Leadership in Sustainable Infrastructure

Leadership en Infrastructures Durables

Vancouver, Canada

May 31 – June 3, 2017/ *Mai 31 – Juin 3, 2017*

IMPROVED DATA STORAGE FOR BETTER SAFETY ANALYSIS AND DECISION MAKING IN LARGE CONSTRUCTION MANAGEMENT FIRMS

Coffland, Megan M.[1,4,5], Kim, Amy A.[2], Sadatsafavi, Hessam[3] and Uber, Mindy M.[4]

[1,2] University of Washington, United States

[3] Cornell University, United States

[4] Skanska USA Inc., United States

[5] megan.coffland@skanska.com

Abstract: Skanska USA Building Inc. has the goal of utilizing past safety data to reduce the number of fatalities, injuries, illnesses and near misses on construction sites in keeping with their Injury Free Environment (IFE) initiative. To this end, following any safety incident, data is collected and stored subsequent to a full incident review that exceeds OSHA reporting requirements. This safety data has been evaluated repeatedly by the company and any valuable information has been mined; to further utilize the safety records, they need to be compared against other project data. This data, along with information on the use of Building Information Modeling on projects, and accounting from the Seattle, WA office from 2006 thru April of 2016 has been made available for investigation. This data represents 556 projects with 1953 separate data points, each representing a total of up to 20 separate pieces of information. This data was kept in separate locations and spread over multiple years, often with overlapping items. The objective of this study is to discuss the lessons learned to move to a centralized system to support data-driven business operations. More specifically, the study evaluates the data storage, file structures and limited integration for ways to improve data mining for safety assessment. Centralized data will lower the cost and time threshold for future data analysis. The results will be invaluable to not just to this firm, but to construction management companies across all locals in guiding data collection and storage to facilitate future analysis.

1 INTRODUCTION

Initial gathering of safety-related data for statistical analysis from a large multinational construction company found areas for improvement in data management. While a great deal of very valuable data is being gathered on projects, this data is stored in disparate locations. On a practical level, this means that the data is not being fully utilized. The hurdles of tracking down and aggregating data ultimately overshadow the time needed for investigation and ultimately halt any deeper analytics before they initiate. This paper focuses on one of the challenges found in pro-active construction safety monitoring and informed safety investment which is organizing, gathering, and processing construction safety-related data for any effective decision-making.

Previous Research

There is plenty of research on construction safety. However, there is a limited discussion on best practices for safety-related data architecture for large construction management firms that require analyzing vast amounts of structured data. And while new technology such as Building Information Modeling (BIM), mobile devices, and other software are found to reduce job site risks, these studies typically make these assessments based on safety data collected from generally one of three categories: broad utilization of Occupational Safety and Health Administration (OSHA) reporting data, individual site case study, or qualitative information typically based on expert interviews (Weiss 2016).

One of the more intensive questionnaires was collected from a total of nine construction sites in Hong Kong over a limited time frame (Lu, 2016). A similar questionnaire based approach was used in Singapore in 2014 where safety culture and safety incentives were ranked via survey on a scale of 1 to 5 and those responses were analysed to determine the relative impact of safety incentive programs and safety culture (Feng 2014). Some of the most well regarded studies in determining weighted impact of safety best practices both are based on multi-stage surveys and interviews of construction experts (Hallowell 2009). Some studies have noted the impact of Leadership in Energy & Environmental Design (LEED) on safety and made recommendations for improvement, again based on expert interviews (Dewlaney 2011). Hinze took this a step further and compared the recordable injury rate across 57 projects (Hinze 2013). However, this study only took into account OSHA level recordable which does not cover all incidents occurring on a construction site. One of the most comprehensive uses of government data came from the University of Lisbon in 2015, which created a cost-based risk analysis using publicly available OSHA reporting data from construction projects (Sousa 2015). One of the first studies to empirically establish relationships between specific variables and OSHA recordables was Jon Wanberg (Wanberg 2013). However, only a limited number of variables and a limited number of active projects were considered. More recently an intensive investigation into case studies related to the impact of design on safety found a correlation between design and fatalities (Malekitabar 2015). However, only data on fatalities and significant injuries was collected and used in the analysis. Then there were some studies that fell outside the bounds of the major categories and instead focused entirely on systemic literature reviews (Mullan 2015).

