Data Analytics for Project Delivery: Unlocking the Potential of an Emerging Field

Yixue Shen[1], Naomi Brookes[1], Luis Lattuf Flores[1], and Julia Brettschneider[2,*]

[1]Warwick Manufacturing Group (WMG), University of Warwick, Coventry, UK

[2]Department of Statistics, University of Warwick, Coventry, UK

[*]Corresponding author: julia.brettschneider@warwick.ac.uk

Abstract

Purpose

In recent years, there has been a growing interest in the potential of data analytics to enhance project delivery. Yet many argue that its application in projects is still lagging behind other disciplines. This paper aims to provide a review of the current use of data analytics in project delivery encompassing both academic research and practice to accelerate current understanding and use this to formulate questions and goals for future research.

Design / Methodology/ Approach

We propose to achieve the research aim through the creation of a systematic review of the status of data analytics in project delivery. Fusing the methodology of integrative literature review with a recently established practice to include both white and grey literature amounts to an approach tailored to the state of the domain. It serves to delineate a research agenda informed by current developments in both academic research and industrial practice.

Findings

The literature review reveals a dearth of work in both academic research and practice relating to data analytics in project delivery and characterises this situation as having 'more gap than knowledge.' Some work does exist in the application of machine learning to predicting project delivery though this is restricted to disparate, single context studies that do not reach extendible findings on algorithm selection or key predictive characteristics. Grey literature addresses the potential benefits of data analytics in project delivery but in a manner reliant on 'thought-experiments' and devoid of empirical examples.

Originality / Value

Based on the review we articulate a research agenda to create knowledge fundamental to the effective use of data analytics in project delivery. This is structured around the functional framework devised by this investigation and highlights both organisational and data analytic challenges. Specifically, we express this structure in the form of an 'onion-skin' model for conceptual structuring of data analytics in projects. We conclude with a discussion about if and how today's Project Studies research community can respond to the totality of these challenges. This paper provides a blueprint for a bridge connecting data analytics and project management.

Keywords

Project Data Analytics; Datafication; Artificial Intelligence; Digital Transformation; Project Studies; Project performance; Project delivery performance; data

Publication

A later version of this manuscript was published: *International Journal of Managing Projects in Business*, 2024 Apr 2; available from: http://dx.doi.org/10.1108/ijmpb-08-2023-0180

1. Introduction

'Data is the new oil' is a phrase reputed to be coined by the British mathematician and entrepreneur, Clive Humby, in 2006 (Humby, 2006). Since that assertion, data analytics, sometimes referred to as big data analytics (BDA), has had profound effects on wide sectors of the economy and the ways in which they make decisions and generate value (as summarised e.g. in Oesterreich, 2022; Verbeke *et al.*, 2017; Tsai et al., 2015). The achievements in industrial practice were enabled by data analytical tools arising from the accelerated growth of data science as a formal academic discipline during in the first decade of this century (Sarker, 2021). Situated in the intersection of computer science, mathematics, and statistics, data science also operates as a core methodology for artificial Intelligence, a symbiosis manifested in the UK The Alan Turing Institute which was funded in 2015 as joint venture of five leading universities with the aim to lead research into the exploitation of big data for engineering applications. Data science was initially brought to life by the American statistician, mathematician, and scientist John W. Tukey in a seminal paper promoting the recognition of a 'science of data' followed by his influential book offering a toolbox of data exploration methodology (Tukey, 1962; Tukey, 1977). For use in industrial practice, data science has been integrated in business decision processes through formulating transparent and holistic processes implementable by to companies beyond Big Tech such as the SMART framework (Marr, 2015).

However, there is a strong perception that project delivery is yet to match the adoption of data analytics seen in other disciplines (Bevort, 2020; Lucas, 2021; Rowland, 2020). Project delivery refers to the comprehensive process of completing and delivering a project to the client or end-users. It encompasses all the activities, tasks, and phases involved in taking a project from its initiation to its conclusion (AIPM, 2022). In response to this situation, professional project associations have introduced a plethora of activities to get to grips with the impact of the data now available on project delivery (PMI, 2020). The UK's Association for Project Management published a report in 2020 (APM, 2020) looking at the state of the art and science of project data analytics and painted a picture of ad-hoc, limited adoption and confusion in the profession surrounding digital terminology.

Yet the benefit from effective use of data in projects appears huge. In the specific context of infrastructure project delivery, feasibility studies suggest savings of £23bn per annum in the UK alone (Brookes & Lattuf, 2021). Project delivery is in an excellent position to take advantage of data analytics. Project delivery has involved the creation and gathering of data since its inception with the use of such tools as the Gantt Chart and critical path analysis (Geraldi & Lechter, 2012; Hammad *et al.*, 2020; Jugdev *et al.*, 2013). Projects have legacies of digital data from computerised management software such as Primavera. In the face of such digital data riches, it is paradoxical that the project delivery profession has underused the opportunities to exploit them through better analytics.

In the light of practitioners' delay in a systematic adoption of project data analytics, it is interesting to examine how the Project Studies research community has responded to the datafication challenge. In the past data analytics has been underserved by research, but more recently there have been signs that research into these challenges is beginning to be recognised. The International Journal of Managing Projects in Business, in particular, has published papers on digitalization (Walker & Lloyd-Walker, 2019), big data (Fox & Do, 2013; Olsson & Bull-Berg, 2015), project management assessment (Williams, *et al.*, 2014; Perkins, *et al.*, 2020), data visualization (Killen, 2017).

This paper aims to construct a comprehensive understanding of the current use of data analytics in project delivery. It posits that until the practitioner and research community have reached this understanding, they will not be able to fully leverage the rewards associated with effective use of data in project delivery. The aims of this paper, therefore, are two-fold:

1. to critically review the current application of data analytics to project delivery from both a practitioner and research stance to gain an understanding of current status of activity
2. to use that review to develop a research agenda that will assist in understanding and overcoming the data analytical challenges in a project delivery context

It is important to understand the innate challenges in pursuing these aims in the context of data analytics in projects. Having examined the use of data analytics through both lenses – academic research and practice – and following initial searches, we observed a rapidly growing yet fragmented amount of literature in this space. Theoretical constructs, definitions, and standards have not sufficiently matured and not yet widely accepted, which is a natural complexity in a nascent field. In addition, data analytics in projects is approached by paradigms from both organisational studies, leading to friction at the interface and giving raise to debates around the necessity for paradigm shifts. To overcome these challenges, our paper comprises developing an integrative literature review based on both white and grey literature to identify key features of current data analytics use. It then uses this review to design a research agenda aiming at supporting knowledge creation around the existing and potential application of data analytics in the area of project management. It concludes by examining how far this research agenda is practicable in the current Project Studies environment.

2. Devising a methodology to capture the status of data analytics in projects: conducting an integrative literature review

Carrying out a literature review field is an established first step in any research investigation and is particularly relevant to support the formation of streams in a newly emerged research area. A wide variety of approaches have been formulated to accomplishing this task and these have been encapsulated in a standard body of knowledge (Fink, 2019). However, performing a literature review in a nascent area, such as the application of data analytics to projects, bears additional complexities and uncertainties and demands a tailored approach to systematising the knowledge in the absence of widely agreed definitions of constructs, frameworks, or standard. A lack of clarity in conceptual terminology implies a lack of search terms to dissect then area in question. Furthermore, in the context of a nascent research area, the outputs from any such review are required to be transformative. Rather than an incremental development of existing knowledge through synthesis, an integrative review aims to establish preliminary conceptualizations.

The dearth of maturity of the field also provides difficulties in the availability of literature. As activity in the Project Studies research community related to data analytics in projects is not yet an established field and cannot be used to effectively gauge the pervasiveness of the phenomenon in practice. This means that the literature reviewed will need to involve both white and grey literature. Not at last, the use of grey literature in Project Studies has recently become more recognised and ways of dealing with some of the more difficult aspects of such literature established (Aubry *et al.*, 2021; Denicol *et al.*, 2020; Papadonikolaki *et al.*, 2022). Our investigation adapted these methods for the context of an immature literature domain.

Given the heterogenous quality of the literature especially in the grey area, careful consideration needs to be given to the type of literature review adopted. An 'integrative' literature review approach appears eminently suitable for a nascent area such data analytics in project delivery. Snyder (2019) emphasizes that the integrative review should result in the advancement of knowledge and rather than being constraint to descriptive or historic approaches only and should result in the generation of new conceptual frameworks. This capability matches our requirement for a transformative output.

This paper creates a novel approach to literature reviewing by fusing the aforementioned two methods. Firstly, an integrative review serves to synthesise existing knowledge while putting an emphasis on the development of a conceptual structuring framework. Secondly, the inclusion of grey literature will acknowledge the pace of knowledge gain in a nascent area and the crucial role industrial practice has for generating use cases and for thinking about the future potentials of data analytics for developing data analytics methods tailored to predict project delivery.

The resulting framework is a research output in its own right. If sufficiently robust, it could provide a seminal contribution to structuring further research in this arena. We create the conceptual framework based on comprehensive methodological guidance provided by Torraco (2005):

1. creating a coherent conceptual structuring
2. using the structuring to create search terms
3. determining the databases to be searched
4. defining inclusion and exclusion criteria
5. using a template analysis approach to thematically analyse the core literature

Lawrences seminal work (Lawrence *et al.*, 2015) on 'grey' literature posits it as literature that is produced outside an organised mode of production. In their 2017 paper, Adams *et al.* refine that definition to distinguish between 'white' literature which is peer reviewed and 'grey' literature which is not (Adams *et al.*, 2017). Grey literature has the advantages of contemporariness and avoiding publication bias. Its value has also been critically appraised in health care research, a domain well-known for its strong tradition of publishing systematic reviews (Adams *et al.*, 2017; Russell, 2005). Grey literature does have its difficulties, however. The quality of its review process will be less rigorous than in a blind peer-reviewed journal or may even be non-existent. Issues of quality were addressed by only including grey literature with significantly credibility which typically includes book chapters, government reports, company publications and think tank publications. Low retrievability grey literature such as emails, blogs, presentations was excluded. In the pursuit of the exploration of the emerging field of project data analytics (PDA), the trade-off was in favour of including grey literature.

3 Results of the review: towards a status review and a research agenda

3.1 Developing a coherent conceptual structure of project data analytics

The first stage of an integrative review is to develop a coherent conceptual structuring of the area under review. This stage is essential when reviewing a nascent body of literature such as project data analytics, because we encounter the inconsistent use of constructs and the absence of a pre-existing framework.

In order to overcome the problem of existing definitions being overlapping, recursive, and multi-faceted, a 'functional' approach to creating a conceptual framework was employed. A function-based definition groups items according to their functional interaction with their environment (Cooper *et al.*, 2007). The association of functions with data analytic conceptions enabled a clear delineation between these constructs. We identified the functions encompassed by data activities to create the definitions. This was achieved through reviewing in data processing terms of definitions and resulted in the identification of six data functions:

1. capturing data from the environment
2. preparing data
3. presenting data
4. modelling data
5. making predictions using data models
6. using data to act upon the environment

The delineation of the data function involved in analytic activities was not sufficient to construct a meaningful conceptual framework. The framework also needed to represent the recursive nature of these constructs. This aspect of functionality was represented by an 'onion skin' framework where each layer of the framework represents a construct and encompasses the functionality of the construct layers below. This framework is designed to be widely adopted in various types of projects, and data also can be understood in a general way. The resultant conceptual framework produced at the start of the integrative review is given in Figure 1.

