	100%	5	3/8/2016	3/16/2016	Finished	3/16/2016
Drop	DeckDrop	Structure	Area D	Level 4	3	5	3/16/2016	3/22/2016	25%	0	3/16/2016		Started	3/17/2016
Inst.	Decking	Structure	Area C	Level 5	3	5	3/16/2016	3/22/2016	60%	0	3/14/2016		Started	3/17/2016
Inst.	Reinforcing	Structure	Area A	Level 5	3	4	3/16/2016	3/21/2016	40%	0	3/15/2016		Started	3/17/2016
Inst.	Columns	Structure	Area B	Level 5	3	4	3/16/2016	3/21/2016	0%	0	3/18/2016		Started	3/18/2016
Drop	DeckDrop	Structure	Area D	Level 4	3	5	3/16/2016	3/22/2016	50%	0	3/16/2016		Started	3/18/2016
Inst.	Decking	Structure	Area C	Level 5	3	5	3/16/2016	3/22/2016	80%	0	3/14/2016		Started	3/18/2016
Inst.	Reinforcing	Structure	Area A	Level 5	3	4	3/16/2016	3/21/2016	80%	0	3/15/2016		Started	3/18/2016
Inst.	Columns	Structure	Area B	Level 5	3	4	3/16/2016	3/21/2016	30%	0	3/18/2016		Started	3/21/2016
Drop	Deckdrop	Structure	Area D	Level 4	3	5	3/16/2016	3/22/2016	75%	0	3/16/2016		Started	3/21/2016
Inst.	Decking	Structure	Area D	Level 5	3	5	3/23/2016	3/29/2016	0%	0	3/21/2016		Started	3/21/2016
Inst.	Reinforcing	Structure	Area A	Level 5	3	4	3/16/2016	3/21/2016	100%	5	3/15/2016	3/21/2016	Finished	3/21/2016
Inst.	Reinforcing	Structure	Area C	Level 5	3	4	3/23/2016	3/28/2016	0%	0	3/21/2016		Started	3/21/2016
Inst.	columns	Structure	Area B	Level 5	3	4	3/16/2016	3/21/2016	50%	0	3/18/2016		Started	3/22/2016
Pouring	Concrete	Structure	Area A	Level 5	3	1	3/22/2016	3/22/2016	100%	1	3/22/2016	3/22/2016	Finished	3/22/2016
Drop	DeckDrop	Structure	Area D	Level 4	3	5	3/16/2016	3/22/2016	100%	6	3/16/2016	3/22/2016	Finished	3/22/2016
Inst.	Decking	Structure	Area C	Level 5	3	5	3/16/2016	3/22/2016	100%	8	3/14/2016	3/22/2016	Started	3/22/2016
Inst.	Decking	Structure	Area D	Level 5	3	5	3/23/2016	3/29/2016	20%	0	3/21/2016		Started	3/22/2016
Inst.	Reinforcing	Structure	Area C	Level 5	3	4	3/23/2016	3/28/2016	20%	0	3/21/2016		Started	3/22/2016
Inst.	Columns	Structure	Area B	Level 5	3	4	3/16/2016	3/21/2016	75%	0	3/18/2016		Started	3/23/2016
Drop	DeckDrop	Structure	Area B	Level 5	3	5	3/23/2016	3/29/2016	0%	0	3/23/2016		Started	3/23/2016
Inst.	Decking	Structure	Area D	Level 5	3	5	3/23/2016	3/29/2016	40%	0	3/21/2016		Started	3/23/2016
Inst.	Reinforcing	Structure	Area C	Level 5	3	4	3/23/2016	3/28/2016	40%	0	3/21/2016		Started	3/23/2016
Inst.	Columns	Structure	Area B	Level 5	3	4	3/16/2016	3/21/2016	100%	5	3/18/2016	3/24/2016	Finished	3/24/2016

Figure 12. Activity Data Screen

Screen 3: Climate Data Screen

Each day, data describing the weather conditions was collected and stored in the application. In the initial stages of data collection, only the average temperature was recorded. Unfortunately, the information was not enough to analyze its impact on the activity's performance due to temperature variations in the morning and evening. Therefore, two values were recorded; one represented the low temperature and the other the high temperature. These two values provided a better understanding of the climate conditions.

In this screen as shown in Figure 13, data documenting the weather conditions was entered to describe the climate in the location of the project. Basic weather data was necessary for this part of the application, including temperature, high and low, wind speed, climate condition, precipitation and record date.

Temperature	Low	High	Humidity	Wind	Climate Condition	Presipitation	Record Date
65	48	81	68%	15	Overcasted	0	3/1/2016
61	48	73	76%	7	Clear	0	3/2/2016
55	37	72	69%	6	Clear	0	3/3/2016
54	43	64	74%	9	Clear	0	3/4/2016
49	43	55	72%	5	Clear	0	3/5/2016
53	43	63	65%	10	Clear	0	3/6/2016
53	43	63	73%	5	Clear	0	3/7/2016
53	36	70	73%	6	Overcasted	0	3/8/2016
55	37	73	70%	4	Overcasted	0	3/9/2016
60	50	70	88%	6	Overcasted	0	3/10/2016
60	57	63	70%	8	Overcasted	0	3/11/2016
57	50	63	64%	5	Clear	0	3/12/2016
55	50	59	78%	5	Overcasted	0	3/13/2016
60	52	68	80%	13	Clear	0	3/14/2016
56	48	64	41%	14	Clear	0	3/15/2016
58	52	64	32%	4	Clear	0	3/16/2016
57	50	64	44%	7	Clear	0	3/17/2016
44	36	52	70%	14	Rain	1	3/18/2016
45	32	57	36%	10	Clear	0	3/19/2016
49	32	66	41%	16	Clear	0	3/21/2016
63	48	77	50%	19	Clear	0	3/22/2016
68	52	84	47%	21	Clear	0	3/23/2016
48	37	59	36%	17	Clear	0	3/24/2016
50	34	66	34%	11	Clear	0	3/25/2016
59	46	72	33%	10	Clear	0	3/26/2016
46	34	57	31%	8	Clear	0	3/27/2016
51	32	70	41%	7	Clear	0	3/28/2016
58	46	70	30%	15	Clear	0	3/29/2016
69	57	81	65%	9	Overcasted	0	3/30/2016

Figure 13. Climate Data Screen

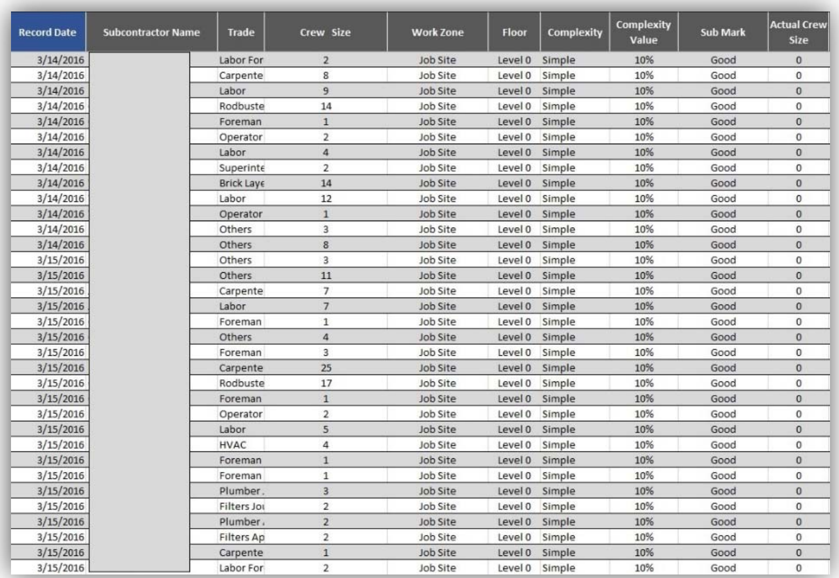
Screen 4: Crew Management Screen

It is important to mention that by the time of the execution of this research, crew reports were submitted by the subcontractors; although they were filled by hand on a physical document. A data collection interface was created to enter this information in a digital format so that the data can be organized and classified in a predetermined data set. The information was used to identify how influential the information from subcontractors' crew composition influenced in the analysis of project progress. Based on the reports provided by the subcontractor, a pattern was identified in the information. Multiple data entries aligned in the composition of the crew reports. A list of the items with the entries that were consistent in the reports was created. Based on the content of the list, a data collection interface was created to insert the information for each subcontractor. Additionally, two other pieces of information such as subcontractor mark and complexity were added to the design to explore its usability. The table with the items collected during this study is shown in Table 1.

