Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 4, 2025

IMPACT OF REAL-TIME DATA INTEGRATION ON DELAY PREDICTION ACCURACY IN LARGE-SCALE PROJECTS: THE MODERATING EFFECT OF PROJECT COMPLEXITY

Muhammad Nadeem[*1], Fayyaz Ali[2], Ayesha Urooj[3], Nawaid Hasan[4]

[*1,3,6]*Department of Computer Science and Information Technology, Sir Syed University of Engineering and Technology*
[2]*Department of Software Engineering, Sir Syed University of Engineering and Technology*

[*1]munadeem@ssuet.edu.pk, [2]fayyaz.ali@ssuet.edu.pk, [3]aurooj@ssuet.edu.pk, [4]nawaidh@ssuet.edu.pk

[*1]0000-0002-9271-5008, [2]0000-0003-3232-5363, [4]0000-0002-3537-0197

DOI: https://doi.org/10.5281/zenodo.15266507

Keywords
Real-time data integration, Project delay prediction, Project complexity, Predictive modelling, and Delay management strategies. Etc.

Article History
Received on 14 March 2025
Accepted on 14 April 2025
Published on 23 April 2025

Copyright @Author

Corresponding Author: * Muhammad Nadeem

Abstract

Purpose: *In this paper, the effects of real-time data integration on delay prediction accuracy in large-scale projects are examined, taking into consideration the role of project complexity.*

Data and Methodology: *Based on secondary qualitative analysis of four studies examining the live data systems implementation in projects of different sizes, this research synthesises the findings.*

Results: *The study yields the findings that real-time information enhances the prediction of delay in less complex projects while in the complex ones, the effectiveness is reduced unless enhanced by other analytical tools and data processing techniques.*

Conclusion*: This paper reveals that the advantage of real-time data integration is valid only when the project is considered to be complex. It reveals that to leverage the use of data integration in managing delays in projects, the project management strategies have to be customized to the peculiarities of the particular project. This research contributes to the literature by identifying the contexts in which real-time data integration produces the best results and therefore can help to improve project management in numerous industries.*

INTRODUCTION

1.1. Context

In the field of large-scale project management, it is crucial to predict the delays in the project to avoid financial loss and other problems. Traditionally, project implementations have faced several issues which include project overruns and these have been attributed to poor data management and inability to handle project changes on the fly(Saidani et al., 2022). Real-time data integration technologies can be expected to help enhance these forecasts and control the occurrence of these delays. Still, most studies have been conducted on the direct consequences of these technologies with little regard to the level of project complexity. This gap therefore raises to question of how real-time data can be better utilized in different project complexity to improve on the delay prediction. Real-time data integration is crucial in project management because of the advancement in project size and complexity. As cited by Bakici et al. (2021), real-time data enhances project management practices, which in turn increases the capability of anticipating and avoiding delays. This integration

enables the identification of critical path deviations early enough to enable corrective action to be taken. This is supported by Mata et al. (2023) who have also pointed out that real-time data systems increase the effectiveness of decision making hence reducing downtime and optimizing the available resources. Their research establishes the benefits that result from the timely and automated feed of information to all the relevant parties to enable them to respond to any changes in project conditions.

Nevertheless, the part of project complexity in these cases is not as clear. In a recent study, Butler et al. (2020) examined the moderating role of complexity in the context of real-time data systems and their impact on delay prediction; the research showed that when the complexity of the projects is high, the enhancement of real-time data integration in delay prediction is less significant. This implies that there is a curvilinear relationship between data integration and delay prediction results in different project environments. On the other hand, Owolabi et al. (2018) point out that the technology has a rather positive effect, however, the application of more robust methods and techniques and/or project management training for managing complicated projects is needed. Real-time data is seen by both studies as a solution to many traditional delay problems but the application and effectiveness of the solution depends on the inherent complexity of the project.

It is however important to note that although the integration of real-time data has been noted to be useful in the prediction of delays in project management, the literature has not adequately explored the moderating effects of project complexity on these effects. Some research, including by Li et al. (2021), points to the directions of the differences in real-time data integration effectiveness depending on the levels of project sophistication, which may imply the moderating effect. Nonetheless, the nature of this association has not been fully understood. This research gap is important as finding out the relationship between project complexity and the accuracy of delay prediction using actual-time data can help in developing better project management approaches that can be applied to complex projects that may not be manageable using the conventional data integration approaches.

This research fills the gap in the project management literature by analysing the role of project complexity in the relationship between real-time data integration and delay prediction. While the prior studies mostly focus on the advantages of real-time information in general, this research explores the detailed impacts of the complexity of big projects. This is in a bid to establish if different complexity tiers call for separate strategies for data integration in the prediction of delays. Thus, this paper attempts to contribute to the identification of these dynamics in the hope of generating practical recommendations for improving project management practices and tools so that they may be better tailored to match the changing complexities of projects. The objective of this research is three-fold. First, to investigate how real-time data integration affects delay prediction accuracy in large-scale projects. Second, to analyse the moderating effect of project complexity on the efficacy of real-time data systems in predicting project delays. Third, to identify and recommend project management strategies that are tailored to the complexity levels of projects for optimal use of real-time data.

This research is important as it extends the knowledge of conditional advantages of real-time data integration in project management according to the project complexity. It is therefore important to acknowledge these intricacies as they help in enhancing the delay prediction that is essential in large-scale projects that require minimal costs. These results will be beneficial to the project managers as they will be able to identify how they can apply data integration approaches in their projects depending on the circumstances. Hence this increases the efficiency of the decision-making process and increases the effectiveness of the project in different industries.