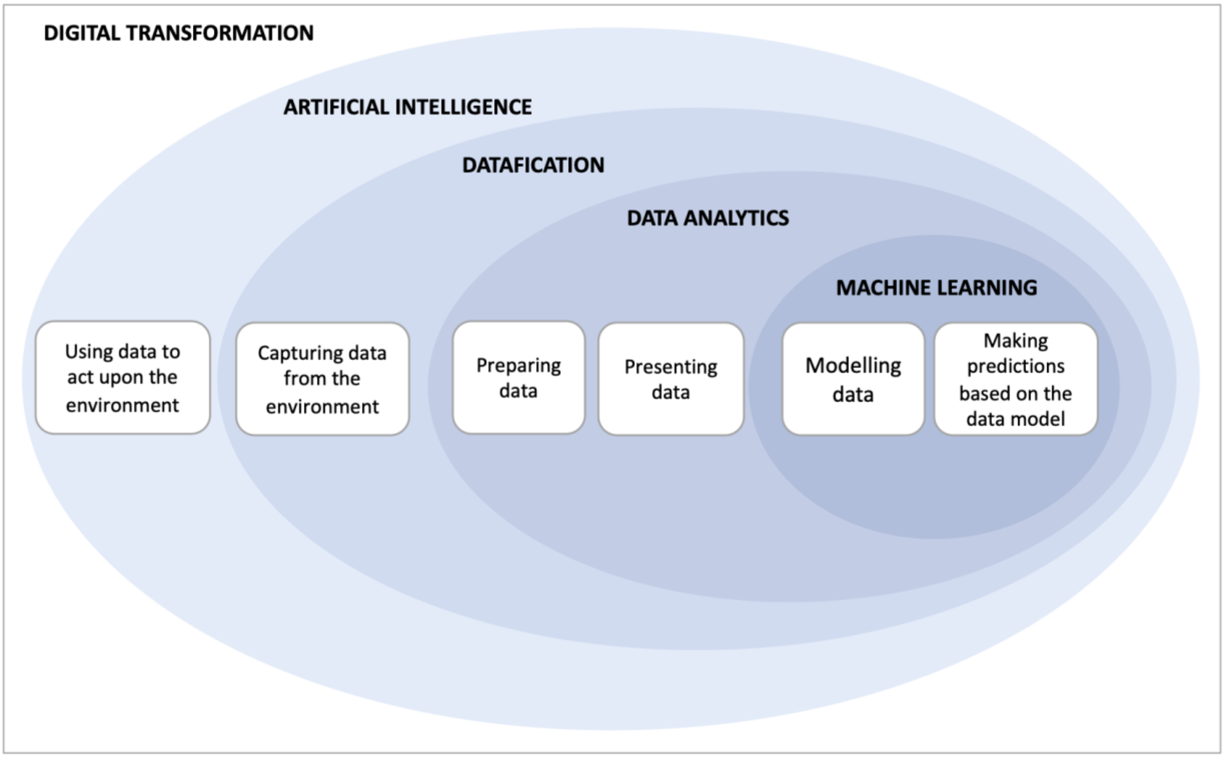

Figure 1: The 'onion-skin' conceptual structuring of data analytics in projects

The onion skin model depicts a progression from the broad context of digital transformation to more specialized and advanced capabilities, highlighting the interconnectedness of these layers in the context of data analytics in project delivery. It can be used to determine which components are preconditions for other components. A further explanation of each layer of the conceptual framework is given below:

Machine Learning in Project Delivery – Machine learning is a subset of artificial intelligence that focuses on developing algorithms and models that enable computers to learn and improve from experience without being explicitly programmed. Machine learning is a sophisticated form of data analytics, where algorithms learn from patterns in data to make predictions or automate decision-making processes. It represents the advanced capabilities within the broader field of data analytics. Algorithms do this by first spotting statistically significant patterns between outcomes and inputs. The mechanisms identified in these ` training data' will then be used to predict future outputs from given inputs. In project delivery terms, machine learning is used firstly to spot historic patterns between some characteristics of a projects (e.g. size, type, geographic locations, procurement type) and some aspect of project delivery performance (e.g. duration). The algorithm will then use the patterns to predict a delivery performance for a new project based on its given characteristics. Each time the process is used, the error between the predicted performance and the actual performance is fed-back so that the algorithm gets more accurate the more it is used. Machine learning therefore incorporate two functional aspects: modelling project data (both in terms of project characteristics and project delivery) and making prediction using that data model.

Data Analytics in Project Delivery – Data analytics involves the analysis of data to derive meaningful insights, patterns, and trends that can inform decision-making and improve business processes. It turns raw data into actionable insights for project managers, allowing them to make informed decisions. Data analytics is the application of analytical techniques to the data produced through datafication. Data analytics also triggers a new level of datafication, creating a dynamic interplay between the two components. Therefore, it is an iterating and repeating process. Data analytics in project delivery can be used both descriptively (simply presenting data to assist in decision making) and predictively (enabling prediction of delivery performance to be made for projects such that better decisions can be made surrounding their delivery). In functional terms, data analytics in project delivery incorporates the functions associated with machine learning and those associated with preparing data (ensuring its quality) and presenting data (in a way which is easily assimilable.)

Datafication in Project Delivery – Datafication is the process of transforming various activities, processes, and aspects of projects into data for analysis and decision-making. Datafication provides the raw material for artificial intelligence by converting real-world activities into digital data that Artificial Intelligence systems can analyze to generate insights. Project Datafication encompasses the functionality of data analytics (preparing data, presenting data, modelling data, and making predictions with those model) but also crucially involves the capturing of data from the environment. The creation and retention of digital data in projects is far from new and has been widespread since the advent of computerised project control packages, such as Primavera and Microsoft Project. Given the longevity and proliferation of digital project delivery data, it is surprising that data analytics has not been more quickly adopted in projects. However, data analytics requires a high level of data quality to produce meaningful results. Though datafication is already prevalent in project delivery, the quality of that data is less certain.

Artificial Intelligence in Project Delivery – Artificial intelligence involves the development of computer systems that can perform tasks that typically require human intelligence, such as learning, reasoning, problem-solving, and decision-making. Artificial intelligence Artificial intelligence Artificial intelligence Artificial intelligence is a subset of digital transformation, representing the advanced capabilities that emerge when machines can learn from data and perform intelligent tasks. In the context of this conceptual structuring, Artificial Intelligence encompasses the functionality of datafication (which in turn encompasses data analytics and machine learning) and adds the ability to use data to act upon the environment. It is this 'closing of the loop' of taking data in from the environment, processing it and then changing the environment on the basis of this analysis that yields the ability to act independently in a way that simulates intelligence. In project delivery terms, this could mean that an intelligent agent could replace human activity by taking in data on project delivery performance, analysing it, and then effecting change on the basis of that analysis such as rescheduling tasks or reallocating resources. The ability to make that change relies on some form of robotic process interaction with its environment.

Digital Transformation in Project Delivery – Digital transformation is a very broad term that refers to the application of digital technologies into all aspects of projects, fundamentally changing how organizations operate and deliver value. It sets the broader context for the integration of digital technologies, providing the foundation for subsequent layers. In the context of this conceptual structure, it encompasses the functionality of data analytics, datafication, machine leaning and artificial intelligence to projects. Data transformation provides the environment under-pinning the use of digital data in delivering projects.

It is important to stress that our goal is to create a structure that enables us to delineate the constructs of interest heterogeneously and, simultaneously, to explain their interactions. We do not claim any legitimacy for our framework beyond its role in our integrative review.

3.2 Creating the core literature for the status review

3.2.1 Creating the search terms and exclusion rules

Having created a coherent conceptual structuring of project data analytics and its related constructs enables us to organise the review of project data analytics drawing on white and grey literature. The utility of the conceptual structuring was first evident in the selection of search terms. The search terms used to search both white and grey literature are given in Table **I**.

Table I: Search Terms involved in the Integrative Review

Construct	**Search Terms**
Machine Learning	"machine learning" AND "project management" "machine learning" AND Project" AND "decision making"
Data Analytics	"data analytics" AND "project management" "data analytics" AND Project" AND "decision making"
Datafication	"datafication" AND "project management" "datafication" AND Project" AND "decision making"
Artificial Intelligence	"artificial intelligence" AND "project management" "artificial intelligence" AND Project" AND "decision making"
Digital Transformation	"digital transformation" AND "project management" "digital transformation" AND Project" AND "decision making"

Items were excluded from the body of literature if they met any of the following criteria:

1. Their unit of analysis was not project or programmes delivery but a sector.
2. Their unit of analysis was not project or programme delivery but a functional activity.
3. They concentrated on aspects that were not related to project data analytics.

An example for an exclusion based on the first criteria would be referencing AI in the construction sector as an association with projects was only implicit. The second exclusion criteria would for example apply to papers that discussed the use of machine learning to understand conflict resolution (albeit some of this was happening in a project organising context such as Construction). The third exclusion criteria would encompass papers that looked at AI in projects but concentrated on robotic process automation.

Exclusion took place initially based on titles. Where titles were insufficient, abstracts or executive summaries (where available) were considered. Finally, if exclusion could not be made on the basis of the title and abstract, the content of the whole item was reviewed.

3.2.2 Identifying the bodies of literature to be reviewed

White literature was interpreted as peer reviewed journal articles in SCOPUS in published in English. The process of identifying the reviewed white literature is presented in Figure 2:

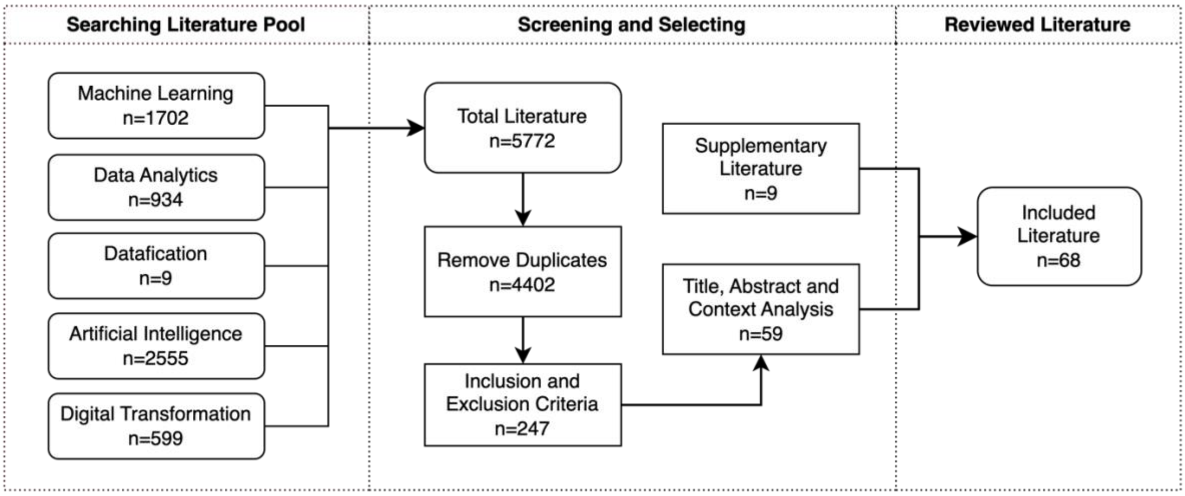

Figure 2: Process of identifying academic research literature

Table II shows a range of journals with varying publication counts, reflecting diversity in the sources of white literature. It indicates that *International Journal of Project Management*, *Project Management Journal*, *Automation in Construction,* and *Journal of Construction Engineering and Management* are the most prominent journals among the listed ones. Journals with at least 3 publications are also considered influential or relevant in the field. The 'Others' category indicates the presence of a wide range of additional sources, each contributing a single publication to the literature. It includes, among others, *IEEE Transaction on Engineering Management, Engineering Management Journal, Construction Management and Economics*.

