Cite the Paper: Flynn, M., RazaviAlavi, S., Germi Seresht, N. (2021) Real-Time Analytics and Visualisation of Construction Safety Data Generated Through a Cloud-Based Application. The IDoBE International Conference 2021, London, UK.

REAL-TIME ANALYTICS AND VISUALISATION OF CONSTRUCTION SAFETY DATA GENERATED THROUGH A CLOUD-BASED APPLICATION

M. Flynn[1], S. RazaviAlavi[2] and N. Gerami Seresht[3]

[1, 2, 3] *Faculty of Engineering and Environment, Department of Mechanical and Construction Engineering, Northumbria University, Sutherland Building, Newcastle Upon Tyne, NE1 8ST, UK*

Within the construction industry, Health & Safety (H&S) is one of the leading concerns. Although a significant amount of H&S data is collected, they are not being used effectively for decision making. One of the main issues with using H&S data is that they are generated in an unstructured format, requiring time and effort to transform the data into useful information. Thus, this research aims to develop a cloud-based application to collect structured H&S field data from live construction sites. A prototype was developed using Microsoft Power Platform and Microsoft SharePoint cloud-based storage to demonstrate the practicality of the application. In this application, the structured data is automatically analysed against the required ISO 45001 H&S outputs. To enable informed and decisive decisions, real-time data analytics dashboards are created by Microsoft Power BI. The resultant analytics provides valuable insight into the safety climate of live construction sites.

Keywords: Cloud-based Application, Real-time Analytics, Safety Climate, Visualisation

INTRODUCTION

Health & Safety continues to be one of the leading concerns within the construction industry. Each year countless lives are lost, due to inadequate safety awareness and protocols. Jin et al., (2019) stated that the construction industry is considered one of the most unsafe industries. Owing to the nature of the construction industry, it is often challenging to gain accurate insights into the safety of construction sites. Additionally, it is harder to obtain real-time structured data that could identify safety trends and allow for real-time decision-making. The primary focus of the current research is to investigate data analytics in construction safety and develop a structured data pipeline of H&S data generated by a custom-built application. Furthermore, machine learning and its principals are reviewed as potential add-ons to the custom application. Ghodrati *et al.* (2018) mention that the introduction of ML models to analyse safety data could be of assistance to the authorities to design proactive strategy plans. Accordingly, this project will examine the development of custom H&S auditing application (using MS Power Apps) within a building contractor's organisation and how the extracted data links back to the data storage platform. From the storage platform, machine learning operations will be adopted to identify trends within the data in an attempt to identify safety climates on individual construction sites. All real-time analytics generated by the custom application will be presented on MS Power BI dashboards to provide the management with the ability to make informed decisions regarding health & safety in real-time. The central question under investigation is, "Does the introduction of an end-to-end health and safety construction data pipeline improve the identification of site safety climates and thus enable predictions of safety incidents?" The aspect of machine learning pertinent to this research will mainly target the ISO 45001 compliance requirements. These include risk assessment, performance monitoring and response to emergency situations. The objective of this research is to analyse the construction data generated by the custom health & safety application using data analytics methods. Furthermore, it would provide real-time updates

[1] mark.flynn@northumbria.ac.uk [2] reza.alavi@northumbria.ac.uk [3] nima.gerami@northumbria.ac.uk