The research starts with a literature review to define the theoretical foundation for real-time data integration and project intricacy. Subsequently, a secondary qualitative synthesis of certain studies is done to provide supporting findings. The research is aimed at a systematic review of data on the integration of real-time systems and their efficiency on various levels of project complexity. The results from this analysis are then applied to make conclusions on the correlation between real-time data and delay prediction. Last but not least, recommendations for using these findings in enhancing the practices of

project management based on different levels of project complexity are provided.

1.2. Literature Review

1.2.1. Introduction

The literature review section presents research findings and results from past studies related to real-time data integration and the role of project complexity in the accuracy of delay prediction in large-scale projects. It starts with exploring the applicability of real-time data systems to make the methods used in project management more effective. The review then discusses different techniques for identifying delays and their shortcomings in formulating real-time solutions. Last, based on the research model, the moderating influence of project complexity on the impact of real-time data on the enhancement of the prediction accuracy is discussed. The section also identifies gaps in existing research.

1.2.2. Real-Time Data Integration in Project Management

Real-time data integration in project management is the instantaneous merging of project data from diverse sources in order to enable the project manager to readily and timely use the data in project decision-making processes (Fobiri et al., 2022). This process is about linking BIM, PMIS, and IoT smart devices to provide a consistent view of project performance indicators, schedules, and resources. Real-time data integration helps to control the projects improving the reaction on its status, performance indicators and possible risks, so it boosts the accuracy of delay prediction and increases the general project efficiency; however, its efficiency depends on some issues such as data integration, complex structure, and the management of large volumes of information, which can become a problem when working for big projects (Chen & Zhang, 2024).

Sbiti et al. (2021) reviewed previous literature on Building Information Modeling (BIM) and Last Planner System (LPS) integration with a focus on possible synergy between 4D BIM-based visualisation and last planner systems with respect to scheduling. However, they pointed out a lack of a sufficiently integrated BIM–LPS framework and recommended the automation of schedule generation using BIM data and WBS (Work Breakdown Structure). They also agreed that linking BIM with ERP (enterprise resource planning) and documents could be useful for field operations since it would help increase the information flow rate. Similarly, Khalid et al. (2017) discussed the idea of connecting the Building Management System (BMS) data to that of the BIM to get live performance information on buildings. Their idea was based on the idea of employing web socket functionality to make the interaction of the BMS and BIM smooth to facilitate through an interactive application that could be used by facility managers to monitor and control the infrastructures effectively. Although this system has demonstrated favourable feasibility for managing buildings, its broader applicability in project management is limited by the difficulties involved in scaling such integration.

Elaborating above, Churacharit and Chutima (2022) dealt with the current integration of the project management system, with reference to a liquefied natural gas station construction project. In their research, the authors integrated PMIS with IoT for long-distance monitoring and achieved 100% on-time project completion against 75% the previous year from the project site. This integration improved schedule compliance and overcame costs linked with site inspection and travel. In addition, Lee et al. (2019) presented a cloud-based framework that combines government open data with construction site sensor data. The integration was to enhance project management functions including schedule control, resources and worker safety. By integrating weather and environmental data into the construction site data, the system provided a more real-time kind of site management useful for dynamic construction sites. In particular, Pan and Zhang (2021) introduced a digital twin for which BIM, IoT, and data mining enhance project management of buildings. The research showed that it is possible to use IoT devices to feed real-time data to build high-fidelity virtual models that can be analyzed for bottlenecks and tasks. It is possible to make preventive decisions in result of which project managers could predict issues and adjust the workforce distribution. These pioneering studies demonstrate one application of real-time integration for enhancing the likelihood of successful outcomes in projects. The study also revealed several concurrent issues—data compatibility,

increased system, and variable adaptability across project types.

1.2.3. Relationship between Real-Time Data Integration and Delay Prediction Accuracy

The degree of accuracy in predicting the potential additional time needed for project completion as allowed by various inputs and methodologies is known as delay prediction accuracy (de Barcelos Tronto et al., 2008). It encompasses the capacity of predicting quantitative measures of cost and time overruns associated with a project, based on historical data, real-time information and project characteristics. As per Ayers et al. (2024), high accuracy in delay prediction is important because it allows stakeholders to take pre-emptive actions, manage resources optimally and make relevant adjustments to calendar time. Sources of such variation comprise data quality and detail, the forecast models selected, which may be statistical or based on machine learning, or probabilities of project environment fluctuations or other project unpredictable factors. Even with modern predictive technologies and data integration, one of the main issues that can hardly be solved is to receive high accuracy of delay prediction, and this quite often requires an understanding of the context of the project and potential problems.

Cebecauer et al. (2018) provide a framework for estimating urban travel time. Their probabilistic principal component analysis is a multivariate hybrid method that, along with local smoothing, can predict fluctuations in travel time even when there are significant data losses. However, Wu et al. (2023) present a train delay prediction framework that utilises GTFS, which is more of a standardised and accessible feed than the bus and paratransit feed; however, is process is complex. The research proposed a methodology of applying the Long Short-Term Memory (LSTM) model to enhance predictive accuracy in train services; however, seemingly more reliable backup to missing or inconsistent data is not claimed as in Cebecauer et al. (2018).