Table II: Number of publications by journal

Journal	Number of publications
International Journal of Project Management	12
Project Management Journal	9
Automation in Construction	5
Journal of Construction Engineering and Management	5
International Journal of Managing Project in Business	3
International Journal of Construction Management	3
Journal of Management in Engineering	3
Journal of Building Engineering	3
Sustainability	3
Others	22

To ensure a minimum level of rigour in the included items the grey literature was confined to:

1. major consulting firms that featured on the Financial Times Top Management Consultants List (FinancialTimes, 2021) and niche control consulting firms specialised in projects
2. professional project associations
3. business press with a global reputation

Searches were carried out by visiting the organisation's website and using its own search functionality. The resultant core literature contained 101 items with 68 items identified from white literature and 33 from grey literature. The supplementary document gives references of all items in the core literature.

3.3 Reviewing the Status of Data Analytics in Project Delivery

An immediate observation on reviewing the core literature is its scant nature. From the large pool of items that have been written on data analytics and related constructs, 101 items were identified that specifically related to project delivery using the inclusion and exclusion criteria specified in section 3.1. This indicates the nascent nature of applying data analytics to project delivery and, in crude terms, the presence of far more 'gap' than 'knowledge.'

The conceptual structuring model devised at the start of the integrative review process helped unravel the differences between the approaches taken by the white versus the grey literature and hence by academic research versus (industrial) practice (see Figure 3). Most academic research concentrates on applying machine learning to predict project delivery performance, whereas practice focuses on other aspects of data analytics and artificial intelligence. This suggests that the discourse among practitioners is not benefitting from the learning held in academia on predictive machine learning in projects and, *vice versa*, that the project delivery research is not addressing the more holistic aspects of data analytic implementation in projects.

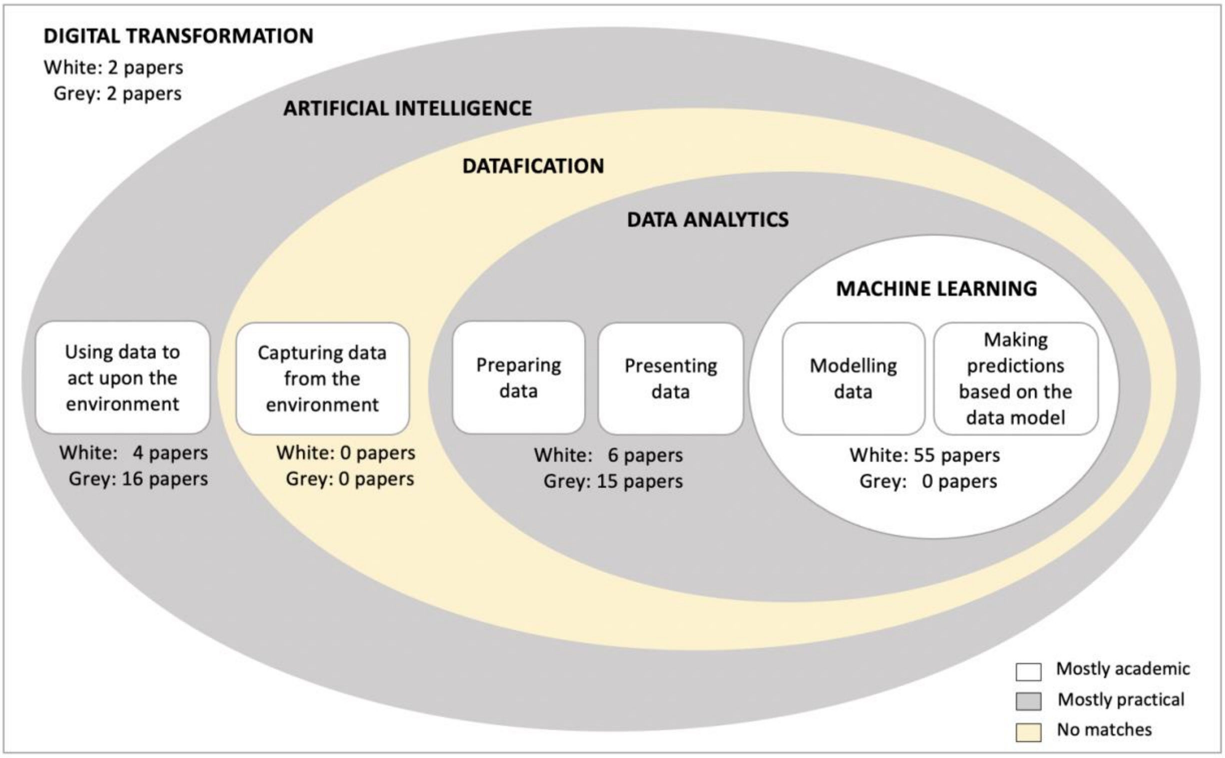

Figure 3: Distribution of papers in the integrative review core literature

We use the 'onion skin' conceptual model derived in the first stage of this integrative review (see Figure 3) as a framework to structure the reflections on data analytics in project delivery. These are summarised in Table III and expanded upon in the ensuing sections.

Table III: Using the Conceptual Structuring Model to Capture the Current Understanding of Project Data Analytics (PDA)

<table>
<tr><th></th><th>Functional Element of Conceptual Structuring</th><th>White Literature

(i.e. Academic Research)</th><th>Grey Literature

(i.e. Practice)</th></tr>
<tr><td rowspan="3">Encompassed by PDA</td><td>Machine Learning</td><td>• Over 50 papers described in detail using machine learning to predict performance.
• The literature is disparate and comprising non-referential stand-alone studies. Very few meta-studies exist that synthesise research on such aspects as algorithm selection.
• No attempt is made to identify which machine learning algorithms (c.f. support vector machines, neural networks, regression models) perform optimally in the project management domain.
• The literature is based on synthetic or constraint datasets that are not sufficiently similar to real-world data of reflecting the characteristics of large scale project data typically encountered in practice.
• With the exception of some attempts relating to Construction and Software, no extendible predictive characteristics (i.e. validity across more than one context) were identified.</td><td>• Some examples of simple statements that machine learning could be used predictively in a project delivery context were found.
• No descriptions of how machine learning is being applied or could be applied could be identified.</td></tr>
<tr><td>Preparing data</td><td colspan="2">• The body of both white and grey literature about data preprocessing is very scant.
• Both white and grey literature offer very limited guidance on data quality improvement and benchmarks enabling successful analysis in a project context.</td></tr>
<tr><td>Presenting data</td><td colspan="2">• There was no guidance on what are the most effective types of data to assist decision-making for effective project delivery.
• There was no guidance for the best way to present data to assist in decision-making.</td></tr>
</table>

PDA		• An extremely scant body of literature that does contain case examples of data analytics	• A substantive body of literature that speculates on what impact PDA will have on project delivery is using hyperbole and lacks reference to the current maturity of PDA. • Identified benefits include improved project delivery performance and its better understanding and ability to make estimates. • Empirical evidence on the use of PDA is minimal. • Guidance on adopting PDA is very limited.
Encompassing PDA	Artificial Intelligence	• Sources addresses AI are extremely rare and remain speculative in that they lack any empirical research.	• While a substantive body of literature addresses the impact PDA (as a component of AI) will have on project delivery, it remains speculative.
	Datafication	• There were no references to datafication in a PDA context in the white literature and only one isolated use of the term one grey literature source.	
	Digital Transformati on	• Two papers discussed digital information as the description of the project outcome (e.g. Building Information Modelling), but was not addressed in terms of project delivery data management and analysis itself.	• Two papers that made a very brief reference to digital transformation and data analytics.

Machine Learning: Most of the reviewed white literature were use cases for using machine learning techniques to assist project delivery. The first point to note about the specific studies is that the datasets used were either synthetic or consisted of 'real-world data' specifically curated for the investigation in question. Secondly, these datasets were not routinely generated by participating project organisations. They did not address any challenges they faced in getting data from practice in a usable format. Thirdly, the datasets contained data on typically less than one hundred projects. In practice, the number of projects in relevant datasets is often much higher.

Papers used a variety of machine learning algorithms in their investigation. Examples were found of Support Vector Machine (Cheng & Hoang, 2014; Cheng *et al.*, 2010a; Cheng *et al.*, 2010b; El-Sawalhi, 2015; Kong *et al.*, 2008; Petruseva *et al.*, 2017; Petruvesa *et al.*, 2013; Wauters & Vanhoucke, 2014), Neural networks (Darko *et al.*, 2023; Kim *et al.*, 2004; Naik & Radhika, 2015; Petroutsatou *et al.*, 2012; Wang *et al.*, 2012; Yahia *et al.*, 2011), Random forest (Shoar *et al.*, 2022; Yaseen *et al.*, 2020) and Regression (Kim *et al.*, 2004; Kim *et al.*, 2013; Mahmoodzadeh *et al.*, 2022; Petruseva *et al.*, 2017; Yip *et al.*, 2014). No consistent findings were evident in what type of algorithm worked best given particular characteristics of the investigation dataset in conjunction with specific predicted characteristics, size or sectoral base.

Overall, the white literature was successful in demonstrating that a wide variety of characteristics could be used to predict project delivery outcomes for those projects in terms of cost, duration and risk. *Prima facie*, this is an astounding finding, because understanding what reliably predicts project delivery performance is a 'holy grail' for those concerned with project delivery. However, no attempt was made to extend predictors beyond the highly contextualised area in which they were derived. This is amounts to a limitation of the current body of literature since sectors of project work were not equally represented with over a third of the core items being based in the construction sector. It is also important to note that the 'white' literature came from a side variety of journals including many not associated with Project Studies. This is important when considering whether the current understanding of Project Studies needs to be extended to encompass work on data analytics.

In the grey literature, mention was made of the applicability of machine learning but was not developed beyond that. In particular, none of the items explained the application of machine learning to predicting project delivery performance.

'Preparing Data' and 'Presenting Data': No reference was found in white or grey literature to the need to prepare data for modelling or to explain how the results of descriptive or predictive analytics should be presented to improve decision-making in project delivery. This suggests that neither practice nor academia is investigating how to curate real-world project delivery data for use in project analytics and how the level of data curation would impact the results of decision-making processes. These requirements are foundational to the use of data analytics in project delivery, and it is impossible to see how project data analytics can be used without such seminal knowledge.

'Data Analytics' and data analytical considerations in 'Artificial Intelligence': In contrast to the locus of activity on machine learning, grey literature was far more evident than white in dealing with holistic considerations of data analytics to project delivery. A substantive body of literature was found that referred either project data analytics in its own right or as a subset of artificial intelligence in projects. In some articles, project data analytics had actually been 'mis-labelled' as artificial intelligence. The functional stance of the onion skin model was instrumental in identifying these situations. Authors talked about artificial intelligence in terms of providing greater insight or improved predictive capability, both of which functions are encompassed by project data analytics.

The first point to note was that most of the grey literature referred to speculative benefits that project data analytics (either in its own right or through the mechanism of artificial intelligence) would bring to project delivery. This assessment of benefits appeared to be based on no empirical review of current applications but simply used 'thought experiments' to envisage what would happen to project delivery if project data analytics was more widely adopted (Schemelzer, 2019). Speculative benefits included very general statements on project delivery improvements and some, more specific, claims to be able to make more accurate estimates in project delivery and to be able to gain better insight through using technologies such as machine learning. With the exception of one item (Brookes & Lattuf, 2021), no attempt was made to quantify the benefits of project data analytics in financial terms.

The widespread speculation of benefits contained in the grey core literature contrasts with the dearth of empirical studies that capture the use of project data analytics. Extremely limited items were located that gave use-cases for project data analytics and these were greatly lacking in detail. Only two (Logikal, 2021; PMI, 2019) items gave a quantitative estimation of the use of project data analytics. These quantitative responses were so divergent (one claimed an 81% usage of project data analytics in organisations surveyed and the other less than 5%) that neither can be relied upon without a much greater understanding of the survey methodology used than is provided with these items. The lack of empirical reporting may explain a similar dearth in 'how-to' studies which give advice on implementing project data analytics. Eight items were found that gave very broad and generalised advice on how to implement project data analytics which had no justifying provenance and appeared simplistic. A much narrower body of literature was found that gave detailed, structured, and well-founded advice on implementing project data analytics. This literature was exemplified by the CAPSTONE framework developed by Accenture in their report 'Building More Value With Capital Projects: How to drive returns through data-driven digitization.' (Accenture, 2020).