Table 1. Crew Report Data List

Crew Report Data List	
Current Data Collected	Research Data
Accident	Complexity
Crew Size	Crew Size
Date	Date
Equipment	Sub. Mark
Project	Level
Reponsible	Subcontractor Name
Safety Meeting	Trade
Site Conditions	Work Zone
Trade	
Weather Descrition	
Work Description	
Work Hours	

The information contained in the crew management screen included the name of the company (subcontractor), trade description, crew size, work zone, floor, complexity, complexity value and subcontractor mark (sub mark). The "subcontractor name" column was covered to protect the identity of the company. This information can be found in Figure 14. An enlarged figure can be found in Appendix E.



Record Date	Subcontractor Name	Trade	Crew Size	Work Zone	Floor	Complexity	Complexity Value	Sub Mark	Actual Crew Size
3/14/2016		Labor For	2	Job Site	Level 0	Simple	10%	Good	0
3/14/2016		Carpente	8	Job Site	Level 0	Simple	10%	Good	0
3/14/2016		Labor	9	Job Site	Level 0	Simple	10%	Good	0
3/14/2016		Rodbuste	14	Job Site	Level 0	Simple	10%	Good	0
3/14/2016		Foreman	1	Job Site	Level 0	Simple	10%	Good	0
3/14/2016		Operator	2	Job Site	Level 0	Simple	10%	Good	0
3/14/2016		Labor	4	Job Site	Level 0	Simple	10%	Good	0
3/14/2016		Superinte	2	Job Site	Level 0	Simple	10%	Good	0
3/14/2016		Brick Laye	14	Job Site	Level 0	Simple	10%	Good	0
3/14/2016		Labor	12	Job Site	Level 0	Simple	10%	Good	0
3/14/2016		Operator	1	Job Site	Level 0	Simple	10%	Good	0
3/14/2016		Others	3	Job Site	Level 0	Simple	10%	Good	0
3/14/2016		Others	8	Job Site	Level 0	Simple	10%	Good	0
3/15/2016		Others	3	Job Site	Level 0	Simple	10%	Good	0
3/15/2016		Others	11	Job Site	Level 0	Simple	10%	Good	0
3/15/2016		Carpente	7	Job Site	Level 0	Simple	10%	Good	0
3/15/2016		Labor	7	Job Site	Level 0	Simple	10%	Good	0
3/15/2016		Foreman	1	Job Site	Level 0	Simple	10%	Good	0
3/15/2016		Others	4	Job Site	Level 0	Simple	10%	Good	0
3/15/2016		Foreman	3	Job Site	Level 0	Simple	10%	Good	0
3/15/2016		Carpente	25	Job Site	Level 0	Simple	10%	Good	0
3/15/2016		Rodbuste	17	Job Site	Level 0	Simple	10%	Good	0
3/15/2016		Foreman	1	Job Site	Level 0	Simple	10%	Good	0
3/15/2016		Operator	2	Job Site	Level 0	Simple	10%	Good	0
3/15/2016		Labor	5	Job Site	Level 0	Simple	10%	Good	0
3/15/2016		HVAC	4	Job Site	Level 0	Simple	10%	Good	0
3/15/2016		Foreman	1	Job Site	Level 0	Simple	10%	Good	0
3/15/2016		Foreman	1	Job Site	Level 0	Simple	10%	Good	0
3/15/2016		Plumber .	3	Job Site	Level 0	Simple	10%	Good	0
3/15/2016		Filters Jou	2	Job Site	Level 0	Simple	10%	Good	0
3/15/2016		Plumber .	2	Job Site	Level 0	Simple	10%	Good	0
3/15/2016		Filters Ap	2	Job Site	Level 0	Simple	10%	Good	0
3/15/2016		Carpente	1	Job Site	Level 0	Simple	10%	Good	0
3/15/2016		Labor For	2	Job Site	Level 0	Simple	10%	Good	0

Figure 14. Crew Management Screen

Screen 5: Activity Data Visualization Screen

The interactive data visualization screen was created through the implementation of multiple data sets and data filters that enabled the user to perform queries. Chart organization was a fundamental component in the data visualization application, so a screen was designed to exhibit only the graphical representation of the data. By creating multiple plots of the activity information, and the other variables considered in this study, the users could visualize and interpret project data quickly. Additionally, on this screen, multiple controllers were added to facilitate the navigation of the information.

A critical component in the application design was the representation of the dynamic nature of the construction information. For this reason, it was considered that by adding enabling buttons to control the charts, the users would be engaged in the process to generate various data queries. Equally important, the buttons that controlled the content of the plots demonstrated the dynamism of the charts as they adjusted to the new filters in real time, which allowed data to be flexible. This flexibility was critical to perform data relationships because it motivated the user to manipulate the application; this was an important achievement for this study. Users were able to provide valuable insights for the development of what specific datasets can be considered relevant for closer study.

Navigation buttons were located on the top ribbon of the screen, including work zone, floor, operation, activity, subcontractor name, and a time bar. As shown in figure 15. An enlarged figure is also found in Appendix F.

Figure 15. Data Navigation Ribbon

Additionally, in the activity data visualization screen, multiple data plots were organized to facilitate the visualization of the information. As shown in figure 16, the activity progresses the history of the installation of the columns and is graphically represented in the blue bars. In the same way, the chart shows the work area and the floors where column activity was executed. Because the "record date" was documented, it was possible to visualize the activity percent complete by day.

Furthermore, a visual representation of the crew sizes was provided in the chart with pink bars. The chart provides the amount of personnel assigned each day by the subcontractor responsible for the activity. Lastly, a plot with the planned finish and actual finish was displayed to identify the activity's duration visually. This chart was also represented with bars (dark blue and purple).

Figure 16. Activity Data Visualization Screen

Data Analysis

Data analysis included the examination of the audio files with the recorded responses from the three participants. Based on their answers to the questions, some of the most important aspects that they considered during the execution of a project regarding project control and monitoring were highlighted. Also, descriptions of the key obstacles that the professionals expressed interest in solving were evaluated during this process. Equally important, some of the concerns that the contractors had when documenting activity statuses during the monitoring of the project for performance assessment were investigated. These observations contributed to the improvement of their performance and the search for new, innovative ways to find solutions. In the same way, a description of the current processes that the team used to keep their schedule active was provided.

With the results obtained from this analysis, it was possible to identify what data was needed for the development of this study. As a result, a list of items that covered part of the information currently used by the GC for project control and monitoring was formulated. Additionally, a map for data collection and processing was designed to produce data visualization. These two components facilitated the development of an application that allowed the participants to visualize current activity progress and gain access to the activity's historical data.

Finally, the experiences documented by the participants after navigating the information with the data visualization application on the follow-up sessions were recorded to monitor their assimilation with and responsiveness to the new scheme. Their observations during this process were registered to find patterns in the different perceptions provided by each professional.

