

THE DEVELOPMENT OF AN ACTION PRIORITY MATRIX AND TECHNOLOGY ROADMAP FOR THE IMPLEMENTATION OF DATA-DRIVEN AND MACHINE-LEARNING-BASED PREDICTIVE MAINTENANCE IN THE SOUTH AFRICAN RAILWAY INDUSTRY

L. Nethamba^{1*} & S. Grobbelaar¹

ARTICLE INFO

Article details

Presented at the 2nd International Conference on Industrial Engineering, Systems Engineering and Engineering Management, held from 2 to 4 October 2023 in Somerset West, South Africa

Available online 17 Nov 2023

Contact details

* Corresponding author
lindelani.Nethamba@gmail.com

Author affiliations

1 Department of Engineering and Technology Management,
University of Pretoria, Pretoria,
South Africa

ORCID® identifiers

L. Nethamba
<https://orcid.org/0000-0002-2310-6735>

S. Grobbelaar
<https://orcid.org/0000-0001-8148-2440>

DOI

<http://dx.doi.org/10.7166/34-3-2958>

ABSTRACT

In improving railways for the future, artificial intelligence and machine learning were identified as top-priority technology systems that enable data-driven methods and predictive maintenance. A local survey using semi-structured interviews showed that the railway industry lags behind in adopting and implementing data-driven and machine-learning methods for predictive maintenance. Insights from international studies were found to be relevant in South Africa. Other implementation barriers were identified in the socio-economic and socio-political areas of South Africa. An action priority matrix and technology roadmap was developed to guide the South African railway industry towards the implementation of data-driven and machine learning-based predictive maintenance. The action priority matrix was developed by using a two-round Delphi technique to rank the prioritisation of the required activities. The research showed the importance of considering insights from both international studies and the local context when adopting and implementing technology systems to improve business objectives.

OPSOMMING

In die verbetering van spoorweë vir die toekoms, is kunsmatige intelligensie en masjienleer geïdentifiseer as top-prioriteit tegnologiestelsels wat data-gedrewe metodes en voorspellende instandhouding moontlik maak. 'n Plaaslike opname wat semi-gestruktureerde onderhoude gebruik het, het getoon dat die spoorwegbedryf agterbly met die aanvaarding en implementering van datagedrewe- en masjienleermetodes vir voorspellende instandhouding. Daar is gevind dat insigte uit internasionale studies relevant is in Suid-Afrika. Ander implementeringshindernisse is in die sosio-ekonomiese en sosio-politieke gebiede van Suid-Afrika geïdentifiseer. 'n Aksieprioriteitmatriks en tegnologie-padkaart is ontwikkel om die Suid-Afrikaanse spoorwegbedryf te lei tot die implementering van datagedrewe- en masjienleer-gebaseerde voorspellende instandhouding. Die aksieprioriteitmatriks is ontwikkel deur 'n tweeronde Delphi-tegniek te gebruik om die prioritisering van die vereiste aktiwiteite te rangskik. Die navorsing het getoon hoe belangrik dit is om insigte uit beide internasionale studies en die plaaslike konteks in ag te neem wanneer tegnologiestelsels aangeneem en geïmplementeer word om besigheidsdoelwitte te verbeter.

1. INTRODUCTION: BACKGROUND AND PURPOSE OF RESEARCH

Data-driven decision-making and machine learning are increasingly becoming the main objectives for many organisations in the Fourth Industrial Revolution [1]. A survey that Navis did with global railway industry executives revealed that automation, integration of systems, artificial intelligence, and machine learning are the top priorities for improving railways in the future [2]. These technology systems are to ensure that

the high service demands of railways can be met through having improved asset and infrastructure availability, reliability, safety, and quality. It is also a requirement that maintenance managers make smart decisions to ensure the protection of asset conditions and network availability [3].

This research study was aimed at assisting the railway industry of South Africa to implement data-driven decision-making and machine learning-based predictive maintenance. The study began by performing a scoping literature review to evaluate the implementation progress in international studies in order to provide insight into data-driven and machine learning methods [4]. The review process was also aimed at identifying the benefits, barriers, and enablers of implementation. A qualitative case study through semi-structured interviews was then done to identify the local issues and the local status of the technology. The case study was combined with a two-round Delphi technique with a panel of twenty-five railway experts to rank the prioritisation of the activities required for implementation. The results of the Delphi technique were used to develop a specialised technology roadmap that could serve as a high-level integrated plan and framework for implementation. The benefits and limitations of the developed technology roadmap were then identified.