White literature in this area comprised limited of papers which, in much the same vein as the grey literature in this area, confines themselves to either speculative theoretic predictions of benefits or asking practitioners to predict benefits and report on the outcomes of this. Notable exceptions were the works that did consider data more broadly; though their focus was less concerned with project delivery data, they were more concerned with the data involved in models of the project output, such as BIM (Whyte, 2019; Whyte *et al.*, 2016). The adoption of data analytics in BIM-related projects has gained substantial traction as more data is generated and managed within BIM (Garcia *et al.*, 2022). Integrating Artificial Intelligence into BIM requires careful data analytics consideration such as data collection, data preprocessing, feature selection, model development, etc (Abdulfattah *et al.*, 2023).

'Datafication' and 'Transformation': There is a dearth in both grey and white literature in the consideration of project data analytics in datafication or transformation. When considered *en masse*, the review findings suggest that data analytics in project delivery is under-researched by both the academic and practice communities. The lack of empirical investigations in both white and grey literature provides evidence that data analytics is not yet widely applied to project delivery. No knowledge has been created on how to prepare and present data in the specific context of project delivery analytics. Both of these functions are crucial to the effective implementation of data analytics in this field. Research does exist, in white literature, on the use of machine learning but no extendible findings on the type of algorithm or predictive characteristics have been generated. Some grey literature does address the application and benefits of project data analytics, but this knowledge does not appear to have been created systematically or using empirics. In summary, our review suggests that there are critical knowledge gaps that need to be 'plugged' if full advantage is ever to be taken in applying data analytics to project delivery.

3.4 Developing a Research Agenda

Section 3.3 identified trenchant 'knowledge gaps' in the application of data analytics to project delivery. The first step in 'plugging' those gaps would be to articulate a succinct research agenda. We focussed our agenda on those activities encompassed in our 'onion-skin' framework by data analytics. Going beyond the research questions concerning the nature and potential of PDA itself, we also as questions about holistic aspects of implementing project data analytics relevant for guiding higher-level decisions about whether to invoke PDA and how to implement it in 'real-world' industrial settings. Where possible, we root our agenda's suggestions in specific inadequacies of extant project data analytic knowledge as revealed in our integrative review and arrived at the research agenda summarised in Table IV which is expanded upon in the subsequent section. We also provided more specific details, methodologies, and examples related to each question would make it more comprehensive and actionable for future research.

Table IV: Developing a Research Agenda for Project Data Analytics (PDA)

		Research Questions	**Methodologies and Examples**
Encompassed by PDA	**Machine Learning**	• What are the project characteristics that can predict project delivery performance in an extendible fashion (i.e. across contexts)?	To comprehensively address these two questions, an exhaustive literature review can be conducted. Meta-analysis is useful to identify commonalities and variations in project characteristics across diverse industries and contexts. Researchers can undertake a comparative analysis of different machine learning algorithms, employing historical project data as the basis for evaluation.
		• Which machine learning algorithms are most appropriate for predicting project delivery characteristics and how can this be established?	

		• How do we know that a project for which we want to predict performance belongs to the same reference class as the predictive model?	A reference class framework based on project characteristics is valid in addressing this question. Subsequently, statistical tests need to be employed to validate the relevance of the reference class to prediction accuracy.
	Preparing	• How can the problems of data quality inhibiting project delivery be overcome?	A case study or cross case analysis approach can be employed to investigate projects with data quality issues, implementing data cleansing techniques. The goal is to provide practical insights into overcoming data quality challenges and enhancing the accuracy of project delivery predictions.
		• How can different types of data be integrated?	To address the integration of different data types, it is necessary to explore advanced techniques such as schema matching and ontology alignment. By combining structured project management data with unstructured textual data from project reports, it can demonstrate how ontology alignment can create a unified dataset for analysis.
		• What data assumptions are made and do they restrict the range of potential results?	A sensitivity analysis with statistical tests is effective to identify critical data assumptions and their impact on predictions. By varying assumptions about data completeness or accuracy, the effect on the outcomes of predictive models can be observed and reported. Statistical tests are be applied to compare model performance under different data assumption scenarios.
		• Do the data assumptions hinder the ability to produce accurate results or the ability to interpret the results?	
	Presenting data	• How can the temporal nature of projects be captured by data visualisation?	A mixed-methods approach can be used to ensure a comprehensive exploration of the complex dynamics involved in these research questions. The quantitative aspect of the research will involve data collection through surveys, experiments, and quantitative analysis to assess the effectiveness, impact, and preferences associated with various data visualization techniques, balanced visual representations, and data presentation formats. Simultaneously, the qualitative aspect will include in-depth interviews, focus groups, and qualitative analysis to capture subjective experiences, perceptions, and contextual insights from stakeholders.
		• How can a balance reflecting the different levels of importance between dimensions measuring successful project delivery be communicated visually?	
		• What is the most transparent and effective way to present data to facilitate evidence-based project delivery decision-making?	
Holistic aspects of PDA		• How is Data Analytics currently being conducted in project delivery? ○ Which sectors? ○ What type of projects? ○ What benefits have been accrued from its use?	A comprehensive survey and interviews across industries can be conducted to gather data analytics practices in project delivery. Surveying and interviewing project managers and analysts in various sectors enable researchers to understand the adoption of data analytics, types

		of projects involved, and the perceived benefits achieved.
	• What are the barriers and enablers to adopting PDA and how should they be overcome?	To comprehensively address this question, a mixed-methods study combining surveys, interviews, case studies, or Delphi study can be conducted to identify barriers and enablers.

Research Questions relating to Machine Learning: The integrative review revealed that, unlike for other aspects related to data analytics, there was a body of research that had used machine learning to predict project delivery performance. However, these studies were highly contextual and essentially non-referential, i.e. they rarely learned from each other or aim to produce findings that were extendible beyond their particular environment. Meta-studies are therefore required that take the experiences across a wide range of project contexts to establish:

- What could be a construct of a 'best' algorithm and how can this be evaluated?
- What machine learning algorithms work best in the context of project delivery (c.f. SVM, random forest, neural networks etc.)?
- What patterns exist in the identification of predictive characteristics using these algorithms?

To answer the first question, customarily, 'best' is associated with 'most predictive' but other aspect of the algorithm such as speed of use, fairness, and the utility of format of results may be equally important in a project context. The second question gets tantalisingly close to the 'Holy Grail' of much extant work in Project Studies concerned with establishing project success factors. Machine learning provides a mechanism to interrogate thousands (if not hundreds of thousands of projects) to establish statistically significant results. However, any comparison of algorithms depends on how the pool of projects in the dataset has been selected and the robustness across a variety of project types and data sources is key to deriving a universally applicable methodology.

Even through some studies explored the adoption of reference classing forecasting (Batselier & Vanhoucke, 2016; Natarajan, 2022), the construct of 'reference classes' is still a limitation of the use of machine learning not explored in either the grey or white literature (Whyte *et al.*, 2016). If the project is very novel, the machine learning prediction will not be effective as it will not match existing references classes. Mechanisms for recognising and reporting that a project is too novel to be predicted will need to be adopted.

Research Questions Relating to Data Preparation: Having highlighted the predominant use of relatively artificial data in the substantive work carried out on developing machine learning models in project data we encourage the use of 'real-world' data reflecting data dependencies and data quality issues typical for project delivery data. It tends to have many missing values and outliers due to errors in data entry rather than statistical phenomena. Data can also contain collinear characteristics (i.e. a large number of descriptive characteristics that appear to be attributes of the same factor.) Furthermore, real-world data often contains multi-modal data such as a combination of categorical (e.g. region, type of project) and numerical data (e.g. linear size, time duration for a particular aspect). Similar issues have been discussed extensively in other fields, for example spatial ecology and biomedical (Chapman, 2005; Schmidt, 2021), and we propose to develop a taxonomy of data quality problems along with toolbox to assess them and to fix them tailored to Project Studies. Many types of machine learning algorithm require that all data in entered as a numeric value. No research has yet been conducted on how to undertake this translation in a project delivery context but this question is vital for the effective use of machine learning. Besides, it is common to make data assumptions based on specific PM knowledge or prior experiences to restrict a wide range or simplify the data analytics process. These assumptions are highly susceptible to various use-cases. We introduced the assumptions There is no research can be found that is focus on if the data assumptions hinder the ability to produce an accurate result or on explaining to what extent the assumptions are different from the real project delivery.

Research Questions Relating to Data Presentation: The integrative review suggests that no work is being undertaken in practice or in the research community to investigate the best format and presentation to facilitate insight into modifiable factors for project performance and to impact on project delivery decision-making. This is despite the long-recognised use of the importance of data format in decision making in general (Lucas & Henry, 1981), and numerous investigations being undertaken in supporting decision-making other sectors and in management as a whole whole (Araz *et al.*, 2020; Bach *et al.*, 2019; Franklin *et al.*, 2017).

The longitudinal nature of project data and the interplay between decision-making and intermediate outcomes require custom-made visualisation formats not covered by the usual palette of graphs used in business analytics. Effective data visualization increases understanding. For instance, animated Gantt charts in project planning provide a dynamic representation of project timelines, offering a visual narrative of progress and interdependencies. The use of animation allows stakeholders to track changes over time, identify critical paths, and comprehend the evolving nature of project tasks. Additionally, interactive features, such as zooming or time manipulation, contribute to a clearer communication of project timeline. Visualized Gantt chart data can serve as input for various tools or techniques, including Machine Learning, to support decision-making (Palombarini & Martínez, 2022). To facilitate data-driven project decision-making it is vital to develop optimal formats for data reporting and visualisation. Researchers can draw upon well-established research from other arena including visual representation and interaction techniques (Thomas & Cook, 2006) and take inspiration from the Tufty's data visualisation philosophy summarised in (Tufty, 2001).

Research Questions Relating to Holistic Aspects of Project Data Analytics: The status review revealed that few exemplars of the extant use of data analytics applied to projects have been reported. While failure to be reported does not necessarily mean it does not exist, this gap strongly suggests that very little exploration of what is happening in this arena has taken place by academics or practitioners. Interestingly, where such exploration has been undertaken by practitioners, radically different conclusions on the prevalence of project data analytics have been produced. White literature suggests a sectoral slant in terms of greater uptake of analytics in construction and software project delivery.

An empirical, systematic review of the role project data analytics from a holistic project management perspective would be a vital foundational step in understanding further the structuring of the research agenda and the dissemination of best-practice. This would need to encompass putative taxonomies of project type and ensuing benefits in addition to considerations of sector. Similar reviews have already been undertaken in white literature for supply chains in financial services (Hung *et al.*, 2020), pharmaceutical chemistry (Nguyen *et al.*, 2021), and automotive (Smith *et al.*, 2019) implying that reliable methodologies for such explorations are already in existence.

In summary, by using an integrative review we have developed a novel and potentially transformational research agenda for data analytics in project delivery. It is important, therefore, to understand how consider the reception of this research agenda by the Project Studies research community. How does it fit with existing Project Studies research? Will Project Studies need a 'paradigm shift' to encompass studies in project data analytics?

4. Discussion: does research into data analytics in project studies demand a new project studies paradigm?

4.1 Does data analytics research fit into project studies?

The status review that we have carried out in this paper has clearly identified a dearth of knowledge in applying data analytics to project delivery both in terms of the activities that comprise data analytics in this arena and the activities that encompass data analytics. We know very little about how to garner reliable project data, how to optimally analyse and learn from these data, and how to change our existing management decision-making to take advantage of them. Research on project delivery has a crucial role in addressing these knowledge needs to enable project practice to fully benefit from the transformative benefit of data analytics. We therefore deem it vital to explore if Project Studies will be able to address these needs and understand where it will need to grow to match them. A useful place to start this exploration is in a current understanding of the paradigm of project studies.