1.1 Goal

The objective of this paper is to offer lessons learned from a recent safety data gathering and information processing effort. This paper provides practical directions on how to better organize and manage project data for large construction management companies. Furthermore, this paper discusses how to build a stronger and more successful industry-academia partnership in order to further leverage the data to truly improve the safety of the construction workers.

2 DATA

2.1 Current Status of Data Collection within One Regional Office of a Global Construction Company

2.1.1 Safety Data Permissions

Safety data is available in an online database to all permanent company employees. The broad, nearly universal access to safety data is due to the companies IFE policy and the expectation that safety is a core component to every employee’s job. This database is stored in an off-the-shelf, cloud based system that was initially chosen by the US portion of this international business for the transition from paper to electronic data storage. While employees have access to the data, permission must be granted for that data to be used for research purposes. That permission needs to come from both the Environmental Health and Safety (EHS) Director and the Vice President of Operations at each regional office targeted for analysis.

This permission requirement explains the lack of studies that discuss the more granular level of safety analysis that are available within the academic and research community. In order to maintain employee privacy and to protect proprietary information, private companies will have differing and in most cases limited data sharing policies. Reluctance to share the data can come from multiple sources. Privacy issue would be one of those (Institute of Medicine 2013). Researchers in other fields have identified poor data organization (UCL 2015) such as institutional barriers to data access as another reason for limited data sharing (Beniston 2012). Finally, aligning the incentives would be important as academic researchers’ are primary interested may be in academic publications while the organization would be mainly interested in quantifiable and profitable process or product improvement.

2.1.2 Accessing Safety Data

This safety data is accessible through the secure company site (http://myskanskausb). From there, one can navigate to the “Our Company” tab and down to “Environmental Health and Safety”. On the right side of the screen is an area titled “EHS Application Links”; the first line item titled “Monthly Incident Tracking & Reporting System” should be followed. The web interface allows for reports to be generated and downloaded. The easiest option is to start with the “Incident Listing” option. These reports can then be sorted and filtered by Office, Company, Year, Classification, Incident Type, and Month. The system also includes a Find option for searching for various and specific terms.

Figure 1: Monthly Incident Track & Reporting System Interface

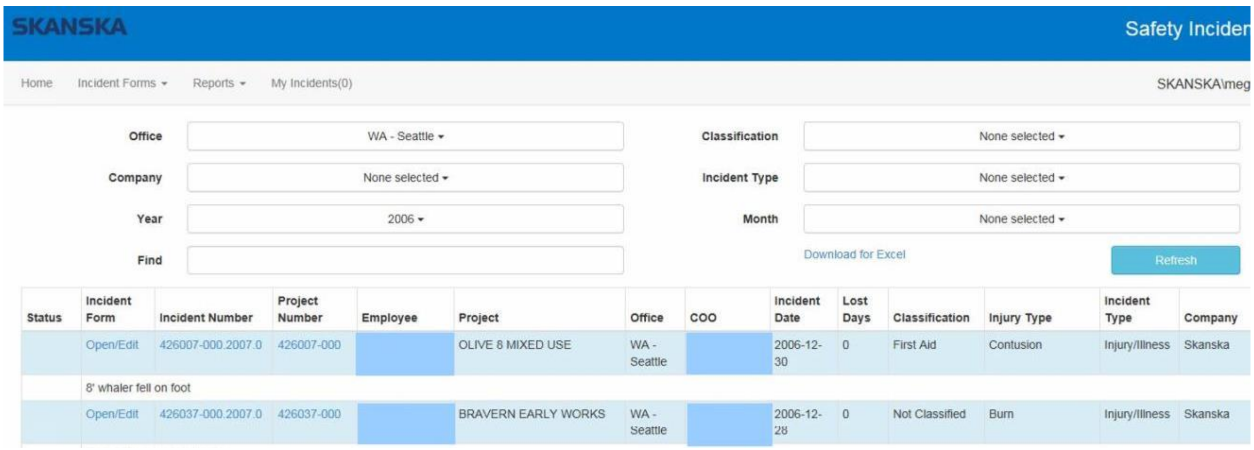


*Figure has been redacted for privacy purposes.

From the usability standpoint, the web-interface is considered user-friendly for most employees. However, anecdotally, employees do not access the web-interface on regular basis to conduct data-mining efforts as more and more information is being collected on even a single project.

2.1.3 Search Parameters

This study sorted for projects from the Seattle offices for each year from 2006-2016. This window was chosen for a number of reasons. Year 2006 was the first year of electronic record keeping with the above mentioned cloud-based data management system and 2016 was the last year to use this particular database system. Additionally, going back to 2006 gave access to safety data prior to broad implementation of BIM. Finally, this span of years provided a substantial sample size and variability in projects for a number of exploratory analyses.