Oneto et al. (2018) elaborated on this by employing real-scale historical train movement data, the latest trend in Big Data technologies. To remedy that, they used the TDPS (Train Delay Prediction System) through in-memory data processing and deep learning algorithms and demonstrated improved ability in predicting delays, when compared to older static rule-based systems. Unlike the smaller-scope applications of Wu et al., Oneto et al.'s system is best suited for recognizing patterns within large pools of data, which profoundly enhances the system's ability to predict the operations of large railway networks. In the same way, as in Cebecauer et al., their method is focused on historical data for delay detection, which is essential for identifying delays tendencies within more extensive railway systems. Another study employed the latest machine learning, namely XGBoost and Bayesian optimization approaches to forecast delays in urban rail systems (Chtioui et al., 2024). Analyzing the literature, it has been found that the historical data-oriented approach taken by Oneto et al., is different from Chtioui et al. stress on real-time delay prediction with a comparatively short 15-minute forecast horizon that is directly fed into control systems. These differences clearly illustrate the contrast in short-term reactive prognoses (Chtioui et al.) as well as longer-term historical determination (Oneto et al. 2018).

However, Heppe and Liebig (2017) provide a different view by taking into consideration the multi-mode trip planning including delay prediction. Their Spatio-Temporal-Random-Field model combines historical and real-time data as Cebecauer et al. or Oneto et al. did but uniquely applies this to personal trip planning rather than network-wide management. Their work demonstrates the significance of real-time data to predict delays and adapt routine selections in a dynamic way that offers a more user-centric technique. Extending the real-time prediction to other fields, besides transport, Shibuo et al. (2016) explore the concept of flood management using a data integration analysis system. They are interested in increasing the accuracy of flood forecasting, using precipitation models and data on the past behavior of the basin as Oneto et al, have done for the transport sector. The application of ensemble forecasting methods is also similar to the multiple steps ahead forecasting approach adopted by Cebecauer et al. and Wu et al. However, the approach followed by Shibuo et al., The physical modelling of the environmental conditions mainly focused on is quite different from the purely statistical models of environments used in transport-related studies. Thus, the findings illustrated that real-time data play a pivotal role in enhancing predictive accuracy, focusing on the importance of historical

data integration; however, various results associated with the data gap and system complexity show the efficiency and applicability of current real-time predictive models.

1.2.4. The Moderating Effect of Project Complexity

Project complexity relates to the number of project-related factors which make it complex and difficult to manage (Zolin et al., 2009). Such complexity can arise in multiple aspects such as organizational, ecological and technical dimensions. Liu & Cross (2016) claimed that technical complexity deals with the design of the project and the interdependencies between numerous elements, while organizational difficulties relate to the different stakeholders, the demanded expertise diversity, and the coordination among different teams. Ignatius et al. (2012) and Hartono et al. (2019) provide evidence that project complexity is a major determinant of outcomes. For example, Ignatius et al. (2012) emphasized on how technological learning, including organizational memory and information interpretation profiles, are associated with different levels of project complexity that underpin new product development (NPD). In the same way, Hartono et al. (2019) study the relations between KMM (Knowledge management maturity) and structural complexity in construction firms, and establish that firms with higher structural complexity are likely to have better performance if they possess higher KMM. However, both studies clarified that complexity increases the significance of the various moderating factors. For instance, similarly to Liu (2015), where complexity risk in information systems projects also exhibits a mixed moderating role regarding control type that reflects how complexity reconstructs the relationship between performance and learning in Ignatius et al., (2012), and knowledge management (Hartono et al., 2019). These studies emphasized the significance of utilizing complexity as a variable by which enhancement of poor performance can be achieved or worsened based on specific management approaches and organizational contexts applied.

Nonetheless, varying with the source of complexity; the studies suggest very different ways of dealing with project complexity roles. Low-complexity projects, Ignatius et al. (2012) stated that organizational memory and information interpretation positively affect project success in cases of complex-project development while in high-complexity projects organizational memory enhances the speed of development but not necessarily the success of the project. On the other hand, Liu (2015) argued that high complexity risk dramatically reduces the quality of behaviour and self-control; at the same time, increases the importance of outcome and clan control, which shows how control and adaptation mechanisms are different in information systems projects under complexity risk. Moreover, Heredia-Rojas et al. (2022) add a higher layer of moderation by investigating how these value creation processes impact project value, while taking into consideration the roles of requirements uncertainty and project complexity. From this, they conclude that concerning project types, low-complexity projects are vastly different from high-complexity projects (Groups 1 and 2 and Group 3). This study's multigraph analysis implies that complexity combined with uncertainty produces distinct value-creation networks compared to other research that mainly centres on structural complexity or complexity risk.

1.2.5. Literature Gaps

Existing literature lacks comprehensive frameworks that present real-time data integration and delay prediction for projects, especially in complex project environments. It is accepted that project complexity acts as a moderating variable; however, it is not being properly measured; thus, limiting understanding of the impact of the various aspects of technological, structural and context complexity on the efficiency of integrated data processing. Past studies have also failed to present a synthesized view of the impact of real-time data integration on delay prediction accuracy by particularly focusing on large-scale projects. There is a need for context-specific and synthesized studies that could find the relationship between real-time data integration and delay prediction accuracy by considering the moderating impact of project complexity.

2. Material and Methods

2.1.Data Sources

The data used for this study was obtained from seven peer-reviewed academic articles and business journals on project management, delay prediction and real-

time data integration. These sources were chosen based on how well they meet the targets of the present study, the methods that were used within the research and the credibility of the results. All the articles include a description of how real-time data systems are used in the context of various projects with an emphasis on projects of various levels of complexity(Vacanas, 2018). This collected data encompassed a detailed description of the project, the beliefs of the project managers and the assessment of the part that real-time data played in the management of project delay. The work summarises and compares the collected and analysed data on the use of real-time data for delay prediction in different projects (Cheong et al., 2023). This study uses secondary data because the quality of the analysis is improved as well as several practical, ethical, and resource-based problems of primary data collection are avoided.