The paradigmatic nature of project studies has been the focus of much debate in its associated research community in recent years (Sydow & Braun, 2018; Tekic *et al.*, 2022; Turner & Müller, 2003; Wawak & Woźniak, 2020). Geraldi has made a monumental contribution to this understanding, most recently in her co-editorship of special issues in PMJ focussing specifically in this area (Geraldi *et al.*, 2021). In her 2018 article with Söderlund, Geraldi provides us with a diffuse explanation of project studies. Project studies are described as (Geraldi & Söderlund, 2018):

> "equivalent to temporary organising and a social science discipline. When considering its relationship with practice, very much one of lived experience"

On a *prima facie* basis, this broad delineation of project studies does not yet provide sufficient room for the research agenda developed by this paper. The field of Project Studies has been shaped by a "dominance of ontological subjectivism and epistemological interpretivism, with a preference for case studies and qualitative methods" (Biedenbach & Müller, 2011). This stance is also reflected in the major journals in the Project Studies field. *International Journal of Managing Projects in Business* (IJMPB) restricts its scope to "highlight models and structures that advance the interests, dignity and well being of all stakeholders, in a sustainable manner" (IJMPB). Coincident with these developments (and mirroring movements found in general management research), Project Studies is concerned with its ability to deliver research that matters for and tangibly interacts with the world of practice but this is still conceptualized in terms that do not include data analytical research. In this spirit, Blomquist et al (2015) still place research that matters firmly in the sphere of social science research (Biedenbach & Müller, 2011). This perspective does not conceptualize that the latest developments for 'research that matters' both to academics and practitioner may come from a field such as data analytics. On the basis of the definitions of Project Studies accepted within its research community and evidenced by the attitudes the majority of its key journals, we must state that the dominant paradigm of Project Studies to date is that of an adjunct of social science and, more specifically, organisation studies.

The research agenda captured in Table IV offers elements that relate to organisation studies in terms of behaviour, change, and innovation within projects. Understanding the barriers and enablers of project data analytics, for example, is achievable using the current paradigm of Project Studies. (It relates very closely to work already published in *IJPM* (Howard *et al.*, 2017).) However, other parts of the research agenda we have articulated (i.e. machine learning using project delivery characteristics, improving data quality, most useful data presentation formats) would not be considered within the scope of Project Studies.

We posit that this potential exclusion is problematic. The interdisciplinary nature and theoretical diversity in Project Studies make this field inherently pluralistic and adaptable. However, Project Studies exhibit an excessive inclination towards traditional paradigms, hindering the embrace of data analytics. Conversely, data analytics specialists may overlook the expertise of project practitioners. Bridging this gap through collaboration is crucial for more effective project outcomes. The implementation of project data analytics in projects demands the mastery of both its 'behavioural' and 'technical' elements and, in all probability, will only be achieved by thinking about both at once. Our findings indicate that there is currently a lack of connection between data analytics and Project Studies. It is necessary to boost the 'technical' perspective that has not received enough attention. Data analytic challenges must become a vital part of the Project Studies remit, especially since a number of challenges in project data analytics are peculiar to the context of projects and are of no interest to the wider data analytics community. We posit that Project Studies needs to undergo, if not a paradigm shift, a transformative paradigm stretches so that issues of utmost importance to future project delivery can be considered within its remit.

4.2 How can project studies adapt to include data analytical research?

In seeking to understand how to enlarge the remit of project studies to include data analytics, inspiration can be found from other research fields that have encountered similar issues. Given the universality of datafication, it is unsurprising that other research fields have realised the advantages that data analytics can bring and are actively employing them in their research. Numerous examples exist of where existing fields of endeavour have embraced the insights that data analytics bring. These fields include prescriptive and descriptive analytics in financial services research (Andriosopoulos *et al.*, 2019); plant and precision agricultural (Olsen *et al.*, 2019; Talaviya *et al.*, 2020; Williamson *et al.*, 2021); ecology (Fink *et al.*, 2014); pandemic monitoring (Dutta *et al.*, 2021; Fiore *et al.*, 2019; Haycock *et al.*, 2020); circadian disruption in shift workers (Zhang *et al.*, 2022); urban analytics (Chohlas-Wood *et al.*, 2015; Lee, 2022); post-disaster human displacement (Oxford, 2021); electricity demand forecasting (Farrokhabadi *et al.*, 2022). Many of these achievements were enabled by collaborations including teams consisting of both domain specialists and data analytics experts.

The inclusion of data analytics in medical research provides an interesting illustration of the stretching of an existing research paradigm to include data analytics. It relates to medical applications (Panesar, 2019) and in particular disease prediction (Ali *et al.*, 2020; Bhatla & Jyoti, 2012; Diwakar *et al.*, 2021; Marimuthu *et al.*, 2018; Sharmila & Chellammal, 2018) and patient monitoring (Haeberle *et al.*, 2019; Lévi *et al.*, 2020). The shifts towards evidence-based medicine and personalized healthcare have prompted the integration of data analytics in medical research and patient care. Interdisciplinary collaboration between healthcare professionals, data scientists, and statisticians has led to the development of predictive models for disease diagnosis, treatment optimization, and patient outcomes. Data analytics is seen as an invaluable mechanism in performing medical research and the wider ethical implications of these approaches are now subject to widespread review. In industrial practice, some sectors that have successfully incorporate advanced analytical methods with the benefit from interdisciplinary collaboration. In ride-hailing services, for instance, analyzing vast datasets requires statisticians for accurate insights, programmers for developing robust algorithms and data base management, and domain experts for understanding the nuances of the transportation industry. Drawing parallels, Project Studies can benefit from a similar interdisciplinary approach by collaborating with experts in data analytics, information technology, and related fields. This interdisciplinary synergy ensures a comprehensive skill set, encompassing statistics, programming, and domain knowledge, ultimately contributing to more robust and insightful project outcomes.

Whilst much work on data analytics in medicine is still published in general computer science journals, this dramatic shift in research approach is increasingly being published in mainstream journals. In the case of medical research this includes the highly impactive journals such as The Lancet (Wiegand *et al.*, 2019) and the New England Journal of Medicine (Rajkomar *et al.*, 2019), and has triggered the launch of newer journals to cover the space, e.g. eBioMedicine and eClinicalMedicine (since 2022 both part of The Lancet's *Discovery Science* suite). Perhaps the most useful starting point in understanding how data analytics should be incorporated into Project Studies would be a special issue of one of the major journals in this field examining this precise issue.

5 Conclusions

5.1 Achievement of aims

The aim of the investigation reported in this paper was two-fold. Firstly, to critically review the current application of data analytics to project delivery from both a practitioner and research stance and thus arrive at an understanding of the status of activity in this space. Secondly, to use that review to develop a research agenda that will assist in understanding and overcoming the data analytical challenges in a project delivery context.

To address the first aim, we developed a novel integrative review methodology that would work in a nascent research area and applied this to data analytics in the context of project delivery. Fundamental to this approach was the development of an 'onion-skin' conceptual structuring framework. This framework is a valuable contribution that could be used in separation from the review to assist discussions in the contested and unstructured arena of what comprises data analytics in the world of project delivery.

We used our 'onion-skin' framework to generate the search terms for reviewing both white and grey literature on data analytics to arrive at a 'status' of the knowledge surrounding this area. This revealed that apart from some pockets of activity surrounding machine learning, there was scant consideration of the other functions such as data preparation or data presentation or of any holistic appreciation of the benefits of data analytics in projects that would be vital to implementing data analytics in project delivery. We developed a research agenda that addressed this dearth of knowledge by combining the results of our review with insights from other fields that had applied data analytics and through our own experiences in our initial scoping studies.

We also used our 'onion-skin' framework to structure our proposed research agenda and devised specific responses to areas of the framework currently underserved by white or grey literature. The agenda gave specific requirements for machine learning for algorithm and variable selection, for data analytics in terms of data preparation and presentation, and for empirical investigation into the actuality of implementing data analytics in project delivery. We then highlighted that though the later item fits into the current manifestation of Project Studies, the more analytical elements do not. We reflected on how the experience of other fields in facing these analytical challenges could assist Project Studies toward achieving the same goal.

5.2 Limitations of our approach

Our research may be limited because we may not have not captured all the literature surrounding the phenomenon of project data analytics. Our suppositions are founded upon our integrative literature review and therefore the warranty of those suppositions depends on the inclusivity of the literature contained in our review. Ensuring the capture of all the relevant literature in a nascent field such as project data analytics is extremely difficult. It is inevitable that we will have missed some of the pertinent literature that relates to the phenomenon that we are investigating. We would argue, however, that the systematic and structured way we searched for literature will have minimised this problem. Furthermore, this methodology also enables our body of literature to be representative enough that any new literature incorporated within it would not disturb the findings of our overall analysis. We posit that the conceptual structure of project data analytics and its associated knowledge gaps would not change even if further items were added.

5.3 Contributions and further work

We deem that the outputs of this paper that can assist further work comprise:

- novel review methodology that could be applied in any nascent area with Project Studies
- a conceptual framework that removes confusion surrounding the terminology involved in data analytics in project delivery

- an understanding of the general scant nature of activity in project data analytics and summaries of where most activity in located
- a research agenda addressing the knowledge gaps in using data analytics in project delivery
- suggestions on how the Project Studies could deliver that agenda given its analytical nature

Future researchers could address each of these aspects. The review of application of new technologies tends to always be a nascent field creating a need for the development of structured review methodologies in these circumstances. The 'onion-skin' model we developed withstood the test of being used to structure a status report and may be worthy of further development and use by other researchers across disciplines. Most importantly, our research agenda can be used to further the development of project data analytics. Through the appreciation of the 'as-is' status of data analytics in project delivery and through the development of a research agenda, we hope that we have bought the day forward when data supports project delivery in a way that benefits individual organisations and society as a whole.

The research findings have significant implications for practitioners by providing tangible action points to improve project management efficacy and address critical considerations in utilizing data for project delivery:

- Emphasize systematic and purpose-driven data collection to establish a comprehensive foundation for analysis. Align captured variables with specific questions to enhance relevance.
- Prioritize the careful selection of variables to determine the scope and depth of insights, especially in predicting project outcomes.
- Formulate a robust data collection and storage plan to serve as a strategic roadmap for gathering, processing, and preserving data throughout the project lifecycle.
- Implement routine data quality checks or a comprehensive data quality assessment framework to ensure accuracy and reliability. Regularly evaluate completeness, coherence, validity, and currency.

Reconnecting this with the popular "new oil" metaphor mentioned in the opening, we must stress that data, as most resources, need human expertise and effort to exploit their value. They present numerous challenges that require attention from the Project Studies research community. This paper provides a blueprint for a bridge linking data analytics and project management.

References

Abdulfattah, B. S., Abdelsalam, H. A., Abdelsalam, M., Bolpagni, M., Thurairajah, N., Perez, L. F. & Butt, T. E. (2023) Predicting implications of design changes in BIM-based construction projects through machine learning. *Automation in Construction*, 155 105057.

Accenture (2020) *Building More Value With Capital Projects: How to drive higher returns through data-driven digitization*, available at: https://www.accenture.com/_acnmedia/PDF-142/Accenture-IndustryX-Building-Value-Capital-Projects.pdf

Adams, R. J., Smart, P. & Huff, A. S. (2017) Shades of grey: guidelines for working with the grey literature in systematic reviews for management and organizational studies. *International Journal of Management Reviews*, 19 (4): 432-454.

Afzal, F., Yunfei, S., Nazir, M. & Bhatti, S. M. (2019) A review of artificial intelligence based risk assessment methods for capturing complexity-risk interdependencies: cost overrun in construction projects. *International Journal of Managing Projects in Business*, 14(2): 300-328.