and a dashboard to allow health & safety operatives in the field to make informed decisions. This research project will develop a custom health & safety auditing application on the MS power apps platform, the data generated by the custom health & safety application will be stored on a SharePoint platform. From this platform, the data will be loaded onto the Azure data bricks platform, where custom Python script machine learning operations can optimise and analyse the data against key performance requirements. Following the analysis of optimised data, the key performance indicators and safety trends would be live linked for visualisation within the Power BI dashboards. Sinelnikov, Inouye and Kerper (2015) mention that the use of Power BI dashboards is to turn data into insights. Therefore, the current research project will strive to provide beneficial results that would enable other departments, within the main contractor's organisation, to adopt real-time analytics and machine learning to streamline their workflows. The structure of the research project entails conducting literature review, an interview with an experienced health & safety manager and the development of a functioning H&S application to provide real-time analytics. However, this research project has several limitations—the leading one being that advanced machine learning models are difficult to design and implement. As a result of this limitation, the application primarily focuses on gathering structured data, which could be presented in real-time dashboards. Although the possibility of introducing machine learning capabilities is discussed, it is not implemented within the application. This leaves room for further studies and research to expand on this topic. From the literature review, it has been discovered that a number of contributors have developed similar applications to offer better predictive capabilities from construction data. Tixier et al. (2016b) shows how models predicting safety outcomes will ensure accurate diagnostics of the safety risks. However, the existing contributors have mainly focused on extracting keywords from existing unstructured construction data via third-party software. The custom application developed within this research project initially outlined the significance of data attributes to ensure that all construction data is gathered in a convenient way to allow easy analysis. As of now, there are numerous data gathering software platforms within the construction industry, none of which are designed to interpret data in a structured manner that would provide insightful output. At best, most of the software platforms store semi-structured observations, which remain unused in the platform and require skilled managers to review, analyse and report their findings. Consequently, the current research investigates the benefits of introducing a custom application that can analyse H&S construction data in real-time and present the trends to a cloud-based dashboard. Furthermore, this project highlights the possibility of adopting machine learning capabilities where existing H&S records can be analysed to provide predictive outputs and identify safety climates.

SAFETY IN THE CONSTRUCTION INDUSTRY

Health & Safety persists as a significant concern within the construction industry. The safety of construction workers is considered one of the main issues faced by the industry in countries worldwide (Patel and Jha, 2015). It is estimated that around 60,000 people die per annum, while thousands sustain serious injuries or contact occupational injuries while working on construction sites (Cheng et al., 2012). Most main building contractors, ensure continued Occupational Health & Safety (OHS) of their staff as they consider it as utmost importance. On the contrary, some building contractors undermine safety management unless it is considered a client requirement or strongly enforced by the law. Oftentimes, this is a by-product of the efforts of the building contractor to save money in a competitive bidding scenario (Ayhan and Tokdemir, 2019). Although the government intervenes through health & safety regulations, and policies, the construction industry continues to be a high-risk sector with a high rate of injury and fatality worldwide (Ghodrati et al., 2018). Hinze and Gambatese (2003) highlighted that major factors affecting the health and safety performance of subcontractors on large construction projects are down to the effects of the main building contractors to promote site safety. Typically, the culture surrounding organisational safety has been outlined as the fundamental values, beliefs and principles that the workforce perceive as active within the organisation (Barbaranelli, Petitta and Probst, 2015). Mohammadi et al. (2018) defined safety culture as the sum of the corporate outlook, safety values and its safety behaviour. A snapshot of the safety culture within an organisation, at a given time, is referred to as the safety climate of the organisation. The safety climate reflects the employee's perceptions and attitudes towards the safety practices being enforced. Patel and Jha (2015) describe the safety climate as one that is dynamic in nature. By implementing safety climate measures in a building contractor's organisation, the management team will possess vital data that can help effectively manage their safety climate and even provide forecasted warnings against potential health and safety incidents. (Flatau Harrison et al., 2018). Recently there have been several studies that outline the benefits of developing a safety climate framework and linking it to safety performances (Jin et al., 2019). Some studies demonstrate an increase in safety investment, such as safety information and communication technologies (ICT). Consequently, they have seen an improvement in safety performance, more specifically a decrease in the Accident Frequency Ratios (AFR) (Feng et al., 2014). Although the construction industry has seen a significant rise in the adoption of information management technologies due to incentives, such as the adoption of BIM level 2 standards, the track record within the health and safety sector has not seen a similar positive change (Aguilar and Hewage, 2013). Instead, several main building contractors continue to record injury-related events and near misses in unstructured textual reports. Since these reports are unstructured, most of the valuable knowledge and experience stored within them remain unexploited. Furthermore, it means that the safety management team must manually analyse the textual data reports to remove the possibility of automated analysis through information technologies (Tixier et al., 2016b). The current research project investigates the adoption of a health and safety initiative within a building contractor's organisation which is to be implemented using a custom-built health and safety application. This application would gather structured safety data for automated analysis of key safety performance indicators.