2.2.Analytical Approach

For data analysis of the current research, the current research uses the enhancing thematic analysis framework which is suitable for synthesizing secondary qualitative data. The research maps evaluates and synthesises empirical and theoretical literature on real-time data and its connection with project complexity to identify trends and gaps. It has used grounded theory as the conceptual framework to avoid prejudice in the development of the conclusions of the research. This approach helps us to come to a more reasonable conclusion and to reveal the dependence between the project's complexity and the efficiency of using real-time data for the project overtime's prognosis(Maina, 2018). The clear and elaborate categorization of themes also contributes to understanding how the levels of complexity affect the usefulness of data and the minimisation of delays in various project settings (Ruggiano and Perry, 2019).

Furthermore, since the cases of this study are based on the outcomes of the studies they can be considered as a form of structured cross-case analysis and therefore the approach used in this study extends the purely thematic one by comparing the real-time data application between the various studies in terms of the size of the project(Wang et al., 2018). This comparative analysis also shows the richness of the uses of real-time data and the conditions under which data systems are most effective. Through the examination of several cases, the research puts together a detailed description of how real-time data flows and project management factors are interconnected and may be used to support better decision-makingbased on data in a complex project (Hughes & Tarrant, 2019). To comprehend the specific regions for which data in real-time is most valuable in regard the delay prediction, it is crucial to examine this level in detail. In addition, it assists in the identification of the gaps that exist in the current project management frameworks and hence makes it easier to develop better frameworks which are in a position to manage the various problems that arise when handling high-complexity projects.

2.3. Integration Techniques

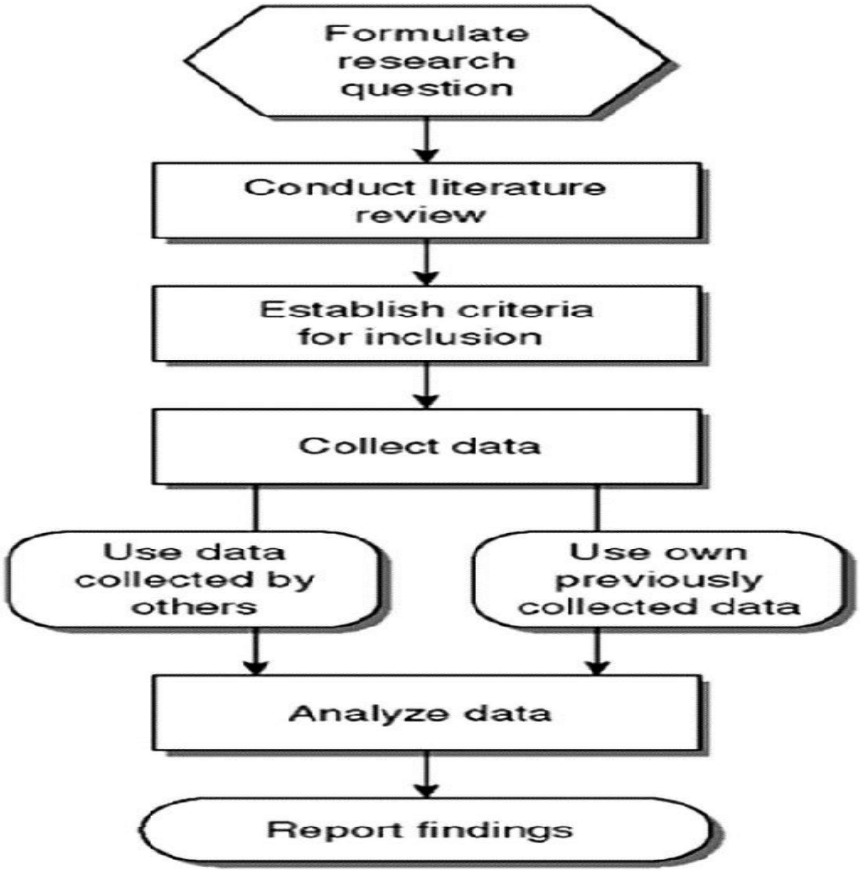
Real-time data integration in project management is a rather intricate process that involves the appropriate transfer of data from one source to another to enhance project decision-making and prevent delays. From the literature reviewed, the integration of real-time data was realised through project management tools that are supplemented with modules of real-time monitoring and analysis. Such platforms can also capture information on various aspects of the project such as the resources, the progress rates and the movement of logistics in the project through the use of sensors and other IoT devices that are integrated into the project framework(Alenazi, 2018).This information is then used with mathematical models to make forecasts and to give on-the-spot suggestions to the project managers. The integration process also includes how the system sends and conveys to the management the level of priority of the information and the projects that the information may be relevant to inform the management of any likelihood of a delay. Also, it has some which are assignable according to the level of the project which includes the dashboard and the reporting options for a specific project (Davidson et al., 2019). It is helpful in the formulation of good strategies towards delay management and this is important in large projects because; size and project complexity may influence how data is dealt with.

The current work uses regression analysis and time series analysis to measure Real-time data impact on delay prediction, and also investigates the influence of various levels of the system's complexity on the given topic. This analysis offers a numerical evaluation of

the correlation between the real-time data feeds and the slippages in the projects and how these data influence the projects' timelines (Rahman, 2018). The results indicate that the effect of the real-time data is contingent upon the project's size and that there is a need to establish a proper methodology concerning data integration. This approach enables not only the assessment of integration techniques but also indicates some factors that affect the delay prediction. Therefore, this makes it possible for us to come up with certain improvements in the project management systems that relate to the size and scope of projects.