Given the research aims, the following research questions were defined:

1. What types of predictive maintenance, data-driven decision-making, and machine-learning technologies could the railway industry implement?
2. What are the benefits of implementation?
3. What are the barriers to implementation?
4. What are the enablers of implementation?
5. What is the expected implementation framework?
6. What benefits could be realised by using the developed technology roadmap for the organisation?
7. What are the realised shortcomings and limitations of using the developed technology roadmap for the organisation?

2. LITERATURE REVIEW

The literature review for this study began by describing the role of Industry 4.0 technologies and capabilities, the proposed framework required for successful implementation, and the process of predictive maintenance in the railway environment. A scoping literature review method was then presented, which was aimed at summarising the existing literature on international studies and answering the first four research questions.

The main themes of the research were defined first:

- Data-driven decision-making is making decisions based on data that involves collecting, analysing, and interpreting data to inform decisions [5].
- Machine learning is a domain that develops algorithms to produce an outcome that is based on previous data or experience [6]. The system can acquire and integrate knowledge that is fed through large-scale observation for improvement without programming [7].
- Predictive maintenance is a proactive approach that uses diagnostic and prognostic methods to ensure effective maintenance. It is widely recognised as one of the most efficient maintenance strategies, as it involves assessing the condition of an asset to plan maintenance activities for it carefully [8].

2.1. Industry 4.0 technologies, the implementation framework, and the process of predictive maintenance in the railway environment

2.1.1. *The role of Industry 4.0 technologies in predictive maintenance*

Industry 4.0 technologies are a wide variety of interdisciplinary technologies with different maturity levels that enable the advancement of digitisation, automation, and integration. The six design principles of Industry 4.0 identified by Hermann *et al.* - decentralisation, interoperability, virtualisation, service orientation, modularity, and real-time capability - are necessary for implementing machine learning in predictive maintenance. Improved decision-making is one of the promises and the potential that Industry 4.0 brings [9]. Machine learning is enabled by effectively capturing, storing, and managing large amounts

of data [3]. Sensor data and real-time condition monitoring techniques are better suited to collecting the state of the asset, and yield more accurate results than visual inspection and wayside condition monitoring systems. This technique, called ‘smart-monitoring’, enables accurate health prognosis through analysing signals from sensors mounted on the asset or infrastructure [8]. The increased number of analysed components and sensors is an enabler, as it improves the amount and quality of data available for the design of a better maintenance strategy [8].

2.1.2. The implementation framework for predictive maintenance

Mulders and Haarman [10] proposed a framework for predictive maintenance relevant to this research study. The framework began by identifying the different maturity levels of an organisation in how they monitor its assets. Level 1 uses visual inspections to gather data and make decisions, and conclusions are based completely on the expertise of the person performing the inspection. Level 2 uses both periodic inspections and instrument readouts. Level 3 uses real-time condition monitoring systems for the assets, and alarms are raised when assets operate beyond certain pre-established levels. Level 4 is where continuous real-time monitoring of the assets is done and predictive techniques are used to make maintenance decisions [10].

The six steps for digital success in predictive maintenance were outlined by Mulders and Haarman [10]. These fall within the project management and organisational alignment of a company, and are enabled through existing Internet of Things infrastructure management [10]:

1. The first step is having the organisation plot a predictive maintenance strategy by first evaluating its maturity level (1-4).
2. In the second step, the organisation should create pilot projects by selecting assets that are suitable for predictive maintenance. These assets will be used to demonstrate business value and set the proof of concepts in place.
3. Cross-functional teams should be created and provided with sufficient resources and the freedom to begin innovative ways of working. The evidence from early success will serve to win support and secure funding for a bigger rollout.
4. The third step requires the organisation to establish the capabilities required for successful implementation to be clearly defined from the lessons learned on the pilot assets. Then data analytics and reliability engineers should continually learn and enable the organisation to have highly skilled and competent people for data analytics.
5. A predictive maintenance strategy should be deployed across the assets of the organisation.
6. Support structures also need to be continually developed through data management.

That framework is important to consider when implementing the technology across the organisation; however, in answering research question five, a technology roadmap that presents specified tasks, important milestones, and technology capabilities was considered more suitable.