Based on the resulting information from the analysis of the follow-up session, the data visualization application required several modifications and adjustments to provide meaningful data to the participants. These changes allowed the exploration of new data relationships and the improvement of data visualization schemes. Providing the participants with the opportunity to experience the utilization of a data visualization application contributed to the learning process of understanding the needs of construction professionals and the limitations of current technologies.

Chapter 4: Results

Applied Research

As mentioned in Chapter 3, this study started with a feasibility analysis that was used to confirm the opportunity for developing the investigations. At the same time the information helped to determine the limitations and obstacles that would prevent this kind of research. After this step, a series of questions to the GC professionals provided fundamental information for the data collection process and analysis. During the interview session the participants provided information that allowed the identification of multiple element that they consider relevant when analyzing project progress. In the same way a table with the responses by participants was created, summarizing the most relevant aspect for project control and monitoring. The table with the interview chart can be found in Table 2. Additionally, documenting project activity data daily through observations was a relevant component of this research. The data gathered in the previous step was stored and processed, so the participants of this study were able to provide their experiences when they were navigating the visualization application.

Feasibility Test

The absence of data structure from the scheduling software made it difficult to separate the information into meaningful elements because it lacked consistency and order. The more specifications that the user can assign to the activity, the better it will be organized and classified.

Table 2. Interview Chart

Table 1.Interview chart - Analysis of Construction Progress Status (Based on Schedule Information)

	Project Engineer	Superintendent	Project Manager	Prevalent Answers	Answers List
Gather Information	Walk work site Schedule Drawings	Walk work site Schedule Costs/Budget	Conversations	Walk work site Schedule	Walk work site Schedule Drawings Costs/Busget Conversations
Time (Gather)	30-1 hour	5-10 minutes	Varies	Vary by professional	
Information Format	2D Drawings Observation	Project Scheduel Meetings Notes Observation	Project Schedule Pictures (Pull Planning) Meetings	Observations Meetings Gant Chart	Observation Gant Chart Meetings 2D Drawings Notes Pictures (Pull Planning)
Time To Perform Analysis	2 hours	Few Minutes	2 days	Vary by professional	Few minutes 2 Hours 2 Days
Concerns	Drawings-Accuracy Schdule complexity Professional Experience	Trust	Professional Experience Trust Accuracy	Trust Professional Experience	Trust Profession Experience Drawings-Accuracy Schedule Complexity Accuracy
Specific Data from the project	Construction Details Quality	Man Power Labor Location Quality	Accuracy	Quality	Construction Details Man Power Labor Location Quality Accuracy
Challenges	Accuracy Level of Detail of Schedule	Track own activities	Amount of data cources. Plug Information back to the system.	Vary by professional	Accuracy Level of Detail of Schedule Track Own Activities Amount of Data Sources Plug Information back to the System
Schedule Update Frequency	Weekly	Weekly	Weekly	Weekly	Weekly
Data Provider	Superintendent	Subcontractors	Subcontractor Superintendent	Superintendent Subcontractor	Superintendent Subcontractor
Information from Schedule			The impact of issues on activity progress.	Vary by professional	Impact of issues on activity progress

Naming Protocol

The design of a naming convention protocol solved the problem of data inconsistencies in the formation of project activity names. The proposed protocol discovered in the feasibility test had two key features; one was the data delimiters and the second was the naming protocol. For the data delimiters, the underscore ("_") character was used. This symbol was used to separate each of the multiple components that describe the activity in the construction schedule. In the same way, to create the naming protocol, it was considered that key values should identify the activities. Consequently, six pieces of data comprised the naming protocol syntax, and they are the following: operation, activity name, category, work zone, floor, and division number. A figure that shows the syntax of the activity name is shown in Figure 17.

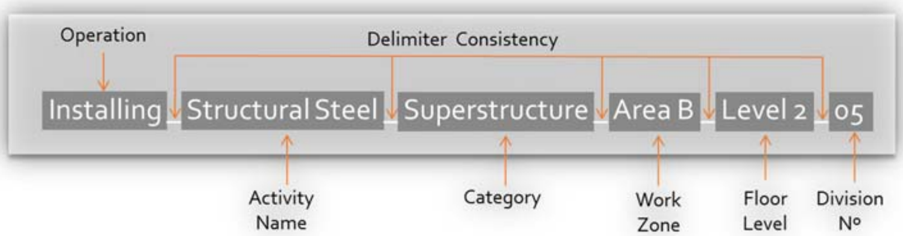


Figure 17. Activity Name Text Syntax

Based on the Master Format® Numbers and Titles classification system, a tool was designed to reduce the time taken to create the activity list and to insert all the pieces of information that the activity needed to be identified according to the proposed protocol. One of the purposes of creating this tool was to reduce the amount of inconsistencies generated by the traditional method, by providing the users with a model that organizes the information. The process to organize and classify the information was automated. Some of the fields within the tool have predefined options to help the user select the information instead of typing it in. Predefined information reduces the opportunity for users to enter data manually, which is one of the causes of poor consistency. The tool proved to be flexible enough to allow the user to enter data manually in the respective fields.

Because the naming protocol designed to determine the name of the activities used the Master Format® Numbers and Titles classification system, the application allowed the user to select the division to work and assign the activity name from a list of options linked to the respective division. By making the activity name field dependent on the divisions, the user accurately assigned names to the activities using a standardized format. The interface of the “activity builder” tool is shown in figure 18.

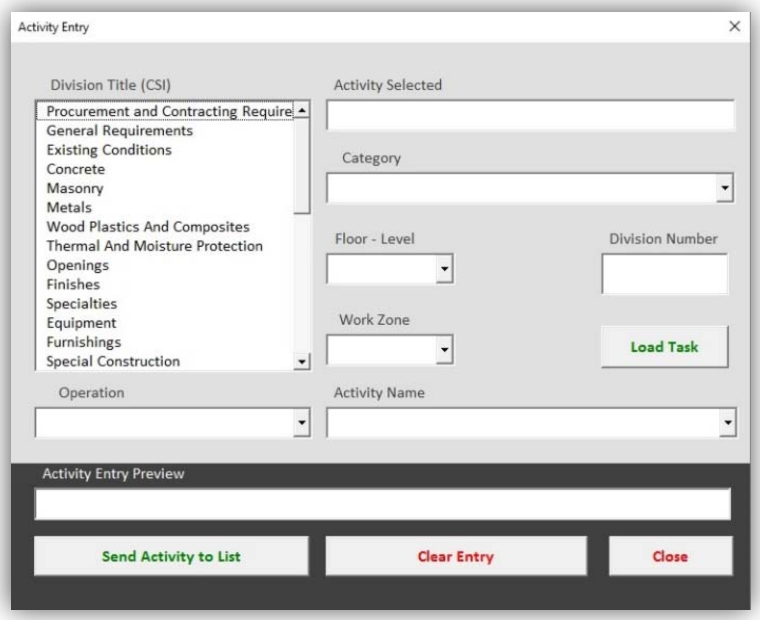


Figure 18. Activity Builder Tool

Data Formatting

Since the data exported from the construction schedule did not possess the syntax required to prove the usability and value of data visualization, it was formatted to adapt the information to a useful order using the tool designed to execute this phase. Each activity imported from P6 was assigned with a new name based on the proposed naming protocol. The new activity name did not alter the meaning of the activities; the tool only provided a syntax that allowed for the organization of the information.

Crew Report

Currently, the general contractor requests that subcontractors document their crew information daily, which is an advantageous opportunity to gather quickly the information and copy the data into the data visualization application. This information was created manually in a physical document, which made interpreting the persons' handwriting problematic. There were multiple inconsistencies in the documentation of subcontractors' data. Also, missing and ambiguous information made the process more difficult. Despite these limitations, the database was created with the available information.