2.4. Moderating Variables

The level of project complexity is the variable of interest in this research in as much as it concerns the relationship between real-time data integration and the accuracy of the delay prediction. To capture the project complexity, some factors like project size, range of the project, technological requirement, and the number of people involved in the project result in project operational and managerial problems (Rezvani, 2018). Therefore, the purpose of managing the project complexity aligns perfectly with the argument advanced here that the impact of real-time data integration may not be constant with a given level of project complexity. Yet it is even more significant when it comes to higher numbers of parameters and conditions which can confuse when it comes to the evaluation of results and forecasting future behaviours (Dufour et al., 2019). This research also focuses on the extent to which project complexity introduces variation in the ability of real-time data systems to enhance delay prediction and management. It also assists in grasping the concept of real-time data in a project and the ways of managing it in a manner specific to the kind of project and increasing the effectiveness of the project in the organization.



Secondary data research framework (Logan, 2020)

The research framework presented below can be integrated into the analysis of project complexity to identify how it moderates the effect of real-time data integration on delay prediction accuracy. This framework therefore provides a systematic approach to the research from the question formulation, data collection and analysis to provide a comprehensive understanding of the relationship between real-time data and project complexity (Maina, 2018). Through detailing certain steps like literature review, criteria setting forth for data inclusion, and thorough data analysis, the framework guarantees that every attribute of project complexity, whether it is size, scope, technological challenge or team dynamics has to be assessed for its effectiveness in data integration. This approach not only improves the credibility of the results but also supports the argument regarding project complexity as a vital moderating variable in this study, which is helpful in the formulation of better and more efficient project management strategies that can be applied depending on the level of project complexity (Vacanas, 2018).

3. Results

3.1. Thematic Analysis

The thematic analysis is done to assess the effect of integrating and processing real-time data on the prediction of project delays for simple and complex projects. The analysis focuses on several themes: the factors related to real-time data supporting project management, project characteristics' impact on the utilization of data integration, and differences in the application of data integration between various project contexts.

3.1.1. Real-time data and how it contributes to the management of project

Data integration in real-time is critical in today's project management, most importantly, in the improvement of the accuracy of delay predictions. The real time data systems implementation enables constant checking and control of the project schedule, resources, and risks, as evidenced by the reviewed studies. For example, in the study by Bakici, Nemeh, and Hazir (2021), it is highlighted that real-time information allows project managers to detect divergence from the plan and address them in time hence mitigating the risk of project delay. Butler Vijayasarathy, and Roberts (2020) also explain how real-time data acquired through IoT devices and project management tools enables real-time decisions to be made, which will enable changes that reflect the current status of the project to be made.

The advantages of real-time data are seen best where there is a lot of change and where adjustments have to be made more frequently. According to Mata, Martins, and Inácio (2023), the real-time data feed can instantly disseminate new information about every project stakeholder and possible risks and delays. This level of integration not only simplifies the communication problem but also enhances the accountability of the project team and at the same time results in higher and better quality project delivery.

3.1.2. Influence of Project Complexity on Data Integration Effectiveness

The real-time data on delay prediction has a significant moderating effect on project complexity. There are sources of project complexity such as the size of the project, technological requirements, number of participants, and dependencies. The analysed researches show that despite the fact that real-time data systems are mostly helpful, their usefulness decreases if the project's complexity is high. For instance, Mata, Martins, and Inácio (2023) found that in very detailed projects, the benefits of getting real time data integration for delay prediction are moderate. This means that, there is an inverted 'U' shaped relationship between real-time data integration and the level of complexity in the system; that is, there is an optimal level of integration beyond which, the complexity of the integration of real-time data becomes detrimental. This is due to the difficulty of dealing with large volumes of data at the same time and the overwhelming effect this creates which hinders decision making. However, according to Owolabi et al. (2018), when personnel is trained well and when project management tools are effective, then the problems of project complexity can be managed effectively. They call for the creation of a set of data integration methods specific to high complexity projects, which might further enhance the reliability of delay prediction practices for the increased complexity of the projects.

3.2. Comparative Analysis of Real-Time Data Integration in Varying Project Complexities

The cross-sectional study of the projects of different levels of difficulty allows to disclose the peculiarities of the effect of real-time data integration on the delay forecast and project management. Analysing work on the scale of relative simplicity of the projects and comparing them to highly complex ones, it is possible to draw conclusions about the fact that the real-time data integration is the context-sensitive solution. In the low and moderate complexity projects, the incorporation of real-time data is quite easy and very efficient. These projects generally have less number of activities, less complex activities and less integration of activities, which makes it convenient for the project managers to analyse and respond to real time data. Hence, real-time data systems in these environments can substantially improve the delays' prognosis since they offer unambiguous and easily understandable information that can be easily applied. For example, the article, "Managing Software Development Projects for Success" by Butler Vijayasarathy, and Roberts (2020) considers software development projects that are of different levels of difficulty. The results provided imply that in less complicated projects, it is possible to apply a plan-oriented approach that uses real-time data. This way and that is how project managers stay abreast of progress of a project according to a set of parameters, see if the project is off track in any way, and correct it. The use of real-time data within these environments provides a degree of accuracy that is hard to achieve when using conventional project management tools. The issue appears when the project complexity rises, and the integration of real-time data also becomes a problem. In comparatively less complex project environments, where there is more dependency between tasks and additional number of factors that are required to be coordinated, the impact of real time data integration of work can be high or low depending on the project management methodology used. Another study in software development projects reveals that in these more complex situations, it is better to use an agility-based approach that relies on real-time information rather than planning-based approach (Ramadan et al., 2020). The use of the agility based approach ensures that project teams are always able to overcome challenges, and continue with the projects since they are flexible enough to adapt to change.