2.1.3. Predictive maintenance in the railway environment

The first level of the implementation framework is condition-based monitoring, which requires smart sensor technology. Smart sensors are necessary to monitor any deviation from ideal operating conditions. The second level is where the data acquired by the sensors is converted into useful information. At this stage, intelligent algorithms using machine-learning and data-mining techniques are used. This process leads to the cyber level, where information from different sources is combined in a central information hub to establish a cyberspace environment. At this level, the state of the operating assets is seen as a complete picture; additional tools can be used for deeper assessment. The next level is referred to as the cognitive level, where expert personnel can interface with the information to propose the required decision. Visualisation makes the cognitive level possible, and makes it an important part of the data-mining and data-preparation process to eliminate erroneous or missing measurements. The last level is the configuration stage, where the final instructions are accepted [11]. This framework is consistent with the predictive maintenance methodology by Yang *et al.* that was presented earlier, which breaks down the process into data-gathering from sensor data and maintenance data. The step is followed by data representation, modelling, and model fusion, after which predictive models are developed [12].

2.2. Scoping literature review

The scoping literature review was considered suitable for this kind of research, in which evidence of implemented strategies is emerging [13]. It was done according to the Arksey and O'Malley framework [14]. Figure 1 shows the flow chart of the scoping literature review process.

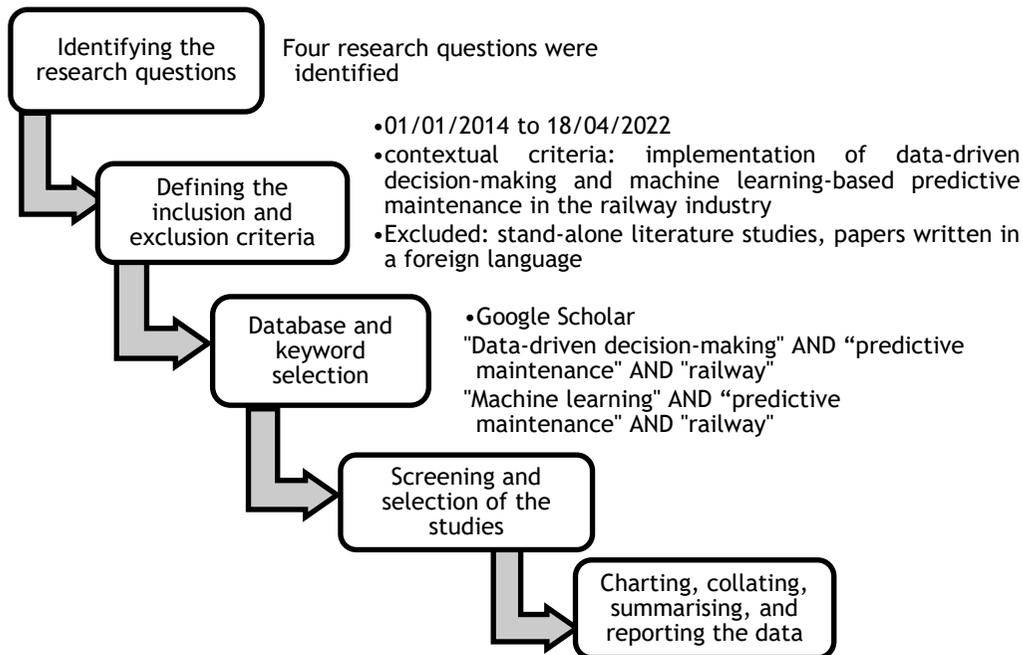


Figure 1: Flow chart of the scoping literature review

2.2.1. Research question one: What types of predictive-maintenance, data-driven decision-making, and machine-learning technology could the railway industry implement?

A review of the literature shows that there is an extensive study of machine-learning-based technologies and artificial intelligence across the world. Research has been done in the following countries: Italy, the United Kingdom, France, Germany, South Africa, Nigeria, Brazil, Russia, Greece, Turkey, Belgium, Sweden, Bangalore, India, Norway, Portugal, Japan, Canada, and China.