The first step utilized for data entry was accessing the digital format (PDF) of the physical records of the reports provided by the subcontractor, documenting the crew's composition. After evaluating the content from the reports, multiple fields were identified as patterns in the information submitted by the subcontractors. Therefore, a data collection interface was created with part of this information to facilitate this process and insert data faster. Additionally, in this interface, two items not listed in the pattern were proposed to prove its usability. These two elements were 1) complexity and 2) grade.

In the case of "complexity," the data tries to assign a value to the complex nature of the tasks performed by the trades. For example, if the mason trade was installing a custom-made architectural element in the project, which requires a high level of experience, the subcontractor and the superintendent may agree on the complexity of the activity and assign the corresponding value. This element had five predefined complexity levels: simple, well structured, complex, alarming complex, and error-prone. Each grade had a numerical value presented in percentage. The percentage assigned to the categories was generated by establishing 100% as the highest level of complexity that an activity can get. Then, because the number of categories was five, the values were distributed in increments of 25 except for the “simple" complexity level because it was considered that this category should not have 0% complexity. The attribution of values resulted in simple 10%, well structured 25%, complex 50%, alarming complex 75%, and error-prone 100%. As for the grade, this data can be assigned to the crew to evaluate their performance. The categories to grade the subcontractor were fair, good, and bad. The last two items added to the crew report did not contain actual data; the information entered in these fields was for simulation purposes to prove their usability.

Second Interview Session

As a result of the second interview session the participants requested changes in the system to visualize different data sets, which reflected the analyses that they currently perform to determine the project performance. Additionally, recording a different dataset was requested to perform other analyses that were not anticipated in this research. In the same way, comments about the usability of the information presented in the data visualization application were recorded.

Summary of Results

In this section, the results of the usability phase confirmed the relevance of data visualization for project control and monitoring. The usability process consisted of two categories: 1) engagement from the professionals and 2) data demand. In the case of this research, "engagement" described how valuable the information displayed was for the participant. On the other hand, "data demand" represented the degree of interest of the participant to request more data sets or the visualization of different information. These two factors were present during the evaluation of the information.

Activity Data Update

Participating in the daily sessions for WWP update provided a unique perspective to experience the process of how this information was obtained and posted on the board. One issue observed was that part of the course of storing the information contained on the board was through a digital image format, which is part of the process to capture and transfer data from the physical to digital. However, the data transferred was not significantly abundant to create a flexible data visualization analysis or data mining for problem-solving sessions.

Based on the analysis of the information collected during this study actual activity progress could provide critical information for the development of extensive analysis and data mining. Thus, it is paramount to keep a record of the activity development on a daily basis and provide actual progress. A particular scenario observed during the WWP session took place when one of the subcontractors announced that his crew was not able to reach the percentage of completion of the activity planned for that day in particular. Moreover, the subcontractor explained the reason why the crew did not complete the task as planned, providing the reasons and the particular delay they caused. The session's moderator updated the information on the board with the information provided by the subcontractor and validated by the superintendent. Although, the percent complete data remained as planned, and it was not modified. Having this data could contribute to a further analysis of project planning. Therefore, more data collected documenting every aspect involved in the construction could result in better investigations of a project's performance and produce a higher level of analysis.

Crew Reports

As mentioned before, this process required manual entry, and the handwriting was not clear in many cases, which made the data transfer more complicated, time-consuming, and inaccurate.

Data Visualization Usability

The usability phase consisted of sessions where the participant had the opportunity to operate the data visualization application and comment their observations. In the first session, one week after data collection started, the participant received the information contained in the database and began operating the program. Based on the participant's opinion, in this case, the project engineer, the dashboard interphase design allowed the user to understand very quickly how to use the advanced filters to dynamically analyze the information displayed in the charts. Additionally, the application facilitated a general understanding of the current state of the activity progress. Because the information collected during that week did not provide enough data to assess the activities performance, the participant was not able to determine if the application would provide value to his analysis process.

Although the application provided sufficient resources to comprehend its capability to produce more information, due to the project engineer requested the inclusion of two plots where a general activity performance could be displayed and a week or a month could be selected for each chart to compare productivity rates. The participant commented, “We monitor the productivity rates of the project, and it would good to have the information to compare project performance of the current and previous month.” It was possible to comply this request because the information to create the charts already existed in the database, which enabled the opportunity for exploration of new datasets. In another comment, the project engineer mentioned was that “this kind of application would allow us to show the owner the progress that we are making in the project, even if the changes are not visible”. The data visualization resulted from the participant’s request is shown in Figure 19.

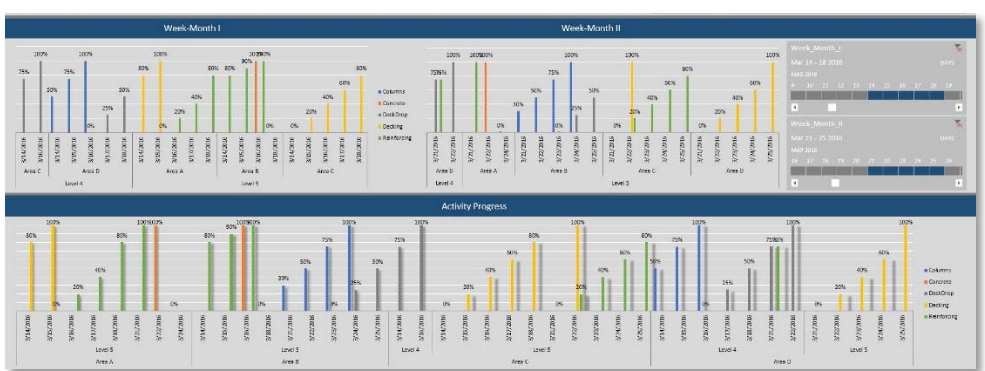


Figure 19. Activity Progress Week/Month Comparison

For the second week's presentation, the session consisted of the demonstration for the program to produce new plots as requested and the capability to continue the exploration of different datasets. In this case, the database contained a larger quantity of data, which permitted the users to perform multiple queries and analyze the information provided in the plots with a dynamic visualization system, so as the user selects the various filters, the plots update the information in real time.

Furthermore, in this session, while the superintendent was examining the different plots, a particular observation was commented during the testing. The superintendent mentioned, "documenting a grade mark to subcontractor's performance; I would be able to consider repeating working with that subcontractor in another project based on my experience working with them." Evaluating subcontractors' performance was a suggestion proposed that could provide value to the project analysis, but it was not clear until construction professionals evaluated the content of the information. The comment from the participant validated the use of this piece of information for historical analysis in future projects. Figure 20 displays the plot produced to visualize the subcontractor mark.

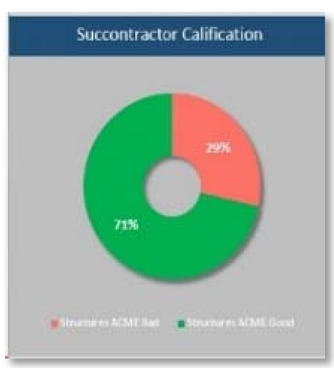


Figure 20. Subcontractor Mark

In the same way, while the project engineer was operating the data visualization application, he identified the benefit of collecting activity complexity. The participant considered that documenting this kind of information would allow him to make decisions regarding the control and monitoring process of the activities with high levels of complexity. The scenario described by the project engineer was that, by knowing the complexity level of an activity he would closely monitor the performance of the trade that is executing the activity, and analyze what would the subcontractor need to complete the task as planned. The comment provided by the project engineer was evidence that data visualization would allow the users to participate proactively in problem-solving analyzes. Figure 21 displays the activity complexity chart.