DATA ANALYTICS AND CUSTOM APPLICATIONS

Due to the ever-expanding pipeline of construction project data, the realm of data analytics and subsequent methods used in this particular field have become increasingly relevant within the construction industry. The spike of data analytics within the construction industry remains at an early stage and falls behind when it comes to adopting technologies within other industries; however, real adoption is beginning to occur (Bilal et al., 2016). As building organisations investigate the possibility of data analytics, they have come to the realisation that most of the technologies available in the market, do not provide the required level of detail for full analysis; therefore, they would need to enhance their internal applications to enable compatibility with specific business challenges. This led to the adoption of development platforms, such as Microsoft Power Apps, where customer do not need to be concerned about the IT infrastructure, and can simply focus on the development of custom business solutions which create business value (Woodhead, Stephenson and Morrey, 2018). Nowadays, organisations can store data from structured projects, gathered by custom business solutions in online data warehouses, which are automatically linked to the custom application. Within the data warehouse, the organisation can organise the structure of project data so that it could be easily analysed against key performance indicators. In recent years, the development of artificial intelligence (AI) and its sub-set machine learning (ML) have resulted in the creation of techniques that are useful in automating activities to discover patterns in databases (Tinoco, Gomes Correia and Cortez, 2011). Within the construction industry, one of the fastest growing applications of machine learning is data mining. Data mining employs machine learning techniques to analyse raw project data and extract high-level information pertinent to key performance indicators for decision making (Tinoco, Gomes Correia and Cortez, 2011). Within the architecture of custom-built applications, it helps to structure the project data by utilising an attribute-based framework to increase the efficiency of the data mining process. An attribute-based framework allows the organisation to capture and encode project data in a standardised format that can be integrated with additional ML models. Tixier et al. (2017) examines benefits of attribute-based frameworks by demonstrating automated identification of safety clashes from attributes extracted from a large set of data in injury reports. As mentioned earlier, most building contractors continue to record their health and safety incidents and report them in unstructured textual reports. As a result, most of the AI and ML applications require natural language processing (NLP) technologies to extract the structured data attributes from the existing textual reports. NLP technologies focus on comprehending and analysing human-produced texts (Tixier et al., 2016b). By adopting an attribute-based framework, within business solution applications, the requirement for NLP technologies will be removed and instead ensure that the organisations health and safety data can be stored and structured for easy analyses.

MACHINE LEARNING & AI CONCEPTS AND WORKFLOWS

Machine learning is considered as the study of computational methods to automate the process of knowledge acquisition from given examples (Indranil Bose, 2001). At present, industries such as retail and manufacturing have utilised machine learning (ML) algorithms to provide greater insight into production trends and to deliver greater product satisfaction to the consumer. The growing success following the adoption of ML, within these industries, has lured other sectors of the economy, such as the construction industry into the adoption of intelligence algorithms to their enterprise software (Bilal and Oyedele, 2020). Some research within the field of ML has primarily focused on cost estimations for competitive bid situations. Custom applications have been developed using neural networks to estimate attributes, such as the number of competitors or the contractor's estimated cost (Tixier et al., 2016a). Recent research has investigated the possibility of using machine learning models to enable the identification of work teams that are at high risk of serious accidents in the field (Choi et al., 2020). It is anticipated that the adoption of construction simulation tools (CST's) will lower the annual rate of accidents and deaths in the construction industry and facilitate greater decision making throughout the project's lifecycle (Bilal and Oyedele, 2020). Some of the major areas of research within machine learning include Rule Induction, Neural Networks, Case-based Reasoning, Genetic Algorithms and Inductive Logic Programming (May Shayboun, 2020). The current research project mainly focuses on the ML areas of Case-Based Reasoning (CBR) and Artificial Neural Networks (ANN) for analyses of project data. Artificial Neural Networks (ANNs) have emerged as the most popular machine learning algorithm to perform risk assessments on structured data (Hegde and Rokseth, 2020). The most commonly used ANN algorithm is the multi-layer perception (MLP) algorithm which usually consists of three layers: the input layer, hidden layer and the output layer (Bilal et al., 2016). A typical neural network requires a list of structure attribute input nodes that are linked to specific key performance indicator output nodes. The connection is set up through a set of hidden nodes that build up the multi-layered network and allow for data analyses to occur (Indranil Bose, 2001). The second ML method chosen for this research paper is the Case-Based Reasoning (CBR) approach. This method uses code safety examples, including the problem description and solution, stored on a database. The CBR system will, in turn, utilise the new case's coded attributes to find safety examples in the same classification stored in the database (Indranil Bose, 2001). The CBR approach relies heavily on producing a structured example database, as the current problems can only be resolved by referring to knowledge and experiences accumulated from previous instances (Su et al., 2019). Limited studies exist that attempt to launch the CBR theory into construction safety and risk management (Su et al., 2019). This research project develops a custom business application that will incorporate the required structured attribute-based framework to enable the functionalities of Artificial Neural Networks (ANN's) and Case-Based Reasoning (CBR).