There is even more variability in this case because of the high complexity of such projects as large scale infrastructure projects or multi-national PPP projects. The paper on "Predicting Completion Risk in PPP Projects using Big Data Analytics" by Owolabi et al. (2028) shows that, contrary to the traditional business environment, the real-time data can still be relevant in the PPP projects; however, the volume and variability of data require more sophisticated analytical approaches. In such cases, conventional real-time data integration tools may not be adequate because the level of risk and uncertainty inherent in such a project is higher than in more straightforward examples. For example, in the PPP projects under study, the researchers used random forests and deep neural networks to analyse the big data that are generated in real time. These models were also useful in establishing some attributes that could not be identified by normal statistical techniques. In the light of the conclusions made in the course of the study it is possible to state that in the case of highly complex projects the success of the real-time data integration depends not only on the quality of the data but also on the ability to analyse and interpret the data with the help of the tools and methods. This result lends credence to the proposition that firms should allocate resources to build sophisticated analytics as a component of the real-time data integration process, most especially for projects that are complex. Moreover, in the course of the comparison of the two approaches, it is possible to identify the need for adaptation of the real-time data integration methodology to the project needs. In less complex environments it could be enough to integrate real-time information into the existing project management practices to gain significant benefits. However, where the level of complexity is higher, the level of sophistication may be required, where such things as specialized tools, sophisticated analysis, and project management techniques that are appropriate to the situation and the level of complexity may be required.

4. Discussion

4.1. Interpretation of Results

The conclusion made in this study is in line with previous studies on the applicability of real-time data integration in project management especially in the context of delay prediction and prevention. Existing literature points to the importance of real-time data in improving the decision-making process because such data can be used to institute preventive measures by the project managers when certain situations occur (Owolabi et al., 2018). This proactive approach is most useful in simple projects, where data can be gathered, analysed and used in project management. Ramadan et al. (2020) work affirms this by pointing out that use of real time data enhances scheduling precision and resource utilization in simple to moderately complex projects. However, as the project complexity rises, the literature also points out the issues of handling and analysing a vast amount of data (Wang et al., 2020). In complex projects, these systems alone may flood the project managers with information thus prolonging the time taken to make a decision instead of preventing the same.

This study also supports the notion that the challenges of real-time data integration in large scale projects require the use of other approaches. For example, integrating real-time data with machine learning algorithms, predictive analytics or Agile or CPM project management methodologies enhances the use of data systems (Awada et al., 2021). The incorporation of such other methodologies will enable management of large volumes of data within project teams as well as improvement of overall project results, in an effort to ensure that real time data does not act as an impediment to project performance. Therefore, real-time data is useful on all project size levels, however, the method of collecting real-time data varies depending on the size of the project.

4.2. Significance of Findings

The result of this study is very useful for the project managers especially those managing large project or project that is complex in nature. Realising the weakness of real-time data systems in high-complexity environment is a way of informing managers on more suitable approaches of increasing the efficiency of the real-time data systems (Kopetz and Steiner, 2022). For instance, the use of machine learning models in complex projects can assist project managers in the extraction of more relevant information from real-time data. Also, there is a possibility of enhancing the effectiveness of real-time data systems by offering specific training for project teams on how to use them.

4.3. Limitations

There are certain limitations of this research which must be stated. Firstly, the use of secondary data sources does not allow to investigate the specifics of real-time data integration in various project environments. Although the above discussed studies offer a general insight on the effect of real-time data on delay prediction and project management, they might not be able to reflect the real life issues faced by the project managers.

5. Conclusion

This research aimed at evaluating the effect of real-time data integration on the delay prediction of large-scale projects with a focus on the moderating role of project complexity. The actual research is based on four primary studies, which were aimed at revealing the impact of project complexity levels on the efficiency of real-time data systems in improving the scheduling of project activities and the calculation of delays. The approach used allowed for the identification of all the literature that is available and how real-time data can contribute to improved decision-making in the context of project management. The results were compared with other studies that have been conducted including the study by Choi et al. (2021) who also noted that advanced analytics and machine learning can help in improving real-time data integration in project systems.

This research offers a lot of evidence to support the fact that real-time data integration enhances the management of project schedules in simple environments; it is recommended that in complex projects, there is a need to involve more complex analytical tools and data analysis techniques. According to previous research, simpler projects are likely to be benefited by real-time data because fewer variables are involved when it comes to complex projects and therefore it is important to handle data and its analysis in a more complicated manner than complex projects (Dufour et al., 2019). Therefore, it is

proposed that organizations should align their real-time data integration strategies with the level of complexity of the project. For basic projects, basic real-time monitoring tools may work well while for complex projects, it may be necessary to employ sophisticated data analytics, possibly coupled with machine learning, to properly monitor and analyse big and different data. The future work should also examine which particular types of analytical methods and machine learning algorithms are the most useful to apply at different levels of project complexity and elaborate these findings across industries to enhance the effectiveness of real-time data integration.

6. References

Alenazi, E. O. (2018). Using BIM for Delay Management in Public Sector Construction Projects in KSA. [PhD Thesis, Loughborough University Reading, UK]. https://core.ac.uk/download/pdf/288351422.pdf

Awada, M., Srour, F.J. and Srour, I.M., 2021. Data-driven machine learning approach to integrate field submittals in project scheduling. Journal of Management in Engineering, 37(1), p.04020104.