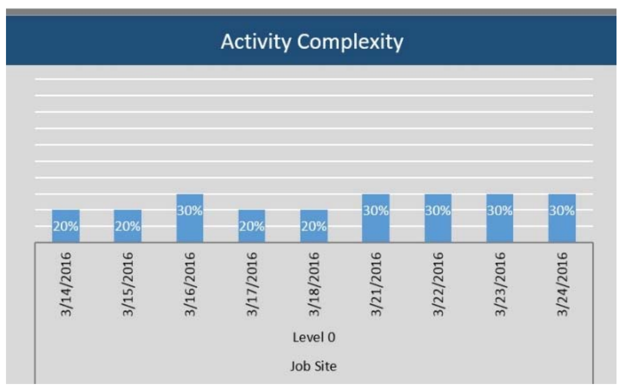


Figure 21. Activity Complexity

In the third week, the project manager explored the plot's content in more detail and performed queries in the "Activity Data Visualization" screen that provided valuable information to visualize the database content. In this particular case, after documenting the WWP data for two and a half weeks, the start and finish dates of a few activities were added into the system's database. While revising the information, the project manager observed that an activity was finished two days after it was scheduled.

Because the session took place in the middle of the week, the contractor did not notice that the activity had a delay. This scenario was evidence that documenting the activity statuses and creating a data repository with daily information from the construction site is critical for the project analysis that could improve decision-making processes. The chart analyzed by the project manager can be seen in Figure 22.

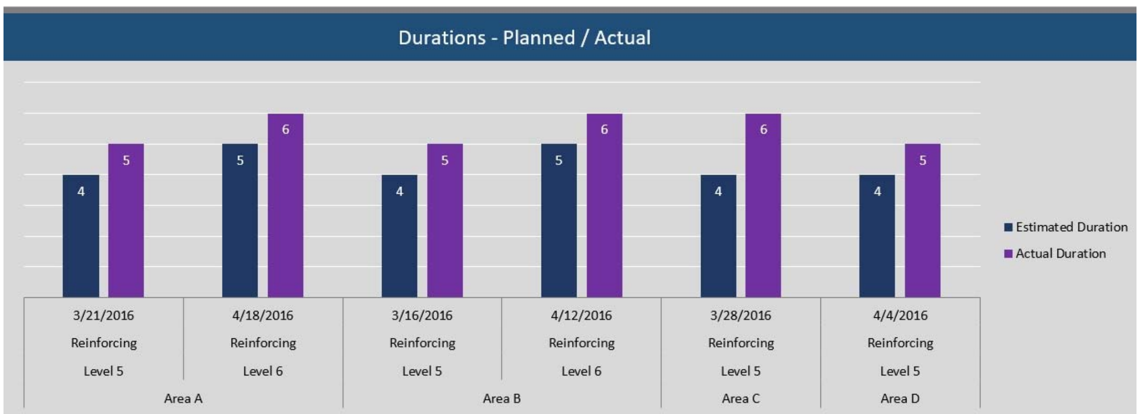


Figure 22. Activity Durations

In the last session, a few more queries were performed by the project manager; the focus was on the capabilities of the application in increasing the information that it can provide. One new request was to provide the percentage of achievement of the planned activities for each week. This operation was performed weekly, and the information is computed manually. This kind of request confirms the value of an application capable of storing and retrieving data in the construction industry.

During this process, it was observed that the application could provide valuable information for activity progress analysis. Additionally, the capability of the user to be familiar with the system and retrieve the information dynamically was relatively easy despite the time that was destined for navigation training. Another observation that was paramount in the usability of construction data visualization was the level of skill required to operate and interpret the data. As the previous statements from the participants, they were able to understand the information quickly and search for specific data to analyze the current state of activity progress.

There were multiple requests by the participants for the expansion of the information that the data visualization application provided, which is evidence that data visualization is capable to provide valuable information for decision-making and project analysis in the construction process. Current practices do not support data workflow by electronic formats, which creates numerous disadvantages regarding documenting the information and labor intensity for data processing.

Based on the research findings, data visualization provides a greater opportunity for construction professionals to identify what activity may compromise the construction schedule and to analyze the multiple factors that could influence their execution. In the same way, it enabled visual analytics of project progress, which is one of the key features of this application. Moreover, by using data visualization systems, construction professionals would be able to disseminate the information across disciplines more effectively because graphic data can be interpreted much more easily than long and complicated data tables. The application of data visualization may result in the formation of more collaborative problem-solving environments where decisions may flow faster. Equally important, the decisions made to update or modify the construction schedule can be supported based on the information contained in the system's database.

Results Discussion

The participants provided important insight when describing the usability of the data visualization application. In the same way, multiple results were found during the execution of this research that provided the opportunity to study other areas of data visualization. These results were fast data comprehension, data communication, and reuse and combination of data.

Based on the observations of participants navigating and interpreting the information displayed in the data visualization application, it was evident that they were able to analyze the information quickly even with few indications of plots' content. The quick comprehension of the information demonstrated that visual data is easy to understand and analyze. In the same way, as the information was easy to interpret, one of the participants (project engineer) identified the opportunity to communicate this information to non-construction professionals. In his comment, the participant saw the possibility of presenting the information to the owner of the project so that a clearer communication of the status of the current project could be established.

Another result of this research was the usability of activity data. Collecting and storing historical activity information allowed the data to be reused and combined with multiple data sets to create different analyses and explore relationships among data. The potential of combining different data sets enhanced the understanding of the participants to analyze the project activities progress.

A significant result of this research was the creation of the data visualization application. The application allowed the transformation of the activity's text from the construction schedule for its classification, organization, and storage to produce project data visualizations.

Chapter 5: Discussion & Conclusions

This document describes the influence of construction data visualization during building operations based on activity progress data. For the execution of this study, it was necessary to understand the current practices of data collection, storage, classification, and structure to analyze its limitations. Based on the constraints found, an application was proposed for organizing and classifying construction schedule data to create a database and use it for data analytics. Additionally, two other elements such as crew size and weather conditions were included in the data collection process; they possessed valuable content for this investigation, so they were also documented and stored in the database. By recording not only the activity statuses but also other factors that influence construction progress, multiple relationships among data were able to be analyzed.

Consequently, a data pipeline that allowed the formation of a database with current activity information was presented and put into practice for usability. Through the implementation of this pipeline, it was possible to use advanced filters that allowed data retrieval in a graphic interphase. Moreover, a dynamic data visualization application was created to provide the participants with an application whereby they could interact and manipulate the information in real time. The application also helped the participants of this study to experience the dynamism of the content of the database when navigating construction activity records. This kind of interactions contributed to the documentation of the ways in which a visual representation of construction data can impact the analytical process to make decisions, and it also allowed them to describe their insights.

Limitations

There were many limitations during the execution of this study. For example, it was limited to only one GC, the number of activities observed were those associated with the concrete structure; the time-frame employed was limited to one month, and lastly, there was a lack of available computer applications to execute this kind of investigation. The primary limitation observed during this research was the lack of a computer program designed to collect project activity data. Equally important was the absence of data syntax in the data created by the participants, to produce construction schedules. These two limitations were critical in the development of this research. Nevertheless, it was still possible to collect enough information to test and validate the assumptions of construction data visualization.

General Contractor

By conducting an investigation with one GC, the study only provided results based on the current practices of that particular organization. These practices influenced the execution of this research, as each company may implement different operation systems, which could lead to different results or exposure to various challenges and limitations. In this case, the GC used Lean Construction instruments to perform their planning and control exercises. Therefore, results revealed how lean principles provided an advantage for data collection. However, the use of Lean Construction is not essential to collect and visualize the construction data.