2.1.4 Report Parameters and Nested Values

When developing the safety data, the following report parameters are required and included universally in the exported data: Date, Project, Incident Type, Injury Type, Days Lost, and Company. Detailed information regarding the Project is nested in a different department and access to Project data requires permission from the Director of Accounting. So, in this case, the data on the Project did not auto-populate but had to be manually retrieved and manipulated to be embedded in a single spreadsheet format for further data analysis. The details of the Project data (Project Number, Project Name, Construction Start Date, Final Closeout Date, and Contract Value) was stored in an excel spreadsheet format for individual years.

2.1.5 First Data Manipulation

Upon further investigation, details of the Project data revealed that the project data changes over time and duplications were found. Construction Start Date, Final Closeout Date and Contract Value tend to change as the project develops and reaches near completion. To ensure only the most recent and therefore accurate information be used in the analysis the following steps were taken:

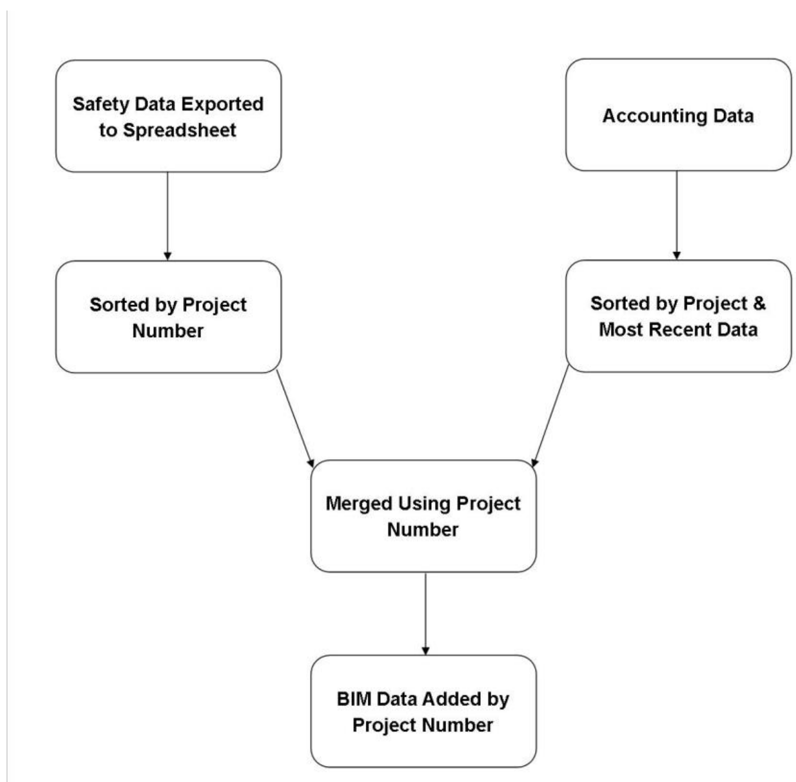
1. 2006 projects were searched using the unique project number and if found in 2007, the 2007 values (including project value, project start and project closeout date) took precedence and the 2006 values were deleted.
2. If no duplicate was found, a new line item was created with the, now verified as final, years data.
3. When this was completed for all items listed in 2007, the 2007 file was closed and the, now updated, 2006 file was saved under a compiled file name.
4. This process was then continued for each successive year until 2015 was complete and all 556 projects were listed with the most current project data available.

The final combined file contains only the final or most recent project data. In this way, a complete project list for the Seattle office was developed.

2.1.6 Adding BIM Data

As previous studies have indicated that the use of smart modeling reduces job site risks, the research team initiated an effort to combine project data with information regarding the use of BIM/Virtual Design/smart modeling. As there was no centralized list of projects that indicated the use of BIM, the Director of VDC for the Seattle and Portland offices manually provided the information to be included in the master spreadsheet. The following information was created: Use of BIM (Yes/No), Market Sector, Part or Full Time Modeller, Phase of BIM Used (Preconstruction, Construction, All), Responsible (Individual VDC Manager assigned to project).

Figure 2: Data Flowchart



2.1.7 Second Data Manipulation

The second data manipulation entailed combining the BIM data to the Project data that included the Safety data. All projects appear on the master project list. However, if the project had no incidents reported then the incident field would receive a N/A. A “Not Given” on the other hand meant that that particular project appeared in the safety report download or record but was missing the information provided by the construction firm.