Ayers, J., Chadburn, O., & Hoque, K. A. (2024). Behaviourally-Informed Early Warning and Anticipatory Action: Unleashing the Potential of Locally-Led Pre-emptive Response and the Forecast-Based Financing it Requires. Current and Emerging Trends in the Management of International Disasters. https://uta.pressbooks.pub/trendsininternationaldisastermanagement/chapter/behaviourally-informed_early_warning_and_anticipatory_action/

Bakici, T., Nemeh, A. & Hazir, Ö., (2021). Big data adoption in project management: insights from French organizations. IEEE Transactions on Engineering Management, 70(10), pp.3358-3372.

Butler, C. W., Vijayasarathy, L. R., & Roberts, N. (2020). Managing Software Development Projects for Success: Aligning Plan- and Agility-Based Approaches to Project Complexity and Project Dynamism. Project Management Journal, 51(3), 262–277. https://doi.org/10.1177/8756972819848251

Cebecauer, M., Jenelius, E., & Burghout, W. (2018). Integrated framework for real-time urban network travel time prediction on sparse probe data. IET Intelligent Transport Systems, 12(1), 66-74. https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/iet-its.2017.0113

Chen, C. P., & Zhang, C. Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. Information sciences, 275, 314-347. https://www.sciencedirect.com/science/article/abs/pii/S0020025514000346

Cheong, H., Lyons, A., Houghton, R., & Majumdar, A. (2023). Secondary Qualitative Research Methodology Using Online Data within the Context of Social Sciences. International Journal of Qualitative Methods, 22, 160940692311801. https://doi.org/10.1177/16094069231180160

Choi, S.-W., Lee, E.-B., & Kim, J.-H. (2021). The engineering machine-learning automation platform (map): A big-data-driven ai tool for contractors' sustainable management solutions for plant projects. Sustainability, 13(18), 10384.

Chtioui, S., Mouelhi, S., Saudrais, S., Azib, T., Ille, M., Morel, M., & Oru, F. (2024). XGBoost in Public Transportation for multi-attribute data: delay prediction in railway systems in real-time. IEEE Access.

Churacharit, C., & Chutima, P. (2022). An integration of project management body of knowledge and project management information system to improve on-time deliverables of liquefied natural gas station construction projects. Engineering Journal, 26(1), 55-73. https://engj.org/index.php/ej/article/view/4412

Davidson, E., Edwards, R., Jamieson, L., & Weller, S. (2019). Big data, qualitative style: A breadth-and-depth method for working with large amounts of secondary qualitative data. Quality & Quantity, 53(1), 363–376. https://doi.org/10.1007/s11135-018-0757-y

de Barcelos Tronto, I. F., da Silva, J. D. S., & Sant'Anna, N. (2008). An investigation of artificial neural network-based prediction systems in software project management. Journal of Systems and Software, 81(3), 356-367. https://www.sciencedirect.com/science/article/pii/S016412120700132X

Dufour, I. F., Richard, M.-C., & Li, J. (2019). Theorizing from secondary qualitative data: A comparison of two data analysis methods. Cogent Education, 6(1), 1690265. https://doi.org/10.1080/2331186X.2019.1690265

Fobiri, G., Musonda, I., & Muleya, F. (2022). Reality capture in Construction Project Management: A review of opportunities and challenges. Buildings, 12(9), 1381. https://www.mdpi.com/2075-5309/12/9/1381

Hartono, B., Sulistyo, S. R., Chai, K. H., & Indarti, N. (2019). Knowledge management maturity and performance in a project environment: Moderating roles of firm size and project complexity. Journal of Management in Engineering, 35(6), 04019023. https://ascelibrary.org/doi/abs/10.1061/(ASCE)ME.1943-5479.0000705

Heppe, L., & Liebig, T. (2017). Real-time public transport delay prediction for situation-aware routing. In KI 2017: Advances in Artificial Intelligence: 40th Annual German Conference on AI, Dortmund, Germany, September 25–29, 2017, Proceedings 40 (pp. 128-141). Springer International Publishing. https://link.springer.com/chapter/10.1007/978-3-319-67190-1_10

Heredia-Rojas, B., Liu, L., Ramírez-Correa, P., & Mariano-Melo, A. (2022). Moderating effects of requirements uncertainty and project complexity on value creation processes and project value: a multigroup analysis via partial least squares structural equation modelling. Mathematical Problems in Engineering, 2022(1), 4106360. https://onlinelibrary.wiley.com/doi/full/10.1155/2022/4106360

Hughes, K., & Tarrant, A. (2019). Qualitative secondary analysis. Sage. https://books.google.com/books?hl=en&lr=&id=uHW9DwAAQBAJ&oi=fnd&pg=PP1&dq=research+methodology,+secondary+qualitative,&ots=RzHAcW5nPS&sig=jsJg3tzHntVWt-ADVVvRwVwR1bQ

Ignatius, J., Leen, J. Y. A., Ramayah, T., Hin, C. K., & Jantan, M. (2012). The impact of technological learning on NPD outcomes: The moderating effect of project complexity. Technovation, 32(7-8), 452-463. https://www.sciencedirect.com/science/article/abs/pii/S0166497212000326

Khalid, M. U., Bashir, M. K., & Newport, D. (2017). Development of a building information modelling (BIM)-based real-time data integration system using a building management system (BMS). Building Information Modelling, Building Performance, Design and Smart Construction, 93-104.