Aghimien, D. O., Ikuabe, M., Aigbavboa, C., Oke, A. & Shirinda, W. (2021) Unravelling the factors influencing construction organisations' intention to adopt big data analytics in South Africa. *Construction Economics and Building*, 21 (3): 262-281.

Al-Tabtabai, H. & Alex, A. P. (2000) Modeling the cost of political risk in international construction projects. *Project Management Journal*, 31 (3): 4-13.

Albogami, S. M., Ariffin, M. K. A. B. M., Supeni, E. E. B. & Ahmad, K. A. (2021) A New Hybrid AHP and Dempster—Shafer Theory of Evidence Method for Project Risk Assessment Problem. *Mathematics*, 9 (24): 3225.

AIPM (2022) *Project delivery 101 | A guide for project manage,* available at: https://aipm.com.au/blog/project-delivery-101-a-guide-for-project-managers/

Ali, F., El-Sappagh, S., Islam, S. R., Kwak, D., Ali, A., Imran, M. & Kwak, K.-S. (2020) A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. *Information Fusion*, 63 208-222.

Amer, F., Jung, Y. & Golparvar-Fard, M. (2021) Transformer machine learning language model for auto-alignment of long-term and short-term plans in construction. *Automation in Construction*, 132 103929.

Andriosopoulos, D., Doumpos, M., Pardalos, P. M. & Zopounidis, C. (2019) Computational approaches and data analytics in financial services: A literature review. *Journal of the Operational Research Society*, 70 (10): 1581-1599.

APM (2020) *Project data analytics: the state of the art and science*, available at: https://www.apm.org.uk/media/46977/pathfinder-report-web-final.pdf

APM (2021) *Getting Started in Project Data*, available at: https://www.apm.org.uk/media/50992/getting-started-in-project-data-web-final.pdf

Araz, O. M., Choi, T. M., Olson, D. L., & Salman, F. S. (2020). Role of analytics for operational risk management in the era of big data. *Decision Sciences*, 51 (6): 1320-1346.

Aubry, M., Boukri, S. E. & Sergi, V. (2021) Opening the black box of benefits management in the context of projects. *Project Management Journal*, 52 (5): 434-452.

Bach, K., Marling, C., Mork, P. J., Aamodt, A., Mair, F. S., & Nicholl, B. I. (2019). Design of a clinician dashboard to facilitate co-decision making in the management of non-specific low back pain. *Journal of Intelligent Information Systems*, 52: 269-284.

Batselier, J. & Vanhoucke, M. (2016) Practical application and empirical evaluation of reference class forecasting for project management. *Project Management Journal*, 47 (5): 36-51.

Benala, T. R., Mall, R., Srikavya, P. & HariPriya, M. V. (2014) Software effort estimation using data mining techniques. In *ICT and Critical Infrastructure: Proceedings of the 48th Annual Convention of Computer Society of India*-Vol I: Hosted by CSI Vishakapatnam Chapter (pp. 85-92). Springer International Publishing.

Bevort, D. (2020) How to Approach Artificial Intelligence for Project Management, available at: https://www.adeaca.com/blog/how-to-approach-artificial-intelligence-for-project-management/ (accessed 15 June 2020)

Bhatla, N. & Jyoti, K. (2012) An analysis of heart disease prediction using different data mining techniques. *International Journal of Engineering*, 1 (8): 1-4.

Biedenbach, T. & Müller, R. (2011) Paradigms in project management research: examples from 15 years of IRNOP conferences. *International journal of managing projects in business*,

Bilal, M., Oyedele, L. O., Kusimo, H. O., Owolabi, H. A., Akanbi, L. A., Ajayi, A. O., Akinade, O. O. & Delgado, J. M. D. (2019) Investigating profitability performance of construction projects using big data: A project analytics approach. *Journal of Building Engineering*, 26 100850.

Brookes, N. (2020) *AI to assist projects has been around for decades – why aren't we using it?*, available at: https://www.apm.org.uk/blog/ai-to-assist-projects-has-been-around-for-decades-why-aren-t-we-using-it/

Brookes, N., Lattuf Flores, L., Dyer, R., Stewart, I., Wang, K. & Dacre, N. (2020) *Project Data Analytics: The State of the Art and Science*. Association for Project Management (APM), available at: https://www.apm.org.uk/v2/media/ovylccxs/pathfinder-report-web-final.pdf

Brookes, N. & Lattuf, L. (2021) *(R)evolution: barriers and enablers to Project Data Analytics*, available at: https://weare.theoaklandgroup.co.uk/project-data-analytics-report

Chaphalkar, N., Iyer, K. & Patil, S. K. (2015) Prediction of outcome of construction dispute claims using multilayer perceptron neural network model. *International Journal of Project Management*, 33 (8): 1827-1835.

Chapman, A. D. (2005). *Principles of data quality*. GBIF.

Chattapadhyay, D. B. & Putta, J. (2021) Risk Identification, Assessments, and Prediction for Mega Construction Projects: A Risk Prediction Paradigm Based on Cross Analytical-Machine Learning Model. *Buildings*, 11 (4): 172.

Cheng, M.-Y., Cao, M.-T. & Mendrofa, A. Y. J. (2021) Dynamic feature selection for accurately predicting construction productivity using symbiotic organisms search-optimized least square support vector machine. *Journal of Building Engineering*, 35 101973.

Cheng, M.-Y., Chang, Y.-H. & Korir, D. (2019) Novel approach to estimating schedule to completion in construction projects using sequence and nonsequence learning. *Journal of Construction Engineering and Management*, 145 (11): 04019072.

Cheng, M.-Y. & Hoang, N.-D. (2014) Interval estimation of construction cost at completion using least squares support vector machine. *Journal of Civil Engineering and Management*, 20 (2): 223-236.

Cheng, M.-Y., Peng, H.-S., Wu, Y.-W. & Chen, T.-L. (2010a) Estimate at completion for construction projects using evolutionary support vector machine inference model. *Automation in Construction*, 19 (5): 619-629.

Cheng, M.-Y., Wu, Y.-W. & Wu, C.-F. (2010b) Project success prediction using an evolutionary support vector machine inference model. *Automation in Construction*, 19 (3): 302-307.

Cheng, M.-Y. & Roy, A. F. (2011) Evolutionary fuzzy decision model for cash flow prediction using time-dependent support vector machines. *International journal of project management*, 29 (1): 56-65.

Cheng, M.-Y., Wu, Y.-W. & Huang, C.-C. (2020) Hybrid Gaussian Process Inference Model for Construction Management Decision Making. *International Journal of Information Technology & Decision Making*, 19 (04): 1015-1036.

Cheung, S. O., Wong, P. S. P., Fung, A. S. & Coffey, W. (2006) Predicting project performance through neural networks. *International Journal of Project Management*, 24 (3): 207-215.

Chohlas-Wood, A., Merali, A., Reed, W. & Damoulas, T. (2015) Mining 911 calls in New York City: temporal patterns, detection, and forecasting.

Choi, S.-W., Lee, E.-B. & Kim, J.-H. (2021) The Engineering Machine-Learning Automation Platform (EMAP): A Big-Data-Driven AI Tool for Contractors' Sustainable Management Solutions for Plant Projects. *Sustainability*, 13 (18): 10384.

Chou, J.-S., Hsu, S.-C., Lin, C.-W. & Chang, Y.-C. (2016) Classifying influential information to discover rule sets for project disputes and possible resolutions. *International Journal of Project Management*, 34 (8): 1706-1716.

Cooper, J. O., Heron, T. E. & Heward, W. L. (2007) Applied behavior analysis.

Costantino, F., Di Gravio, G. & Nonino, F. (2015) Project selection in project portfolio management: An artificial neural network model based on critical success factors. *International Journal of Project Management*, 33 (8): 1744-1754.

Darko, A., Glushakova, I., Boateng, E. B. & Chan, A. P. (2023) Using Machine Learning to Improve Cost and Duration Prediction Accuracy in Green Building Projects. *Journal of Construction Engineering and Management*, 149 (8): 04023061.

Delgado, J. M. D. & Oyedele, L. (2021) Deep learning with small datasets: using autoencoders to address limited datasets in construction management. *Applied Soft Computing*, 112 107836.

Denicol, J., Davies, A. & Krystallis, I. (2020) What are the causes and cures of poor megaproject performance? A systematic literature review and research agenda. *Project Management Journal*, 51 (3): 328-345.

Dikmen, I., Birgonul, M. T. & Han, S. (2007) Using fuzzy risk assessment to rate cost overrun risk in international construction projects. *International journal of project management*, 25 (5): 494-505.

Diwakar, M., Tripathi, A., Joshi, K., Memoria, M. & Singh, P. (2021) Latest trends on heart disease prediction using machine learning and image fusion. *Materials Today: Proceedings*, 37 3213-3218.

Dutta, R., Gomes, S. N., Kalise, D. & Pacchiardi, L. (2021) Using mobility data in the design of optimal lockdown strategies for the COVID-19 pandemic. *PLoS Computational Biology*, 17 (8): e1009236.

Dyer, R. (2021) *Artificial intelligence for project lessons learnt*, available at: https://www.apm.org.uk/blog/artificial-intelligence-for-project-lessons-learnt/

Egwim, C. N., Alaka, H., Toriola-Coker, L. O., Balogun, H., Ajayi, S. & Oseghale, R. (2021) Extraction of underlying factors causing construction projects delay in Nigeria. *Journal of Engineering, Design and Technology*,

El-Kholy, A. (2021) Exploring the best ANN model based on four paradigms to predict delay and cost overrun percentages of highway projects. *International Journal of Construction Management*, 21 (7): 694-712.

El-Sawalhi, N. I. (2015) Support vector machine cost estimation model for road projects. *Journal of Civil Engineering and Architecture*, 9 (9): 846-859.

Ertuğrul, E., Baytar, Z., Çatal, Ç. & Muratli, Ö. C. (2019) Performance tuning for machine learning-based software development effort prediction models. *Turkish Journal of Electrical Engineering & Computer Sciences*, 27 (2): 1308-1324.

Farrokhabadi, M., Browell, J., Wang, Y., Makonin, S., Su, W. & Zareipour, H. (2022) Day-Ahead Electricity Demand Forecasting Competition: Post-COVID Paradigm. *IEEE Open Access Journal of Power and Energy*, 9, pp.185-191.

FinancialTimes (2021) *UK's Leading Management Consultants 2021: the ratings*, available at: https://www.ft.com/content/c02c6de5-7e9b-4fc2-b153-cbf5f02655ad (Accessed 18 February 2021)

Fink, A. (2019) *Conducting research literature reviews: From the internet to paper*. Sage publications.

Fink, D., Damoulas, T., Bruns, N. E., La Sorte, F. A., Hochachka, W. M., Gomes, C. P. & Kelling, S. (2014) Crowdsourcing meets ecology: hemisphere-wide spatiotemporal species distribution models. *AI magazine*, 35 (2): 19-30.

Fiore, S., Elia, D., Pires, C. E., Mestre, D. G., Cappiello, C., Vitali, M., Andrade, N., Braz, T., Lezzi, D. & Moraes, R. (2019) An integrated big and fast data analytics platform for smart urban transportation management. *IEEE Access*, 7 117652-117677.

Fox, S. & Do, T. (2013) Getting real about Big Data: applying critical realism to analyse Big Data hype. *International Journal of Managing Projects in Business*, 6 (4): 739-760.

Franklin, A., Gantela, S., Shifarraw, S., Johnson, T. R., Robinson, D. J., King, B. R., ... & Okafor, N. G. (2017). Dashboard visualizations: Supporting real-time throughput decision-making. *Journal of biomedical informatics*, 71: 211-221.

Fridgeirsson, T. V., Ingason, H. T., Jonasson, H. I. & Jonsdottir, H. (2021) An authoritative study on the near future effect of artificial intelligence on project management knowledge areas. *Sustainability*, 13 (4): 2345.