DATA VISUALISATION

Within the construction industry, data visualisation has significantly increased over the last few years given the adoption of sophisticated project management technologies. Due to the extensive amount of construction project data being produced daily, it is crucial for building organisations to utilise the project data in an easily readable format. Data visualization plays an effective role by enabling the analyst to scan the raw data to identify patterns, detect outliers, and develop hypotheses, which are then verified (Indranil Bose, 2001). Majority of the construction project management technologies provide some sort of data visualisation functionality. However, most of these functions are static, un-customisable and fail to meet the ever-changing challenges faced by the construction industry. This research paper will demonstrate how adopting Power BI application will present structured construction safety data by the custom-built business application. The Power BI dashboards use tables, pivot matrices, and various indicator cards to spotlight a single figure (Becker and Gould, 2019). The cloud-based live Power BI dashboards will be configured to extract key safety performance indicators, such as Accident Frequency Ratio (AFR). Sinelnikov, Inouye and Kerper (2015) demonstrates how communicating key performance indicators to high-level management through visual methods can result in organisations highlighting and adopting corrective or preventive actions quickly and easily. The Power BI dashboards will provide a visual display of real-time safety analytics extracted from site observations and incidents recorded on the safety application. The real-time analytics will attempt to bridge the existing gap between the vast construction databases and the management team (Becker and Gould, 2019). The live pipeline of construction data streaming to the dashboards will assist the management team in making informed decisions, consider a multitude of criteria's and corrective actions, reduce the time needed to make effective decisions, and direct attention to essential elements of a scenario (Esmaeili and Hallowell, 2013). Power BI as an application provides a number of different data analytic functions including sum totals, counts, measures of central tendency, minimums, and maximums (Becker and Gould, 2019). The Power BI application also provides several functions for more advanced calculations and statistical design needs. To meet with the requirements of advanced graphical design, the dashboard portal can be configured to export specific construction data to enable R script codes to manipulate and display complex data sets (Becker and Gould, 2019). For advanced calculations within the Power BI platform, Data Analysis Expressions (DAX) can be utilised. Similarly, structure to excel formulas are customised for dynamic filtering DAX allowing for the manipulation of a specific table or column of construction safety data. By introducing this questionnaire into the new custom application, the safety management team will procure a transparent view of individual workers' perspective of health & safety. Recently, a journal outlined the adoption of Case-Based Reasoning (CBR) technique as a new paradigm of artificial intelligence. The CBR technique solves current problems by referring to knowledge, information and experience gathered from previous instances (Su et al., 2019). The CBR technique uses structured Hazard Status Sets (HSS) and Solution Set (SS) to determine the severity of safety accidents. Furthermore, it relays on codifying past cases with structured hazard and solution sets which could be time consuming. The current research project will further expand on these concepts by devising a new custom application of structure hazard and solution sets. By doing so, the new application will ensure that all health and safety accidents recorded on the site are automatically coded to the correct hazard or solution structure and, in turn, build the required knowledge base to analyse new incidents.

RESEARCH METHODOLOGY

The current project acquired a research design framework which laid-out the procedures for collecting and analysing data. Within this framework, research methods by which the project data may be collected is outlined—these methods include an interview and a custom-built application. Initially, this research project adopted a deductive theory approach by developing a hypothesis surrounding well established knowledge within the area of data analytics and machine learning. Bryman (2016) stated that a deductive theory approach establishes a hypothesis based on well-known knowledge and examines it against empirical quantitative research gathered.