3 PRELIMINARY ANALYSIS

3.1 Data Variables

While the data gathering, extraction, and manipulation finally led to an organized and searchable database, the most unique set of variables in this database is the classification of incidents and use of BIM. The classification of incidents recorded in the database exceeds the OSHA requirements and provides an opportunity to investigate the relationship between project specific characteristics on a broader level of safety performance. Table 1 shows the classification of incidents with detailed descriptions.

Table 1: Incident Classification and Description

Classification*	OSHA Equivalent	Description
First Aid	OSHA First Aid	Same definition as OSHA
Lost Work	OSHA Recordable	OSHA recordable that results in missed time from work for treatment or recovery
Non-Industrial	OSHA Recordable or First Aid	OSHA first aid or recordable that did not occur on a construction site (i.e. at a district office)
Not Classified	Not OSHA Required	No additional information on classification included at time of reporting
Not Recordable	Not OSHA Required	Near miss or injury/illness reported to Contractor that did not meet minimum requirements for first aid
Restricted Work	OSHA Recordable	OSHA recordable that results in physician recommendation of modified duty to return to work. This may include office work or restrictions on lifting, movement or hours.
N/A	Nothing to report for project	This is used for projects that were included, but had nothing reported

*Contract manager’s classification.

Detailed information on the use of BIM will allow testing relationships between the uses of BIM on construction safety in the future. If proven effective, then the analysis would aid in building confidence and justification for the use and further investment in BIM utilization. The BIM project data was further broken down for various market sections as shown in Table 2. This allowed further assessing whether certain sector was more prone to using BIM and if that had any relationship to the number of severity of incidents.

Table 2: BIM Project Market Sector Explanations

Sector	Description
Aviation	Including manufacturing, hangers and airports.
Biopharm	Biopharmaceutical, including labs.
Chemical	Including treatment plants.
Corporate/Commercial	Commercial not associated with aviation, retail, or health care.
Cultural/Religious	Including museums, cultural centers, and performing arts venues.
Government	Funded by federal, state or local governments, not including schools.
Health Care	Includes hospitals, clinics, sustaining work, and research centers.
Higher Education	Includes both publicly and privately funded.
K-12	Includes both publicly and privately funded.
Residential	Includes condos and apartments.
Retail	Commercial retail fronts.
Sports Arena/Stadium	Includes arenas, repairs, and associated support structures.
Not Given	Information was not provided, missing or not obvious.

4 RECOMMENDATIONS

There are multiple lessons learned from this initiative that can be translated to both the private industry and the academic community. For the private industry, there needs to be enhanced data metrics and collection methods to streamline the analysis. Academic partners can provide valuable information by providing the scientific analysis. For this to occur, a good understanding of the current process is essential. By conducting the data collection and organization the gap was identified through this research.

Next, a policy-driven changes need to be implemented within the organization based on the findings for proper project documentation and reporting. This will help improve the safety performance in the long-run.

4.1 Enhanced Data Metrics and Collection

Preliminary results indicate the following items may be helpful in further analysis. Number of people on site or man hours may help evaluate overall complexity or likelihood of trade stacking having impact on injury and incident rates. Similarly a listing of the number of other subcontractors or trades working on site that day with additional breakdown of how many other trades were working in the same zone could be beneficial. Alternatives to both of these include total man hours reported to accounting on public works jobs, and total number of subcontracts and sub tier subcontractors. Metrics like these help identify other exogenous or environmental factors that may be significant yet hard to prove.

In addition, safety reports are recommended to extend to include the following: location of incident, time of day, and time since beginning of shift and/or last break. All of this information is anticipated to help pinpoint and evaluate current “common knowledge or beliefs” on incident rates. These include the idea that longer hours lead to increased incidents.

Other signifiers may include whether or not the work being performed was rework or out of sequence work. This is to help determine whether this leads to lesser planning when evaluating the work.

All of these metrics should be globally selected and evaluated in context to determine if they provide meaningful information. The academic partner has the ability to analyze various metrics and to assess their relative importance based on industry expert's feedback using various research methods and techniques. Furthermore, these analyses could be conducted on a regular basis to identify trends and improvements requiring continued partnerships.