Garcia, J., Villavicencio, G., Altimiras, F., Crawford, B., Soto, R., Minatogawa, V., Franco, M., Martínez-Muñoz, D. & Yepes, V. (2022) Machine learning techniques applied to construction: A hybrid bibliometric analysis of advances and future directions. *Automation in Construction*, 142 104532.

Geraldi, J. & Lechter, T. (2012) Gantt charts revisited: A critical analysis of its roots and implications to the management of projects today. *International Journal of Managing Projects in Business*, 5 (4): 578-594.

Geraldi, J., & Söderlund, J. (2018). Project studies: What it is, where it is going. *International journal of project management*, 36 (1): 55-70.

Geraldi, J., Söderlund, J. & Marrewijk, A. v. (2021) Bright and Dark Spots in Project Studies: Continuing Efforts to Advance Theory Development and Debate. *Project management journal*, 52 (3): 227-236.

Goel, L., Sharma, M., Khatri, S. K. & Damodaran, D. (2021) An empirical analysis of the statistical learning models for different categories of cross-project defect prediction. *International Journal of Computer Aided Engineering and Technology*, 14 (2): 233-254.

Gondia, A., Siam, A., El-Dakhakhni, W. & Nassar, A. H. (2020) Machine learning algorithms for construction projects delay risk prediction. *Journal of Construction Engineering and Management*, 146 (1): 04019085.

Grushka-Cockayne, Y. (2020) Use Data to Revolutionize Project Planning, available at: https://hbr.org/2020/02/use-data-to-revolutionize-project-planning

Hamid, C., Brahim, B., Samir, B. & Houssaine, Z. (2020) THE MACHINE LEARNING IN PROJECT MANAGEMENT: A SYSTEMATIC MAPPING STUDY.

Holzmann, V., Zitter, D. & Peshkess, S. (2022) The Expectations of Project Managers from Artificial Intelligence: A Delphi Study. *Project Management Journal*, 87569728211061779.

Hovnanian, G., Kroll, K. & Sjödin, E. (2019) *How analytics can drive smarter engineering and construction decisions*, available at: https://www.autodesk.com/learn/ondemand/module/identify-strategies-for-application-of-advanced-construction-technologies/unit/6QG3e8OcpP2oSrzvBAWKM3

Innocent, M., Wasek, J. S. & Franz, A. (2018) Predicting military construction project time outcomes using data analytics. *Engineering Management Journal*, 30 (4): 232-246.

IPMA (2020) *Artificial Intelligence impact in Project Management*, available at: https://ipma.world/assets/IPMA_PwC_AI_Impact_in_PM_-_the_Survey_Report.pdf

Haeberle, H. S., Helm, J. M., Navarro, S. M., Karnuta, J. M., Schaffer, J. L., Callaghan, J. J., Mont, M. A., Kamath, A. F., Krebs, V. E. & Ramkumar, P. N. (2019) Artificial intelligence and machine learning in lower extremity arthroplasty: a review. *The Journal of arthroplasty*, 34 (10): 2201-2203.

Hammad, M., Abbasi, A., Chakrabortty, R. K. & Ryan, M. J. (2020) Predicting the critical path changes using sensitivity analysis: a delay analysis approach. *International Journal of Managing Projects in Business*, 13 (5): 1097-1119.

Haycock, C., Thorpe-Woods, E., Walsh, J., O'Hara, P., Giles, O., Dhir, N. & Damoulas, T. (2020) An Expectation-Based Network Scan Statistic for a COVID-19 Early Warning System. *arXiv preprint arXiv:2012.07574*,

Howard, R., Restrepo, L. & Chang, C.-Y. (2017) Addressing individual perceptions: An application of the unified theory of acceptance and use of technology to building information modelling. *International Journal of Project Management*, 35 (2): 107-120.

Humby, C. (2006) Data is the new oil. *Proc. ANA Sr. Marketer's Summit. Evanston, IL, USA*,

Hung, J.-L., He, W. & Shen, J. (2020) Big data analytics for supply chain relationship in banking. *Industrial Marketing Management*, 86 144-153.

Jin, X.-H. & Zhang, G. (2011) Modelling optimal risk allocation in PPP projects using artificial neural networks. *International journal of project management*, 29 (5): 591-603.

Jugdev, K., Perkins, D., Fortune, J., White, D. & Walker, D. (2013) An exploratory study of project success with tools, software and methods. *International Journal of Managing Projects in Business*, 6 (3): 534-551.

Karaca, I., Gransberg, D. D. & Jeong, H. D. (2020) Improving the accuracy of early cost estimates on transportation infrastructure projects. *Journal of Management in Engineering*, 36 (5): 04020063.

Khodakarami, V. & Abdi, A. (2014) Project cost risk analysis: A Bayesian networks approach for modeling dependencies between cost items. *International Journal of Project Management*, 32 (7): 1233-1245.

Killen, C. P. (2017) Managing portfolio interdependencies: The effects of visual data representations on project portfolio decision making. *International Journal of Managing Projects in Business*, 10 (4): 856-879.

Kim, G.-H., An, S.-H. & Kang, K.-I. (2004) Comparison of construction cost estimating models based on regression analysis, neural networks, and case-based reasoning. *Building and environment*, 39 (10): 1235-1242.

Kim, G.-H., Shin, J.-M., Kim, S. & Shin, Y. (2013) Comparison of school building construction costs estimation methods using regression analysis, neural network, and support vector machine. *Journal of Building Construction and Planning Research*. Vol.01, pp 1-7.

Kong, F., Wu, X.-j. & Cai, L.-y. (2008) A novel approach based on support vector machine to forecasting the construction project cost. in *2008 International Symposium on Computational Intelligence and Design*, Vol. 1, pp.21-24.

Krechowicz, M. & Krechowicz, A. (2021) Risk Assessment in Energy Infrastructure Installations by Horizontal Directional Drilling Using Machine Learning. *Energies*, 14 (2): 289.

Kumar, K. H. & Srinivas, K. (2021) Preliminary performance study of a brief review on machine learning techniques for analogy based software effort estimation. *Journal of Ambient Intelligence and Humanized Computing*, 1-25.

Lahmann, M. (2018) *AI will transform project management. Are you ready?*, available at: https://www.pwc.ch/en/insights/risk/ai-will-transform-project-management-are-you-ready.html

Lattuf, L. (2021) *Barriers to project data analytics and how to overcome them*, available at; https://www.apm.org.uk/blog/barriers-to-project-data-analytics-and-how-to-overcome-them/

Lawrence, A., Thomas, J., Houghton, J. & Weldon, P. (2015) Collecting the evidence: Improving access to grey literature and data for public policy and practice. *Australian Academic & Research Libraries*, 46 (4): 229-249.

Lee, T. (2022) *Getting the right equipment to the right people*, available at: https://engineering.berkeley.edu/news/2020/04/getting-the-right-equipment-to-the-right-people/ (Accessed 2 April 2020)

Leu, S.-S., Chen, A.-T. & Yang, C.-H. (2001) A GA-based fuzzy optimal model for construction time–cost trade-off. *International Journal of Project Management*, 19 (1): 47-58.

Li, H., Cao, Y., Lin, Q. & Zhu, H. (2022) Data-driven project buffer sizing in critical chains. *Automation in Construction*, 135 104134.

Li, H. & Love, P. E. (1999) Combining rule-based expert systems and artificial neural networks for mark-up estimation. *Construction Management & Economics*, 17 (2): 169-176.

Ligsuniversity (2022) *The Role of Artificial Intelligence in Improving Project Management*, available at: https://ligsuniversity.com/blog/the-role-of-artificial-intelligence-in-improving-project-management

Liu, L., Liu, D., Wu, H. & Wang, J. (2020) Study on Foundation Pit Construction Cost Prediction Based on the Stacked Denoising Autoencoder. *Mathematical Problems in Engineering*, 2020

Liu, S. & Hao, W. (2021) Forecasting the scheduling issues in engineering project management: Applications of deep learning models. *Future Generation Computer Systems*, 123 85-93.

Locatelli, G., Mikic, M., Kovacevic, M., Brookes, N. & Ivanisevic, N. (2017) The successful delivery of megaprojects: a novel research method. *Project management journal*, 48 (5): 78-94.

Logikal (2021) *2021 Integrated Project Controls & Digital Insights Report*, available at: https://logikalprojects.com/insights/2021-integrated-project-controls-digital-insights-report-2/

Lucas, J. (2021) Project Analytics: Benefits, Challenges and First Steps, available at: https://www.ecosys.net/blog/project-analytics-benefits-challenges-and-first-steps/ (Accessed 22 February 2021)

Lucas Jr, H. C. (1981). An experimental investigation of the use of computer-based graphics in decision making. *Management Science*, 27 (7): 757-768.

Lévi, F., Komarzynski, S., Huang, Q., Young, T., Ang, Y., Fuller, C., Bolborea, M., Brettschneider, J., Fursse, J. & Finkenstädt, B. (2020) Tele-monitoring of cancer patients' rhythms during daily life identifies actionable determinants of circadian and sleep disruption. *Cancers*, 12 (7): 1938.

López, L., Manzano, M., Gómez, C., Oriol, M., Farré, C., Franch, X., Martínez-Fernández, S. & Vollmer, A. M. (2021) QaSD: a quality-aware strategic dashboard for supporting decision makers in agile software development. *Science of Computer Programming*, 202 102568.

Mahdavian, A., Shojaei, A., Salem, M., Yuan, J. S. & Oloufa, A. A. (2021) Data-driven predictive modeling of highway construction cost items. *Journal of Construction Engineering and Management*, 147 (3): 04020180

Mahmoodzadeh, A., Nejati, H. R. & Mohammadi, M. (2022) Optimized machine learning modelling for predicting the construction cost and duration of tunnelling projects. *Automation in Construction*, 139 104305.

Marco, R. (2019) Optimizing Software Effort Estimation Models Based On Metaheuristic Methods: A Proposed Framework.

Marimuthu, M., Abinaya, M., Hariesh, K., Madhankumar, K. & Pavithra, V. (2018) A review on heart disease prediction using machine learning and data analytics approach. *International Journal of Computer Applications*, 181 (18): 20-25.

Marr, B. (2015). *Big Data: Using SMART big data, analytics and metrics to make better decisions and improve performance*. John Wiley & Sons.

Matel, E., Vahdatikhaki, F., Hosseinyalamdary, S., Evers, T. & Voordijk, H. (2019) An artificial neural network approach for cost estimation of engineering services. *International journal of construction management*, 1-14.

McGlynn, J. (2020) *What does AI and data mean for the project profession?*, available at: https://www.apm.org.uk/blog/what-does-ai-and-data-mean-for-the-project-profession/

McKinsey (2018) *How predictive analytics can boost product development*, available at: https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/how-predictive-analytics-can-boost-product-development

Mckinsey (2021) *Project delivery hubs: Analytics for better megaproject outcomes*, available at: https://www.mckinsey.com/capabilities/operations/our-insights/project-delivery-hubs-analytics-for-better-megaproject-outcomes

MPA (2021) *Programmes with purpose: delivering success in government's major projects*, available at: https://majorprojects.org/resources/programmes-with-purpose/

Naik, G. & Radhika, V. S. B. (2015) Time and cost analysis for highway road construction project using artificial neural networks. *KICEM Journal of Construction Engineering and Project Management*, 4 (3): 26-31.

Natarajan, A. (2022) Reference Class Forecasting and Machine Learning for Improved Offshore Oil and Gas Megaproject Planning: Methods and Application. *Project Management Journal*, , 53(5), pp.456-484.

Nemati, H. R., Todd, D. W. & Brown, P. D. (2002) A hybrid intelligent system to facilitate information system project management activities. *Project Management Journal*, 33 (3): 42-52.