Kopetz, H. and Steiner, W., 2022. Real-time systems: design principles for distributed embedded applications. Springer Nature.

Lee, C. M., Kuo, W. L., Tung, T. J., Huang, B. K., Hsu, S. H., & Hsieh, S. H. (2019, May). Government open data and sensing data integration framework for smart construction site management. In Proceedings of the 36th International Symposium on Automation and Robotics in Construction, ISARC (pp. 1261-1267). https://www.academia.edu/download/66360853/d85d74d85e1058a1bc51f337ae9c2127fd5c.pdf

Li, Z., Wen, C., Hu, R., Xu, C., Huang, P., & Jiang, X. (2021). Near-term train delay prediction in the Dutch railways network. International Journal of Rail Transportation, 9(6), 520–539. https://doi.org/10.1080/23248378.2020.1843194

Liu, S. (2015). Effects of control on the performance of information systems projects: the moderating role of complexity risk. Journal of Operations Management, 36, 46-62. https://www.sciencedirect.com/science/article/abs/pii/S0272696315000200

Liu, W. H., & Cross, J. A. (2016). A comprehensive model of project team technical performance. International Journal of Project Management, 34(7), 1150-1166. https://www.sciencedirect.com/science/article/abs/pii/S026378631630028X

Logan, T. (2020). A practical, iterative framework for secondary data analysis in educational research. The Australian Educational Researcher, 47(1), 129–148. https://doi.org/10.1007/s13384-019-00329-z

Maina, P. W. (2018). An investigation into the role of building information modelling in reducing cost and time overruns experienced in 2-dimensional based construction-a survey of construction practitioners in Nairobi County. [PhD Thesis, University of Nairobi]. http://erepository.uonbi.ac.ke/handle/11295/104039

Mata, M. N., Martins, J. M., & Inácio, P. L. (2023). Impact of absorptive capacity on project success through mediating role of strategic agility: Project complexity as a moderator. Journal of Innovation & Knowledge, 8(1), 100327.

Oneto, L., Fumeo, E., Clerico, G., Canepa, R., Papa, F., Dambra, C., ... & Anguita, D. (2018). Train delay prediction systems: a big data analytics perspective. Big data research, 11, 54-64.

Owolabi, H. A., Bilal, M., Oyedele, L. O., Alaka, H. A., Ajayi, S. O., & Akinade, O. O. (2018). Predicting completion risk in PPP projects using big data analytics. IEEE Transactions on Engineering Management, 67(2), 430–453.

Pan, Y., & Zhang, L. (2021). A BIM-data mining integrated digital twin framework for advanced project management. Automation in Construction, 124, 103564. https://www.sciencedirect.com/science/article/pii/S0926580521000157

Rahman, M. (2018). Risk Management and Measurement of Risk management performance in complex projects [Master's Thesis, M. Rahman]. https://oulurepo.oulu.fi/handle/10024/11441

Ramadan, M., Salah, B., Othman, M. and Ayubali, A.A., 2020. Industry 4.0-based real-time scheduling and dispatching in lean manufacturing systems. Sustainability, 12(6), p.2272.

Rezvani, A. (2018). Large scale complex projects—beyond the'iron triangle': An influence of soft skills on project success. https://eprints.qut.edu.au/116591

Ruggiano, N., & Perry, T. E. (2019). Conducting secondary analysis of qualitative data: Should we, can we, and how? Qualitative Social Work, 18(1), 81–97. https://doi.org/10.1177/1473325017700701

Saidani, I., Ouni, A., & Mkaouer, M. W. (2022). Improving the prediction of continuous integration build failures using deep learning. Automated Software Engineering, 29(1), 21. https://doi.org/10.1007/s10515-021-00319-5

Sbiti, M., Beddiar, K., Beladjine, D., Perrault, R., & Mazari, B. (2021). Toward BIM and LPS data integration for lean site project management: A state-of-the-art review and recommendations. Buildings, 11(5), 196. https://www.mdpi.com/2075-5309/11/5/196

Shibuo, Y., Ikoma, E., Valeriano, O. S., Wang, L., Lawford, P., Kitsuregawa, M., & Koike, T. (2016). Implementation of real-time flood prediction and its application to dam operations by data integration analysis system. Journal of Disaster Research, 11(6), 1052-1061.

Vacanas, Y. (2018). The risk of delay in Cyprus construction industry: Causes, effects and control [PhD Thesis, Department of Civil Engineering and Geomatics, Faculty of Engineering and ...]. https://ktisis.cut.ac.cy/handle/20.500.14279/13255

Wang, J., Liu, Y., Ren, S., Wang, C. and Wang, W., 2021. Evolutionary game based real-time scheduling for energy-efficient distributed and flexible job shop. Journal of Cleaner Production, 293, p.126093.

Wang, Y., Liu, Y., & Canel, C. (2018). Process coordination, project attributes and project performance in offshore-outsourced service projects. International Journal of Project Management, 36(7), 980–991.

Wu, J., Du, B., Gong, Z., Wu, Q., Shen, J., Zhou, L., & Cai, C. (2023). A GTFS data acquisition and processing framework and its application to train delay prediction. International Journal of Transportation Science and Technology, 12(1), 201-216. https://www.sciencedirect.com/science/article/pii/S2046043022000090

Zolin, R., Turner, R., & Remington, K. (2009). A model of project complexity: distinguishing dimensions of complexity from severity. In International research network of project management conference (IRNOP) (pp. 1-30). https://eprints.qut.edu.au/29011/.