Project Activities

The small number of project activities was another limitation. The activities observed during this study were those associated with the construction phase of the concrete structure of a new building project, which included the columns, decking, concrete, and rebar. Because the number of items documented from the construction schedule was small in quantity, reporting their statuses did not represent a challenge. Additionally, no significant issue was documented during the observation period. Keeping a record of few activities led to a reduced capability for exploration of project control data.

Observation Timeframe

Another limitation was the amount of days assigned to complete the data collection process for the observation phase. In this case, data was collected for twenty-two (22) working days, not including weekends, for a total of thirty (30) calendars days. This amount of time allowed the collection of twenty-two data points for each activity associated with the concrete structure. The amount of data may not represent the constraints related to the execution of each line item recorded. As the project progresses, multiple elements might add challenges for executions, which could provide valuable data. Additionally, weather conditions were stable. Precipitation was little or insignificant; in the same way, the temperature was not extreme. Ideal weather conditions prevented the analysis of the impact of climate on crew performance.

Lack of Software Applications

The first obstacle found was that the syntax used to produce the construction schedule did not have a well-defined structure that allows the integration of the systems to organize the data in an efficient way. Inconsistencies in the information structure and lack of a regular order to enter information prevented the automation of data processing. Therefore, an application was created to solve this issue. This application consisted in the integration of construction standards and current practices for activity naming conventions. The lack of a software application capable of importing the information from the construction schedule for classification and storage to produce construction data plots contributed to this limitation. Therefore, creating an application that addressed the requirements for an automated process was important. This application reduced the time employed for data entry and the production of data visualization charts.

Conclusion

The intention of conducting this research was to identify the importance of visual analytics in the construction process and understand the value of this instrument as a decision-making mechanism. Findings from this study revealed that there is potential for further development of this kind of research. Several unexplored areas need testing and analyses to discover the full impact of visual data analytics in the construction industry. Construction projects generate data from multiple sources; unfortunately, much of this data is not collected or used during project execution to support decision-making because it requires a diversity of skills to display visual information. The implementation of data valuations could improve how project managers evaluate a project's progress status and support their decisions with reliable information. Based on the findings of this study, the communication of visual data was efficient, and the information was easy to interpret, which could lead to a collaborative problem-solving environment.

During the development of this study, a gap was identified between the physical world, which in this case is the construction site, and the digital data, which resulted in the development of an application that helped to create a bridge for information technology. Also, it was found that by collecting data, storing it in digital format, and establishing a concise syntax, the information from the construction schedule can be used and reused for project analysis. Moreover, documenting construction activity progress on a daily basis increases the possibility of improving project performance.

An important fact is that, through the consistent use of data analytics, the decisions that the team makes to implement a solution that improves construction productivity can be put into practice, and the team can see the results of the new strategy. Therefore, the more data the project team collects from the work site, the better the information produced by the data visualization. Additionally, due to this being a data-driven system, the team can observe measurable results for planned solutions as the project progresses. Another benefit is that project decision-makers are able to analyze quantitative data contained in the data visualization application instead of viewpoints.

Based on this study, the implementation of a data visualization system can be executed at any time during the construction process, although to obtain better results from the visualization application, it should be implemented in the initial stages of the construction project. Due to the limitations of current syntax in the construction schedules, a process of data reconfiguration is critical for adequate storage for data manipulation.

Technology has made it possible to build multiple applications that allow construction professionals to take advantage of the information produced during the planning phase of building construction. However, these efforts to produce construction visualization systems have not been aligned to create a robust data management system that allows data analytics. To date, there is not a significant amount of information or research on the application of data visualization in construction projects. Construction management should be a data-driven decision discipline, where every factor that impacts the project progress and execution should be analyzed to find solutions for an optimal project performance. If the construction industry were data driven, databases would support analysis of performance and productivity and ultimately create simulations that would enable construction professionals to improve their planning and decision making.

Significance of Study

The study could provide information on the current practices of construction activity data collection and the limitations to implementing data visualization. Based on the findings of this study, it would be beneficial to the construction industry considering that data visualization could represent an important source of information to support project decisions increase performance and optimize project control. The greater demand for data justifies the need for more efficient and well-structured data visualization systems. Thus, the participants from the construction project where this study was implemented were able to visualize project activity data and analyze the benefits that it would provide to their analyses.

For researchers, the study would help them uncover critical areas in the data collection of construction activities that were not able to be explored before. Additionally, future researchers may use the information provided in this study to continue the investigation, documenting different data sets and exploring new data combinations to discover unique elements that would improve construction project analysis and data visualization.

References

Abbasside, A., & Gray, D. (2015). Assessing the influence of automated data analytics in cost and schedule performance. *Creative Construction Conference, Science Direct, Procedia Engineering*, 123, 3-6.

Brubeck, B., Mezza, R., & Leanne, D., (2009). Interactive Visualization – Survey. In D. Lane & J., (Ends), *Human Machine Interaction*, 27-46

Behzadan, A., Manassa, C., & Pradhan, A. (2015). Enabling Real-Time Simulation of Architecture, Engineering, Construction, and Facility Management (AEC/FM) Systems: A Review of Formalism, Model Architecture, and Data Representation. *Journal of Information Technology in Construction*, 20, 1-23.

Cheng, T., & Teaser, J., (2010). Real-Time Data Collection and Visualization Technology in Construction. *Construction Research Congress*, *ASCE*, 339-348.

Chiu, CY., & Russel, A. D., (2010). Design of a Construction Management Data Visualization Environment: A Top-Down Approach. *Automation in Construction*, 20, 399-417.

Cybulski, J. L., Keller, S., Nguyen, L., & Sundage, D. (2013). Creative problem solving in digital space using visual analytics. *Computers in Human Behavior*, 42, 20-35.

Dang, T., & Bierstadt, HJ., (2015). 4D Relationships: The Missing Link in 4D Scheduling. *Journal of Construction Engineering and Management*, 142, 1-16

Goedert, J. D., & Meadati, P. (2008). Integrating Construction Process Documentation into Building Information Modeling. *Journal of Construction Engineering and Management*, *ASCE*, 7, 509-516.

Golparvar-Fard, M., Tang, P., Cho, K. Y., & Siddiqui, M. K., (2013). Grand Challenges in Data and Information Visualization for the Architecture Engineering Construction and Facility Management Industry. *Computing in Civil Engineering*, 849-856.

Hartmann, T., Meerveld, H., Vossebeld, N., & Adriaanse, A. (2012). Aligning building information model tools and construction management methods. *Automation in Construction*, 22, 605-613.

Healey, C., & Enns, J. (2012). Attention and Visual Memory in Visualization and Computer Graphics. *IEEE Transactions on Visualization and Computer Graphics*, 7, 1170-1188.

Heer, J., & Shneiderman, B. (2012). Interactive Dynamics for Visual Analysis. *Association for computing Machinery*, 1-26.

Karshenas, S., & Sharma, A. (2010). Visually Scheduling Construction Projects. *Construction Research Congress*, *ASCE*, 490-499.

Kelleher, C., & Wagener, T. (2011). Ten Guidelines for Effective Data Visualization in Scientific Publications. *Environmental Modeling & Software*, 6, 822-827.

Keim, D. A., Mansmann, F, Schneidewind, J., Thomas, J., & Ziegler, H. (2008). Visual Analytics: Scope and Challenges. In S.J. Simoff, M.H. Böhlem, & A. Mazeika (Eds.), *Visual Data Mining*, 76-90.

Kim, C., Park, T., Lim, H., & Kim, H. (2013). On-Site Construction Management Using Mobile Computing Technology. *Automation in Construction*, 415-423.