Nguvulu, A., Yamato, S. & Honma, T. (2012) Project Performance Evaluation Using Deep Belief Networks. *IEEJ Transactions on Electronics, Information and Systems*, 132 (2): 306-312.

Nguyen, A., Lamouri, S., Pellerin, R., Tamayo, S. & Lekens, B. (2021) Data analytics in pharmaceutical supply chains: state of the art, opportunities, and challenges. *International Journal of Production Research*, 1-20.

Niederman, F. (2021) Project management: openings for disruption from AI and advanced analytics. *Information Technology & People*, *34*(6): 1570-1599.

Oesterreich, T. D., Anton, E., Teuteberg, F., & Dwivedi, Y. K. (2022). The role of the social and technical factors in creating business value from big data analytics: A meta-analysis. *Journal of Business Research*, 153: 128-149.

Olsen, A., Konovalov, D. A., Philippa, B., Ridd, P., Wood, J. C., Johns, J., Banks, W., Girgenti, B., Kenny, O. & Whinney, J. (2019) DeepWeeds: A multiclass weed species image dataset for deep learning. *Scientific reports*, 9 (1): 1-12.

Olsson, N. O. & Bull-Berg, H. (2015) Use of big data in project evaluations. *International Journal of Managing Projects in Business*, 8 (3): 491-512.

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.

Oxford (2021) Oxford researchers develop tool to predict human displacement post-disaster, available at: https://www.ox.ac.uk/news/science-blog/oxford-researchers-develop-tool-predict-human-displacement-post-disaster

Palombarini, J. A. & Martínez, E. C. (2022) End-to-end on-line rescheduling from Gantt chart images using deep reinforcement learning. *International Journal of Production Research*, 60 (14): 4434-4463.

Panesar, A. (2019) *Machine learning and AI for healthcare*. Springer.

Papadonikolaki, E., van Oel, C. & Kagioglou, M. (2019) Organising and Managing boundaries: A structurational view of collaboration with Building Information Modelling (BIM). *International Journal of Project Management*, 37 (3): 378-394.

Papadonikolaki, E., Krystallis, I. & Morgan, B. (2022) Digital Technologies in Built Environment Projects: Review and Future Directions. *Project Management Journal*, 87569728211070225.

Papatheocharous, E. & Andreou, A. S. (2012) A hybrid software cost estimation approach utilizing decision trees and fuzzy logic. *International Journal of Software Engineering and Knowledge Engineering*, 22 (03): 435-465.

Petroutsatou, K., Georgopoulos, E., Lambropoulos, S. & Pantouvakis, J. (2012) Early cost estimating of road tunnel construction using neural networks. *Journal of construction engineering and management*, 138 (6): 679-687.

Petruseva, S., Zileska-Pancovska, V., Žujo, V. & Brkan-Vejzović, A. (2017) Construction costs forecasting: Comparison of the accuracy of linear regression and support vector machine models. *Technical Gazette*, 24 (5): 1431-1438.

Petruvesa, S., Zileska, V. & Zujo, V. (2013) Predicting construction project duration with support vector machine. *International Journal of research in Engineering and Technology*, 11 (2): 12-24.

PDATaskForce (2020) *Transforming project performance with data: Unlocking the power of data to deliver project success*, available at: https://pdataskforce.com/wp-content/uploads/2020/09/Transforming-project-performance-with-data.pdf

PMI (2019) AI Innovators: Cracking the Code on Project Performance, available at: https://www.pmi.org/learning/thought-leadership/pulse/ai-innovators (Accessed June 2019)

PMI (2020) *Disruptive Technologies*, available at: https://www.pmi.org/learning/thought-leadership/series/disruptive-technologies.

PMI (2021) *5 Implications of Artificial Intelligence for Project Management*, available at: https://www.pmi.org/learning/publications/pm-network/digital-exclusives/implications-of-ai

Prasad, B. M. G. S., P.V.S. (2018) *Predicting Effort, Duration and Cost in Software Development Using Machine Learning Technique*, available at: https://www.jardcs.org/abstract.php?id=665

Prasad, K. S. N. & Saradhi, M. V. V. (2019) Comprehensive Project Management Framework using Machine Learning, *International Journal of Recent Technology and Engineering*, 8(2S3)

Rajkomar, A., Dean, J. & Kohane, I. (2019) Machine learning in medicine. *New England Journal of Medicine*, 380 (14): 1347-1358.

Rowland, M. (2020) Why next year could be a turning point for project management and AI, available at: https://www.apm.org.uk/blog/why-next-year-could-be-a-turning-point-for-project-management-and-ai/ (Accessed 31 Mar 2020)

Russell, C. L. (2005) An overview of the integrative research review. *Progress in transplantation*, 15 (1): 8-13.

Sanni-Anibire, M. O., Zin, R. M. & Olatunji, S. O. (2020) Machine learning model for delay risk assessment in tall building projects. *International Journal of Construction Management*, 1-10.

Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN computer science*, 2 (3): 160.

Schemelzer, R. (2019) *AI In Project Management*. Forbes, available at: https://www.forbes.com/sites/cognitiveworld/2019/07/30/ai-in-project-management/?sh=6dc9d73b4a00 (Accessed 30 July 2019)

Schmidt, B. Z. (2021). *Career Advice for Young Scientists in Biomedical Research: How to Think Like a Principal Investigator*. Springer Nature.

Sharmila, R. & Chellammal, S. (2018) A conceptual method to enhance the prediction of heart diseases using the data techniques. *International Journal of Computer Science and Engineering*, 6 (4): 21-25.

Shoar, S., Chileshe, N. & Edwards, J. D. (2022) Machine learning-aided engineering services' cost overruns prediction in high-rise residential building projects: Application of random forest regression. *Journal of Building Engineering*, 50 104102.

Siqueira, I. (1999) Automated cost estimating system using neural networks. *Project Management Journal*, 30 (1): 11-18.

Smith, W. S., Coleman, S., Bacardit, J. & Coxon, S. (2019) Insight from data analytics with an automotive aftermarket SME. *Quality and Reliability Engineering International*, 35 (5): 1396-1407.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of business research*, 104: 333-339.

Spalek, S. (2018) *Data analytics in project management*. CRC Press.

Swanson, S. A. (2017) *Real Possibilities: Artificial Intelligence Has The Potential to Allow Project Managers to Focus on Higher-Value Work. The Question Is: Are They Ready?,* available at: https://www.pmi.org/learning/library/ai-potential-allow-pm-higher-value-work-10540

Sydow, J. & Braun, T. (2018) Projects as temporary organizations: An agenda for further theorizing the interorganizational dimension. *International journal of project management*, 36 (1): 4-11.

Talaviya, T., Shah, D., Patel, N., Yagnik, H. & Shah, M. (2020) Implementation of artificial intelligence in agriculture for optimisation of irrigation and application of pesticides and herbicides. *Artificial Intelligence in Agriculture*, 4 58-73.

Tekic, A., Zerjav, V. & Tekic, Z. (2022) Evolution of project studies through the lens of engaged scholarship: A longitudinal bibliometric analysis. *International Journal of Project Management*, 40 (5): 531-546.

Thomas, J. J. & Cook, K. A. (2006) A visual analytics agenda. *IEEE computer graphics and applications*, 26 (1): 10-13.

Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human resource development review*, 4 (3): 356-367.

Tsai, C. W., Lai, C. F., Chao, H. C., & Vasilakos, A. V. (2015). Big data analytics: a survey. *Journal of Big data*, 2: 1-32.

Tufty, H. G. (2001). Value engineering: Mistaken identity?. *Consulting-Specifying Engineer*, 30 (6): 58-58.

Tukey, J. W. (1977). *Exploratory data analysis* (Vol. 2, pp. 131-160).

Tukey, J. W. (1962). The future of data analysis. *The annals of mathematical statistics*, 33 (1): 1-67.

Turner, J. R. & Müller, R. (2003) On the nature of the project as a temporary organization. *International journal of project management*, 21 (1): 1-8.

Twala, B. & Cartwright, M. (2010) Ensemble missing data techniques for software effort prediction. *Intelligent Data Analysis*, 14 (3): 299-331

Venkatasubramanian, K. (2020) *Unlocking data insights to improve project planning and delivery: Four key steps*, available at: https://blogs.oracle.com/construction-engineering/post/unlocking-data-insights-to-improve-project-planning-and-delivery-four-key-steps

Verbeke, W., Baesens, B., & Bravo, C. (2017). *Profit driven business analytics: A practitioner's guide to transforming big data into added value*. John Wiley & Sons.

Walker, D. & Lloyd-Walker, B. (2019) The future of the management of projects in the 2030s. *International journal of managing projects in business*, 12 (2): 242-266.

Wang, Y., Shao, Z. & Tiong, R. L. (2021) Data-driven prediction of contract failure of public-private partnership projects. *Journal of Construction Engineering and Management*, 147 (8): 04021089.

Wang, Y. & Tiong, R. L. (2022) Public–Private Partnership Contract Failure Prediction Using Example-Dependent Cost-Sensitive Models. *Journal of Management in Engineering*, 38 (1): 04021079.

Wang, Y.-R., Yu, C.-Y. & Chan, H.-H. (2012) Predicting construction cost and schedule success using artificial neural networks ensemble and support vector machines classification models. *International Journal of Project Management*, 30 (4): 470-478.

Wauters, M. & Vanhoucke, M. (2014) Support vector machine regression for project control forecasting. *Automation in Construction*, 47 92-106.

Wawak, S. & Woźniak, K. (2020) Evolution of project management studies in the XXI century. *International Journal of Managing Projects in Business*, 13 (4): 867-888.

Whyte, J. (2019) How digital information transforms project delivery models. *Project management journal*, 50 (2): 177-194.

Whyte, J., Stasis, A. & Lindkvist, C. (2016) Managing change in the delivery of complex projects: Configuration management, asset information and 'big data'. *International Journal of Project Management*, 34 (2): 339-351.

Wiegand, T., Krishnamurthy, R., Kuglitsch, M., Lee, N., Pujari, S., Salathé, M., Wenzel, M. & Xu, S. (2019) WHO and ITU establish benchmarking process for artificial intelligence in health. *The Lancet*, 394 (10192): 9-11.

Williamson, H. F., Brettschneider, J., Caccamo, M., Davey, R. P., Goble, C., Kersey, P. J., May, S., Morris, R. J., Ostler, R. & Pridmore, T. (2021) Data management challenges for artificial intelligence in plant and agricultural research. *F1000Research*, 10 (324): 324.

Wu, L., Ji, W., Feng, B., Hermann, U. & AbouRizk, S. (2021) Intelligent data-driven approach for enhancing preliminary resource planning in industrial construction. *Automation in Construction*, 130 103846.

Yahia, H., Hosny, H. & Razik, M. E. A. (2011) Time contingency assessment in construction projects in Egypt using artificial neural networks model. *International Journal of Computer Science Issues (IJCSI)*, 8 (4): 523.

Yaseen, Z. M., Ali, Z. H., Salih, S. Q. & Al-Ansari, N. (2020) Prediction of risk delay in construction projects using a hybrid artificial intelligence model. *Sustainability*, 12 (4): 1514.

Yip, H.-l., Fan, H. & Chiang, Y.-h. (2014) Predicting the maintenance cost of construction equipment: Comparison between general regression neural network and Box–Jenkins time series models. *Automation in Construction*, 38 30-38.

Yun, J., Ryu, K. R. & Ham, S. (2022) Spatial analysis leveraging machine learning and GIS of socio-geographic factors affecting cost overrun occurrence in roadway projects. *Automation in Construction*, 133 104007.

Zhang, Y., Cordina-Duverger, E., Komarzynski, S., Attari, A. M., Huang, Q., Aristizabal, G., ... & Levi, F. (2022). Digital circadian and sleep health in individual hospital shift workers: a cross sectional telemonitoring study. *EBioMedicine*, 81.