The objective of the research was to investigate the perceptions of data analytics for the advantages of health and safety within a main building contractor organisation. To achieve this, the methodology strategy included several different research methods that relied on both primary and secondary data. The methods for obtaining primary data included a semi-structured interview with a highly experienced health and safety manager, and a custom-built application. While the methods for obtaining secondary data included an in-depth literature review concerning the health and safety and data analytics of the construction industry. From the secondary data, three existing case studies were identified as having key material that would aid in the development of the custom-built application.

RESULTS AND DATA ANALYSIS

Custom built health & safety application

The custom-built health and safety application was created to demonstrate the possibilities of implementing a custom-built business solution that would provide visual data insights. Consequently, this would lead to improved decision making when dealing with certain health and safety challenges faced in a main building contractor's organisation. The primary objective was to develop a live pipeline of safety construction data that would provide insights to the higher-level management to improve decision making. The custom application was designed to record specific requirements set out by the 45001 ISO health and safety standards. These requirements include safety observations, incidents, accident reports, near misses and security breaches. Furthermore, the application was designed to document safety records in a codified structure that would enable hassle-free analyses by machine learning tools. In addition, the application was designed on the Microsoft Power Platform—this platform allows data to be easily linked and synced across different data sources. All safety records documented on the application are automatically synced to a custom HSQE SharePoint storage site where all safety records were stored as SharePoint lists. Based on the input data, custom columns were added to the SharePoint list to extract numerical data. From the HSQE SharePoint storage site, the structured safety data was linked to a data visualisation platform called Microsoft Power BI. Due to Power BI, structured safety data was analysed by custom formulas to extract key safety performance indicators (KPI's). The safety KPI's were, in turn, presented in the form of graph, bar and pie charts within safety dashboards. Lastly, the safety dashboards were embedded within the custom application to allow safety managers to review safety KPI's when they used the application. Additionally, the safety data stored within the SharePoint storage site could be linked to a machine learning platform called Azure Data Bricks. Within this platform advanced analysis such as ANN and CBR can be implemented. Below, Figure 1 provides an overview of the custom application.

Data visualisation using Power BI

The Microsoft Power BI application was selected as it linked with a number of different data sources, including the Microsoft SharePoint storage site. For each of the required safety records stored in the application, a safety dashboard was developed. Figure 2 shows the data pipeline linking from the SharePoint Storage site to the individual safety dashboard. This pipeline was set up to automatically refresh each of the safety dashboards every 30 minutes. This way, the latest real-time analytics would be presented within each of the dashboards, highlighting key insights regarding each of the safety record types. Below, Figure 2 displays the safety observation dashboard which is embedded within the observation record area. As of now, the individual dashboard is set up to highlight the top five requiring safety incidents within each record category. However, further development can be carried out to highlight the major key safety performance indicators.