Lee, N., & Rojas, E.M., (2013). Visual representation for monitoring project performance: Developing novel prototypes for improved communication. *Journal of Construction Engineering and Management*, 8, 994-1005.

Mahalingam, A., Kashyap, R., & Mahajan, C. (2009). An Evaluation of the Applicability of 4D CAD on Construction Projects. *Automation in Construction*, 19, 148-159.

Mantmann, F., Ficher, F., & Keim D. (2012). Dynamic Visual Analytics-Facing the Real-Time Challenge. *Expanding the Frontiers of Visual Analytics and Visualization,* 69-80.

Rojas, E., & Lee, N. (2007). Visualization of Project Control Data: A Research Agenda. *Computing in Civil Engineering, ASCE*, 26-34.

Root-Berstein, R., & Root-Berstein, M. (2004). Artistic Scientists and scientific artists: The link Between Polymathy and Creativity.

Russell, A.D., Chiu, CY., & Korde, T. (2009). Visual Representation of Construction Management Data. *Automation in Construction*, 18, 1045-1062.

Song, K., Pollais, S. N., & Pena-Mora, F. (2005). Project Dashboard: Concurrent Visual Representation Method of Project Metrics on 3D Building Models. *Computing in Civil Engineering, ASCE*, 1-12.

Sun, GD., Wu, YC., Liang R. H. & Liu SX. (2013). A Survey of Visual Analytics Techniques and Applications: State-of-the-Art Research and Future Challenges. *Journal of Computer Science and Technology*, 25, 825-867.

Wang, L. (2007). Using 4D modeling to Advance Construction Visualization in Engineering Education. *Computer Integrated Construction Research Program, Department of Architecture Engineering, The Pennsylvania State University*, Technical Report No. 51, iii-127.

Appendix A: Interview Questionnaire I

Interview Questions

1. Describe the process you use to gather information about current status of the project each time that you visit the job site?

2. How long does the process described in question #1 typically take to complete?

3. In what format is this information is presented to you?

4. How much time does it take you to analyze schedule information specific to the project?
   Why?

5. Do you have concerns about the accuracy and reliability of how the information is been collected related to the project status?
   If so, please explain further?
   Why you think is in accurate?

6. What specific information do you look for when visiting a job site?

7. What are the challenges and your concerns when documenting project activities' statuses?

8. Who provides you with the information at the job site?

9. How often the original schedule that was created for the project is updated with the actual durations for the project? (PM, PE, Superintendent)

10. Is the information about actual status exclusively accessible though job site visits?
    a. Explain why, or why not?

11. What schedule related information could be helpful in your positions to make informed decisions about the current status of the project?

Appendix B: Interview Questionnaire II

Follow up Questions (Usability)

1. How valuable do you consider the schedule information presented?
   a. Explain why, or why not?

2. Do you consider that the graphic representation of data is beneficial for your analysis?
   a. Explain why, or why not?

3. Did the visual representation of the information collected improved your analysis of the project status of the project?

4. How difficult was for you to navigate through the information in the visualization tool?

5. Would you consider implementation of this system to your daily operations? Please explain

Appendix C: Data Import Screen

LEV3 AREA C							18	1/19/2016	2/11/2016	100	18	42388	2/11/2016
Installing	Deck	Structure	Area C	Level 03		3	5	1/19/2016	1/25/2016	100	5	42388	1/25/2016 Finished
Installing	Reinforcing	Structure	Area C	Level 03		3	5	1/26/2016	2/1/2016	100	5	42395.33333	2/1/2016 Finished
Installing	Rough MEP	Structure	Area C	Level 03		3	3	1/28/2016	2/1/2016	100	3	42397.33333	2/1/2016 Finished
Pouring	Concrete	Structure	Area C	Level 03		3	1	2/2/2016	2/2/2016	100	1	42402.33333	2/2/2016 Finished
Installing	Columns	Structure	Area C	Level 03		3	4	2/8/2016	2/11/2016	100	4	42408.33333	2/11/2016 Finished
LEV3 AREA B							30	12/18/2015	2/1/2016	100	30	42355	2/1/2016
Installing	Deck	Structure	Area B	Level 03		3	18	12/18/2015	1/15/2016	100	18	42355	1/15/2016 Finished
Installing	Reinforcing	Structure	Area B	Level 03		3	19	12/21/2015	1/19/2016	100	19	42359	1/19/2016 Finished
Installing	Rough MEP	Structure	Area B	Level 03		3	19	12/21/2015	1/19/2016	100	19	42359	1/19/2016 Finished
Pouring	Concrete	Structure	Area B	Level 03		3	1	1/20/2016	1/20/2016	100	1	42389	1/20/2016 Finished
Installing	Columns	Structure	Area B	Level 03		3	4	1/27/2016	2/1/2016	100	4	42396.33333	2/1/2016 Finished
LEV3 AREA A							20	1/11/2016	2/5/2016	100	20	42380.33333	2/5/2016
Installing	Deck	Structure	Area A	Level 03		3	6	1/11/2016	1/18/2016	100	6	42380.33333	1/18/2016 Finished
Installing	Reinforcing	Structure	Area A	Level 03		3	4	1/20/2016	1/25/2016	100	4	42389	1/25/2016 Finished
Installing	Rough MEP	Structure	Area A	Level 03		3	6	1/18/2016	1/25/2016	100	6	42387	1/25/2016 Finished
Pouring	Concrete	Structure	Area A	Level 03		3	1	1/26/2016	1/26/2016	100	1	42395.33333	1/26/2016 Finished
Installing	Columns	Structure	Area A	Level 03		3	4	2/2/2016	2/5/2016	100	4	42402.33333	2/5/2016 Finished
LEVEL 4							30	2/2/2016	3/14/2016	30	9	42402.33333	3/14/2016
LEV4 AREA A							17	2/10/2016	3/3/2016	17.65	3	42410.33333	3/3/2016
Installing	Deck	Structure	Area A	Level 04		3	2	2/10/2016	2/12/2016	100	2	42410.33333	2/12/2016 Finished
Installing	Reinforcing	Structure	Area A	Level 04		3	9	2/15/2016	2/25/2016	0	0	42415.33333	2/15/2016 Started
Installing	Rough MEP	Structure	Area A	Level 04		3	8	2/16/2016	2/25/2016	0	0	42416.33333	2/25/2016 Planned
Pouring	Concrete	Structure	Area A	Level 04		3	1	2/26/2016	2/26/2016	0	0	42426.33333	2/26/2016 Planned
Installing	Columns	Structure	Area A	Level 04		3	5	2/26/2016	3/3/2016	0	0	42426.33333	3/3/2016 Planned
LEV4 AREA C							17	2/15/2016	3/8/2016	0	0	42415.33333	3/8/2016
Installing	Deck	Structure	Area C	Level 04		3	6	2/15/2016	2/22/2016	0	0	42415.33333	2/22/2016 Planned
Installing	Reinforcing	Structure	Area C	Level 04		3	6	2/22/2016	2/29/2016	0	0	42422.33333	2/29/2016 Planned
Installing	Rough MEP	Structure	Area C	Level 04		3	5	2/23/2016	2/29/2016	0	0	42423.33333	2/29/2016 Planned
Pouring	Concrete	Structure	Area C	Level 04		3	1	3/1/2016	3/1/2016	0	0	42430.33333	3/1/2016 Planned
Installing	Columns	Structure	Area C	Level 04		3	4	3/3/2016	3/8/2016	0	0	42432.33333	3/8/2016 Planned
LEV4 AREA D							16	2/22/2016	3/14/2016	0	0	42422.33333	3/14/2016
Installing	Deck	Structure	Area D	Level 04		3	6	2/22/2016	2/29/2016	0	0	42422.33333	2/29/2016 Planned
Installing	Reinforcing	Structure	Area D	Level 04		3	5	3/1/2016	3/7/2016	0	0	42430.33333	3/7/2016 Planned
Installing	Rough MEP	Structure	Area D	Level 04		3	4	3/2/2016	3/7/2016	0	0	42431.33333	3/7/2016 Planned
Pouring	Concrete	Structure	Area D	Level 04		3	1	3/8/2016	3/8/2016	0	0	42437.33333	3/8/2016 Planned
Installing	Columns	Structure	Area D	Level 04		3	4	3/9/2016	3/14/2016	0	0	42438.33333	3/14/2016 Planned
LEV4 AREA B							20	2/2/2016	2/29/2016	45	9	42402.33333	2/29/2016
Installing	Deck	Structure	Area B	Level 04		3	6	2/2/2016	2/9/2016	100	6	42402.33333	2/9/2016 Finished