4.2 Project Reporting and Data Integration

It is recommended that part of project closeout includes a report out on consolidated project data, including, but not limited to: all safety incidents, project data, BIM data, and even green construction data such as Leadership in Energy & Environmental Design (LEED) data. As LEED provides an improved perspective on construction workers' safety, it may be a vital that this information gets assessed more comprehensively. The current system does not allow for easy manipulation of data to do this kind of assessment.

Table 3: Recommended Closeout Report Data

Sector	Description
Safety	All currently collected data with the addition of man hours or number of people on site the day of incident.
Dates	Project start data, substantial completion date, and any delays or accelerations to project schedule.
Financials	Overall project value, revenue and actual vs. projected on both.
Sector	Per table 2 and owner and owners representative type.
Delivery Method	Design-Build, GC/CM, GMP, etc.
BIM	Extent of BIM use on project, including any new tools used.
Green Construction	List of any certifications including the year/version of certification achieved and specific credits achieved.

4.3 Project Reporting

It is further recommended that this data be stored in a searchable data base equivalent to the current safety database used by Skanska. The easy access of data for analytics is vital to future research. Storage in a central location would allow all time used on analytics to be associated with data analysis instead of tracking down, sorting and combining data. The data architecture should be carefully considered and designed for the necessary analysis.

Acknowledgements

The Seattle and Portland regional offices of Skanska USA Inc. provided all data and financial sponsorship for this research. Special thanks to Chris Toher, Dave Harrison, Lew Guerrette, Greg Smith, and Chip Smith for aid in navigating permissions and providing data. Additional support was provided by the University of Washington Department of Civil and Environmental Engineering and Capital Planning and Development.

References

Beniston, M., Stoffel, M., Harding, R., Kernan, M., Ludwig, R., Moors, E., Samuels, P. and Tockner, K. 2012. Obstacles to Data Access for Research Related to Climate and Water: Implications for Science and EU Policy-Making. *Environmental Science & Policy*, **17**: 41-48.

Dewlaney, K. and Hallowell M. 2012. Prevention Through Design and Construction Safety Management Strategies for High Performance Sustainable Building Construction. *Construction Management & Economics*, **30**(2): 932-943.

Feng, Y., Teo, E., Ling, F. and Low, S. 2013. Exploring the Interactive Effects of Safety Investments, Safety Culture and Project Hazard on Safety Performance: An Empirical Analysis. *International Journal of Project Management,* **32**(6): 932-943.

Hallowell, M. and Gambatese, J. 2009. Construction Safety Risk Mitigation. *Journal of Construction Engineering and Management,***135**(12): 1316-1323.

Hallowell, M., Hinze, J., Baud, K. and Wehle A. 2013. Proactive Construction Safety Control: Measuring, Monitoring, and Reporting to Safety Leading Indicators. *Journal of Construction Engineering and Management*, **139**(10).

Hinze, J., Hallowell, M. and Baud, K. 2013. Construction-Safety Best Practices and Relationships to Safety Performance. *Journal of Construction Engineering and Management*, **139**(10).

Institute of Medicine (US). 2013. “Sharing Clinical Research Data: Workshop Summary 3, Barriers to Data Sharing.” NCBI.gov. https://www.ncbi.nlm.nih.gov/books/NBK137824/ (accessed February 14, 2017).

Lu, M., Cheung, C., Li, H. and Hsu, S. 2016. Understanding the Relationship Between Safety Investment and Safety Performance of Construction Projects Through Agent-Based Modeling. *Accident Analysis & Prevention, **94**: 8-17.*

Malekitabar, H., Ardeshir, A., Sebt, M., and Stouffs, R. 2015. Construction Safety Risk Drivers: A BIM Approach. *Safety Science*, **82**: 445-455.

Mullan, B., Smith, L., Sainsbury, K., Allom, V., Paterson, H. and Lopez, A. 2015. Active Behaviour Change Safety Interventions in the Construction Industry: A systemic Review. *Safety Science*, **79**: 184-194.

University College London. 2015. “Lack of Government Data-Sharing Hindering Social Mobility Research.” UCL. https://www.ucl.ac.uk/news/news-articles/0315/200315-social-mobility-data-sharing (accessed February 14, 2017).

Wanberg, J., Harper, C., Hallowell, M. and Rajendran, S. 2013. Relationship Between Construction Safety and Quality Performance. *Journal of Construction Engineering and Management*, **139**(10).

Weiss, J. 2016. “4 Ways ERP Reduces Risk on the Job Site.” CMiC Global. http://cmicglobal.com/erp- reduces-risk-on-the-job-site/ (accessed February 14, 2017).