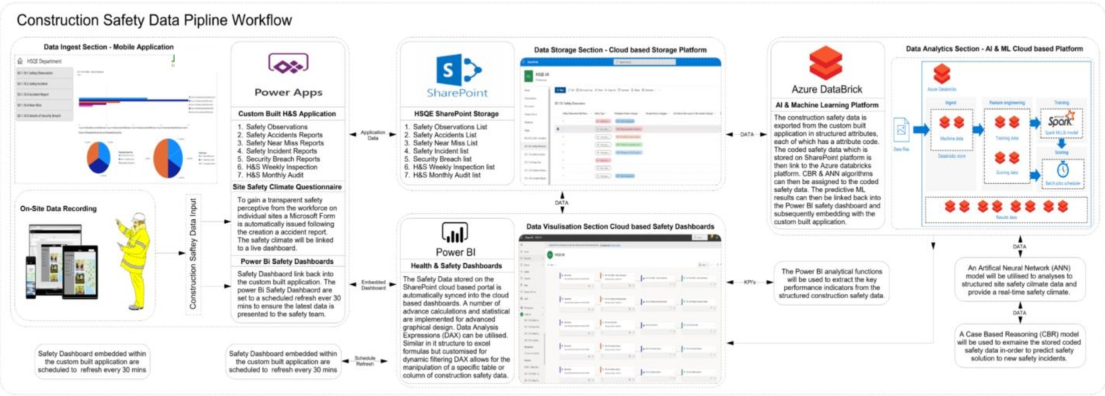

Figure 1: Custom Application Overview

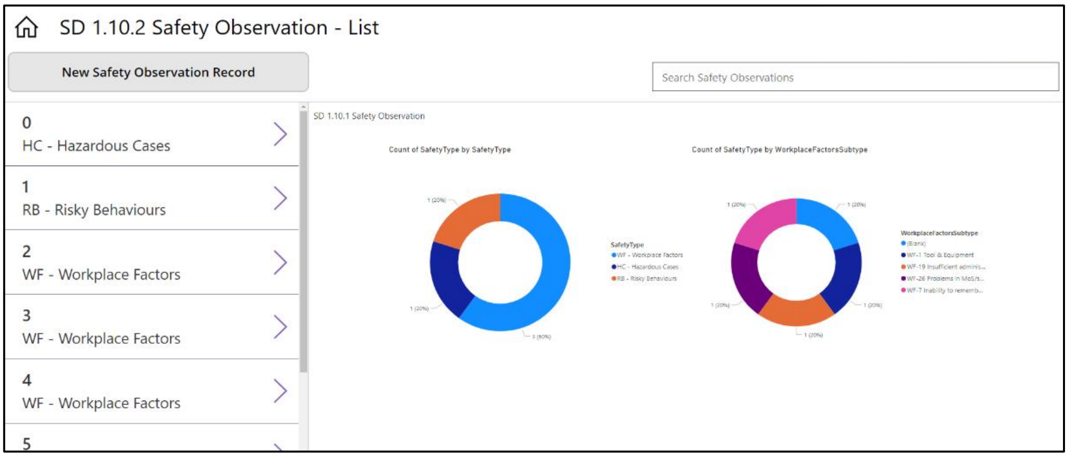

Figure 2: Safety Observation Dashboard

Below, Figure 3 provides an initial outline of the high-level dashboard which is embedded within the outermost level of the application. The proposed dashboard is the first visualisation that the safety managers will see upon entering the application. It will present statistics and safety reports from all safety record categories—present within the application.

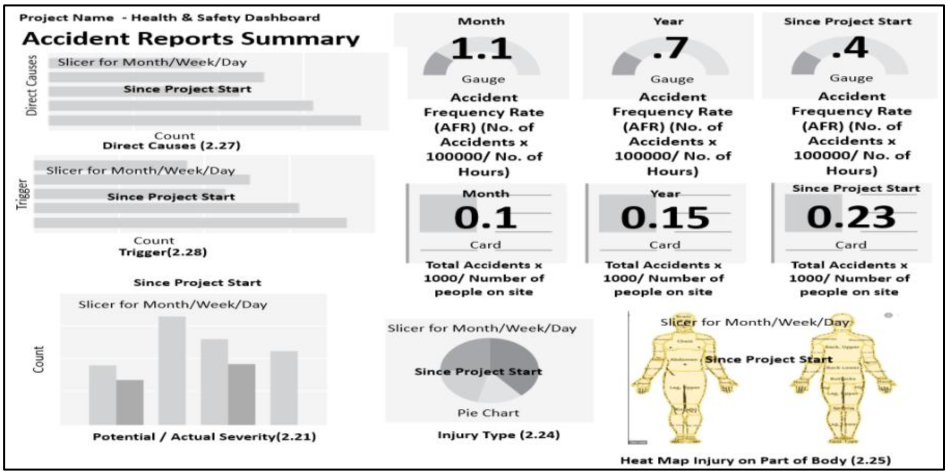

Figure 3: Health & Safety Dashboard Example

Machine learning and artificial intelligence functions

During contemporary times, machine learning and artificial intelligence are one of the most rapidly growing sectors in the field of technology. With significant advancements being made each year, technology has undergone the development of computers that mimic human intelligence and also have the ability to improve automatically through experience (Mitchell, 2015). Thus far, the benefits of these improvements can be seen in the fields of commerce, science and government. By capturing and mining large amounts of data, organisations can improve services and productivity (Mitchell, 2015). As outlined previously, the custom-built application primarily focused on the development of data visualisation. However, some of the techniques require machine learning to be enabled.