Appendix D: Activity Data Screen

Action	Activity	Category	Work Zone	Floor	Division N°	Estimated Duration	Planned Start	Planned Finish	% Complete	Actual Duration	Actual Start	Actual Finish	Status	Recod Date
Inst.	Columns	Structure	Area D	Level 4	3	5	3/10/2016	3/16/2016	50%	0	3/10/2016		Started	3/14/2016
Drop	DeckDrop	Structure	Area C	Level 4	3	5	3/12/2016	3/17/2016	75%	0	3/12/2016		Started	3/14/2016
Inst.	Decking	Structure	Area A	Level 5	3	4	3/9/2016	3/15/2016	80%	4	3/9/2016		Started	3/14/2016
Inst.	Decking	Structure	Area C	Level 5	3	5	3/16/2016	3/22/2016	0%	0	3/14/2016		Started	3/14/2016
Inst.	Reinforcing	Structure	Area B	Level 5	3	4	3/8/2016	3/15/2016	80%	0	3/8/2016		Started	3/14/2016
Inst.	Columns	Structure	Area D	Level 4	3	5	3/10/2016	3/16/2016	75%	0	3/10/2016		Started	3/15/2016
Inst.	Decking	Structure	Area A	Level 5	3	4	3/9/2016	3/15/2016	100%	4	3/9/2016	3/15/2016	Finished	3/15/2016
Inst.	Decking	Structure	Area C	Level 5	3	5	3/16/2016	3/22/2016	20%	0	3/14/2016		Started	3/15/2016
Inst.	Reinforcing	Structure	Area B	Level 5	3	4	3/8/2016	3/15/2016	90%	0	3/8/2016	3/15/2016	Started	3/15/2016
Inst.	Reinforcing	Structure	Area A	Level 5	3	4	3/16/2016	3/21/2016	0%	0	3/15/2016		Started	3/15/2016
Inst.	Columns	Structure	Area D	Level 4	3	5	3/10/2016	3/16/2016	100%	5	3/10/2016	3/16/2016	Finished	3/16/2016
Pouring	Concrete	Structure	Area B	Level 5	3	1	3/16/2016	3/16/2016	100%	1	3/16/2016	3/16/2016	Finished	3/16/2016
Drop	DeckDrop	Structure	Area C	Level 4	3	5	3/12/2016	3/17/2016	100%	4	3/12/2016	3/16/2016	Finished	3/16/2016
Drop	DeckDrop	Structure	Area D	Level 4	3	5	3/16/2016	3/22/2016	0%	0	3/16/2016		Started	3/16/2016
Inst.	Decking	Structure	Area C	Level 5	3	5	3/16/2016	3/22/2016	40%	0	3/14/2016		Started	3/16/2016
Inst.	Reinforcing	Structure	Area A	Level 5	3	4	3/16/2016	3/21/2016	20%	0	3/15/2016		Started	3/16/2016
Inst.	Reinforcing	Structure	Area B	Level 5	3	4	3/8/2016	3/15/2016	100%	5	3/8/2016	3/16/2016	Finished	3/16/2016
Drop	DeckDrop	Structure	Area D	Level 4	3	5	3/16/2016	3/22/2016	25%	0	3/16/2016		Started	3/17/2016
Inst.	Decking	Structure	Area C	Level 5	3	5	3/16/2016	3/22/2016	60%	0	3/14/2016		Started	3/17/2016
Inst.	Reinforcing	Structure	Area A	Level 5	3	4	3/16/2016	3/21/2016	40%	0	3/15/2016		Started	3/17/2016
Inst.	Columns	Structure	Area B	Level 5	3	4	3/16/2016	3/21/2016	0%	0	3/18/2016		Started	3/18/2016
Drop	DeckDrop	Structure	Area D	Level 4	3	5	3/16/2016	3/22/2016	50%	0	3/16/2016		Started	3/18/2016
Inst.	Decking	Structure	Area C	Level 5	3	5	3/16/2016	3/22/2016	80%	0	3/14/2016		Started	3/18/2016
Inst.	Reinforcing	Structure	Area A	Level 5	3	4	3/16/2016	3/21/2016	80%	0	3/15/2016		Started	3/18/2016
Inst.	Columns	Structure	Area B	Level 5	3	4	3/16/2016	3/21/2016	30%	0	3/18/2016		Started	3/21/2016
Drop	Deckdrop	Structure	Area D	Level 4	3	5	3/16/2016	3/22/2016	75%	0	3/16/2016		Started	3/21/2016
Inst.	Decking	Structure	Area D	Level 5	3	5	3/23/2016	3/29/2016	0%	0	3/21/2016		Started	3/21/2016
Inst.	Reinforcing	Structure	Area A	Level 5	3	4	3/16/2016	3/21/2016	100%	5	3/15/2016	3/21/2016	Finished	3/21/2016
Inst.	Reinforcing	Structure	Area C	Level 5	3	4	3/23/2016	3/28/2016	0%	0	3/21/2016		Started	3/21/2016
Inst.	columns	Structure	Area B	Level 5	3	4	3/16/2016	3/21/2016	50%	0	3/18/2016		Started	3/22/2016
Pouring	Concrete	Structure	Area A	Level 5	3	1	3/22/2016	3/22/2016	100%	1	3/22/2016	3/22/2016	Finished	3/22/2016
Drop	DeckDrop	Structure	Area D	Level 4	3	5	3/16/2016	3/22/2016	100%	6	3/16/2016	3/22/2016	Finished	3/22/2016
Inst.	Decking	Structure	Area C	Level 5	3	5	3/16/2016	3/22/2016	100%	8	3/14/2016	3/22/2016	Started	3/22/2016
Inst.	Decking	Structure	Area D	Level 5	3	5	3/23/2016	3/29/2016	20%	0	3/21/2016		Started	3/22/2016
Inst.	Reinforcing	Structure	Area C	Level 5	3	4	3/23/2016	3/28/2016	20%	0	3/21/2016		Started	3/22/2016
Inst.	Columns	Structure	Area B	Level 5	3	4	3/16/2016	3/21/2016	75%	0	3/18/2016		Started	3/23/2016
Drop	DeckDrop	Structure	Area B	Level 5	3	5	3/23/2016	3/29/2016	0%	0	3/23/2016		Started	3/23/2016
Inst.	Decking	Structure	Area D	Level 5	3	5	3/23/2016	3/29/2016	40%	0	3/21/2016		Started	3/23/2016
Inst.	Reinforcing	Structure	Area C	Level 5	3	4	3/23/2016	3/28/2016	40%	0	3/21/2016		Started	3/23/2016
Inst.	Columns	Structure	Area B	Level 5	3	4	3/16/2016	3/21/2016	100%	5	3/18/2016	3/24/2016	Finished	3/24/2016

Appendix E: Crew Management Screen

Appendix F: Crew Screen