The main technologies and capabilities discussed in the study are real-time condition-monitoring and Internet-of-Things sensors, multi-sensors, accelerators, on-board and off-board data collecting systems, and instrumented railway vehicles, all of which enable the generation and collection of data; an integrated asset data management system such as the Systems Applications and Products in Data Processing (SAP) enterprise resource planning (ERP) tool; failure mode effects and criticality analysis (FMECA); and remote inspection [3][8]. Data mining was used to refine the existing datasets, and Matrix Laboratory (MATLAB2018b) was used to predict the remaining useful life and time-to-failure from existing historical failure data and maintenance data [15][16].

Cloud computing was selected as an effective tool for analysing operating data [11]. The common machine-learning techniques used were supervised machine learning, regression, classification, support vector machines, deep neural networks, and rule-based methods [17][18][12][19][20]. Unsupervised machine learning was used in cases where there was a lack of available labelled data [17]. Other important capabilities discussed in the studies are big data analytics, interoperability, decentralisation, modularity, and security.

Specific machine-learning models/algorithms used in the literature are listed below:

- SAP's ERP to develop machine-learning classifiers [3]
- Deep-learning algorithms alongside regression and classification training [18]
- Deep neural networks and convolution neural networks [21]
- Multiple classification methods [22]
- Support vector machines [22]
- Statistical and symbolic artificial intelligence [5]
- MATLAB2018b's diagnostic algorithm [16]
- Cepstral analysis [23]
- Genetic algorithm resource management and supervised machine learning [24]
- An auto-regressive moving average model [25]
- Combining classification, support vector machine regression, and clustering [12]
- Combining supervised and unsupervised machine learning techniques [17]

2.2.2. Research question two: What are the expected benefits of implementation?

The benefits of data-driven and machine-learning-based approaches include efficient decision-making in maintenance activities, planning and budgeting of maintenance activities, and better forecasting of maintenance needs [3][26][27][28]. Running safer trains with reduced adverse events and service interruptions also benefits the railway industry [15][29]. In addition, using these technologies provides the benefit of real-time asset health monitoring, enabling timely detection of deviations [16]. Furthermore, remote inspections can identify and resolve issues for assets operating in hazardous areas [30][31]. Maintenance managers can rely on systems and algorithms to provide precision [33][17]. Thus it may also significantly reduce the time required to address potential malfunctions, thereby increasing the lifespan of railway components and assets [16][32]. Furthermore, incorporating key performance indicators into the modelling enhances decision-making [15][34], and financial advantages, including avoiding costly corrective and unplanned maintenance and preventing wasteful inventory planning, may be achieved [15].

2.2.3. Research question three: What are the barriers to implementation?

The barriers to the implementation of data-driven decision-making and machine learning-based predictive maintenance can be divided into three main categories: technological, organisational, and ethical or socio-political. Technological barriers include the lack of data required to develop the machine learning models, the lack of the products required by the system, the challenge of selecting the correct model that would be useful for the required application, and the non-interpretability of the models and methods being used [33][34][19][35]. Organisational barriers include the risk of misaligning the systems with the business goal or the system not meeting the business-level goals, regulatory underdevelopment, and stakeholder resistance [8][36]. There will also be a requirement to upskill the workforce in data-science skills, and a high level of investment in the technologies is required [37][35]. A lack of standardisation and work organisation makes implementing the strategy difficult [35]. The organisational mindset and culture can also be resistant to change to data-driven decision-making methods. Top management in the organisation being ignorant can result in a lack of support for the strategy [36]. Uncertainty in the cost benefit of the system as it relates to the entire asset management of the assets is also an organisational barrier [8]. Ethical barriers relate to concerns over the replacement of jobs and over-reliance on technology [30].

2.2.4. Research question four: What are the enablers of implementation?

The enablers of data-driven decision-making and machine-learning-based predictive maintenance can be divided into three main categories: technological, financial, and organisational-level. The technological enablers are the availability of quality historical data and failure data of the assets, efficient data management processes, wheel acceleration data, and the ability to combine complex track data into a single parameter, as well as data fusion [3][29][28][27][26][16][34]. Computing power and the ability to integrate information across many sources and platforms are enablers. Interoperability, visualisation, decentralisation, real-time capability, information and communications technology infrastructure, modularity, skilled human capital, and cyber-security are all enablers of the strategy's implementation [38]. Financial enablers include having access to low-cost software and the ability of an organisation to

have an integrated cost-benefit analysis of the entire asset management system [39]. Organisational enablers are a collaboration of initiatives and partnership networks [9]. A governance structure that oversees the activities of the system with the ability to veto and effective communication channels are also enablers in the organisation [40]. A leadership that steers towards data-driven methods and a composite team of experts enable strategy implementation [41]. The organisation also needs to have the appropriate promotion methods for data flow, clearly defined predictive maintenance requirements, and a change in the organisational culture and mindset that supports the strategy [42].