Interview results

From the interview with the H&S manger, a number of interesting insights were gained. It was clear from the response given to a question "Has the culture of the construction industry prevented the adoption of data analytics and machine learning tools?" that the H&S manager had faced this problem a number of times. He outlined that some of the older managers struggled to adopt to the introduction of new technologies but with the correct level of upskilling the managers fully embraced the new technologies. He also advised that most project managers and safety advisors would be very interested in the ML functionally to point them in the correct direction when reviewing site safety information. Another main insight extracted from the interview was the strong agreement from the H&S manager around a question "How beneficial would the development of a custom-built application be to your organisation/industry sector?". The H&S manager highlighted that his experience with third party software platforms had not always been great and most of such platforms are designed to fulfil only certain activities while they do not evolve with the changing requirements of the construction industry. The H&S manager agreed that the development of a custom-built application would be hugely beneficial and would remove the reliance on third party software platforms. The final insight that was achieved from the interview related to a question "Do you think the adoption of machine learning tools will result in the reducing of jobs within your organisation/ industry sector?". The H&S manager advised that he fully disagreed with this statement. He outlined that the machine learning tools will never be 100% accurate and that they might miss something that only human would be able to pick up. He could instead see the ML tools and human working together but that the introduction of ML tools would not remove the requirement for managerial jobs within the construction industry. The takeaway from the interview was that the H&S manager could see the benefits of adopting custom-built applications that deal with specific business challenges but the data output from these applications could not be trusted 100% and a human interaction would always be a need to ensure the correct actions are carried out. The main results from the data analysis show that the custom-built application allows for the easy creation of structured safety records. It also shows that further machine learning techniques can be easily implemented due to the structured attribute database linked to the application. The result from the interview shows that the current operatives can recognise the benefits that data analytics can provide to high level management within a main building organisation, but they do not find the introduction of artificial intelligence and machine learning technologies removing the requirement for human managerial jobs. There was also a strong view from the operatives that the proposed machine learning safety predictions could not be fully trusted as they felt human analysis would still be required.

CONCLUSION

For this project, the main research aim was to investigate the benefits of introducing a custom-built safety application for the safety department of a main building contractor's organisation. This was done to help identify safety climates and predict safety trends. By using quantitative and qualitative research analysis, it can be deduced that the introduction of an end-to-end health and safety construction data pipeline would improve the identification of site safety climates. In addition, it would allow for the prediction of safety incidents. By including structured attributes, within the application, the organisation can link structured data into advanced data analytic tools to identify key safety performance indicators and implement a Case-Based Reasoning (CBR) approach to predict safety solutions. Based on the analysis of the interview transcript, several findings emerged. The health and safety manager agreed that the adoption of data analytics and machine learning models would be beneficial. However, the predictive elements of machine learning could not be completely trusted. Furthermore, the H&S manager also mentioned that he believed the introduction of AI & ML would not remove the requirement of safety management jobs within the H&S Sector.

As the research project mainly focused on data visualisation concepts and processes, future research could further build upon the advanced machine learning concepts that were outlined within the project. This would enable organisations to implement safety climate predictions and Case-Based Reasoning functions—organisational key performance safety indicators could also be investigated.

REFERENCES

Aguilar, G. E. and Hewage, K. N. (2013) IT based system for construction safety management and monitoring: C-RTICS2. *Automation in Construction*, **35**, pp. 217-228.

Ayhan, B. U. and Tokdemir, O. B. (2019) Predicting the outcome of construction incidents. *Safety Science*, **113**, pp. 91-104.

Barbaranelli, C., Petitta, L. and Probst, T. M. (2015) Does safety climate predict safety performance in Italy and the USA? Cross-cultural validation of a theoretical model of safety climate. *Accident Analysis & Prevention*, **77**, pp. 35-44.

Becker, L. T. and Gould, E. M. (2019) Microsoft Power BI: Extending Excel to Manipulate, Analyze, and Visualize Diverse Data. *Serials Review*, **45**(3), pp. 184-188.

Bilal, M. and Oyedele, L. O. (2020) Guidelines for applied machine learning in construction industry—A case of profit margins estimation. *Advanced Engineering Informatics*, **43**. 101013.

Bilal, M., Oyedele, L. O., Qadir, J., Munir, K., Ajayi, S. O., Akinade, O. O., Owolabi, H. A., Alaka, H. A. and Pasha, M. (2016) Big Data in the construction industry: A review of present status, opportunities, and future trends. *Advanced Engineering Informatics*, **30**(3), pp. 500-521.

Bryman, A. (2016) Social research methods. 5ed. Oxford: Oxford University Press.

Cheng, C. W., Leu, S. S., Cheng, Y. M., Wu, T. C. and Lin, C. C. (2012) Applying data mining techniques to explore factors contributing to occupational injuries in Taiwan's construction industry. *Accident Analysis & Prevention*, **48**, pp. 214-22.

Choi, J., Gu, B., Chin, S. and Lee, J.-S. (2020) Machine learning predictive model based on national data for fatal accidents of construction workers. *Automation in Construction*, **110**. 102974

Esmaeili, B. and Hallowell, M. (2013) Integration of safety risk data with highway construction schedules, *Construction Management and Economics*, **31**(6), pp. 528-541.

Feng, Y., Teo, E. A. L., Ling, F. Y. Y. and Low, S. P. (2014) 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), pp. 932-943.

Flatau Harrison, H., Griffin, M. A., Gagne, M. and Andrei, D. (2018) Assessing shortened safety climate measures: Simulating a planned missing data design in a field setting, *Safety Science*, **104**, pp. 189-201.

Ghodrati, N., Yiu, T. W., Wilkinson, S. and Shahbazpour, M. (2018) A new approach to predict safety outcomes in the construction industry, *Safety Science*, **109**, pp. 86-94.

Hegde, J. and Rokseth, B. (2020) Applications of machine learning methods for engineering risk assessment – A review, *Safety Science*, **122**.

Hinze, J. and Gambatese, J. (2003) Factors that influence Safety Performance of specialty Contractors. *Journal of Construction Engineering and Management*, 129(2), pp. 159-164.

Indranil Bose, R. K. M. (2001) 'Business data mining a machine learning perspec_2001_Information Management. *Information and Management*, **39**, pp. 211-225.

Jin, R., Zou, P. X. W., Piroozfar, P., Wood, H., Yang, Y., Yan, L. and Han, Y. (2019) A science mapping approach based review of construction safety research. *Safety Science*, **113**, pp. 285-297.

May Shayboun, Kifokeris, K., Koch C. (2020) Machine Learning for Analysis of Occupational Accidents Registration Data. *In*: Scott, L and Neilson, C J (Eds) *Proceedings of the 36th Annual ARCOM Conference, 7-8 September 2020, UK, Association of Researchers in Construction Management*, 485-494.

Mohamed, S. (2002) Safety Climate in Construction Site Environments, *Journal of Construction Engineering and Management*, **128**(5), pp. 375-384.

Mohammadi, A., Tavakolan, M. and Khosravi, Y. (2018) Factors influencing safety performance on construction projects: A review. *Safety Science*, **109**, pp. 382-397.

Patel, D. A. and Jha, K. N. (2015) Neural Network Approach for Safety Climate Prediction, *Journal of Management in Engineering*, **31**(6).

Sinelnikov, S., Inouye, J. and Kerper, S. (2015) Using leading indicators to measure occupational health and safety performance, *Safety Science*, **72**, pp. 240-248.

Su, Y., Yang, S., Liu, K., Hua, K. and Yao, Q. (2019) Developing A Case-Based Reasoning Model for Safety Accident Pre-Control and Decision Making in the Construction Industry. *International journal of environmental research and public health*, **16**(9).

Tinoco, J., Gomes Correia, A. and Cortez, P. (2011) Application of data mining techniques in the estimation of the uniaxial compressive strength of jet grouting columns over time. *Construction and Building Materials*, **25**(3), pp. 1257-1262.

Tixier, A. J. P., Hallowell, M. R., Rajagopalan, B. and Bowman, D. (2016a) Application of machine learning to construction injury prediction. *Automation in Construction*, **69**, pp. 102-114.

Tixier, A. J. P., Hallowell, M. R., Rajagopalan, B. and Bowman, D. (2016b) Automated content analysis for construction safety: A natural language processing system to extract precursors and outcomes from unstructured injury reports. *Automation in Construction*, **62**, pp. 45-56.

Tixier, A. J. P., Hallowell, M. R., Rajagopalan, B. and Bowman, D. (2017) Construction Safety Clash Detection: Identifying Safety Incompatibilities among Fundamental Attributes using Data Mining, *Automation in Construction*, **74**, pp. 39-54.