3. THEORETICAL AND CONCEPTUAL FRAMEWORK

Data-driven methods require robust strategic planning, implementation, sustaining, organisational culture, and a mindset change [42]. An organisation’s challenges in becoming a data-driven organisation include data, management, decision-making, and an organisational mindset and culture [1]. For this reason, a technology roadmap is the most suitable tool to show how an organisation can take advantage of the new opportunities provided by technology while also allowing alignment within different teams [43]. An understanding of the local issues to be considered and addressed is also necessary; this has been done using the technology, organisation, and environment theoretical framework developed by Tornatzky, Fleischer and Chakrabarti [44]. The framework was used to predict an organisation’s intention to adopt information systems and to understand the variables that influence the adoption of new technology [45]. For this reason, it was considered a suitable framework for determining the factors in implementing data-driven decision-making and machine learning-based predictive maintenance, as it addresses the variables involved in this study. Figure 2 shows the theoretical framework for the research study.

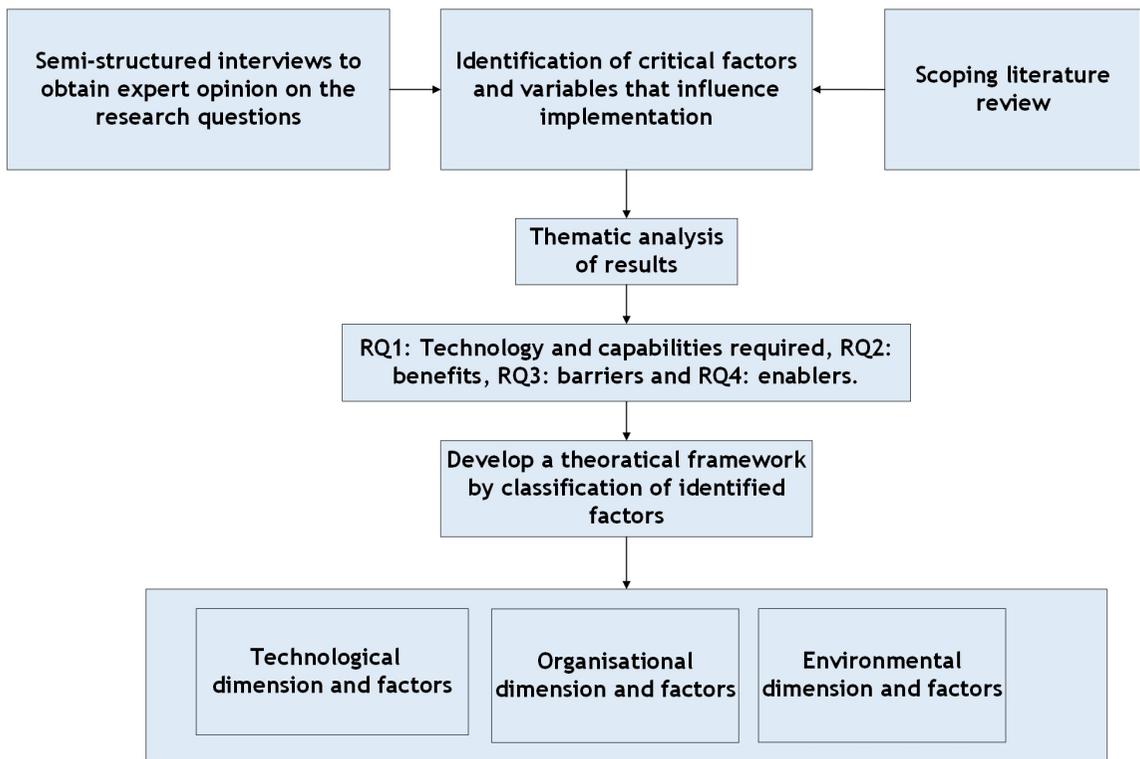


Figure 2: Theoretical framework

The technology roadmap was developed using a conceptual framework that illustrates how the elements were derived and how decisions were made. Figure 3 shows the conceptual framework for the research study.

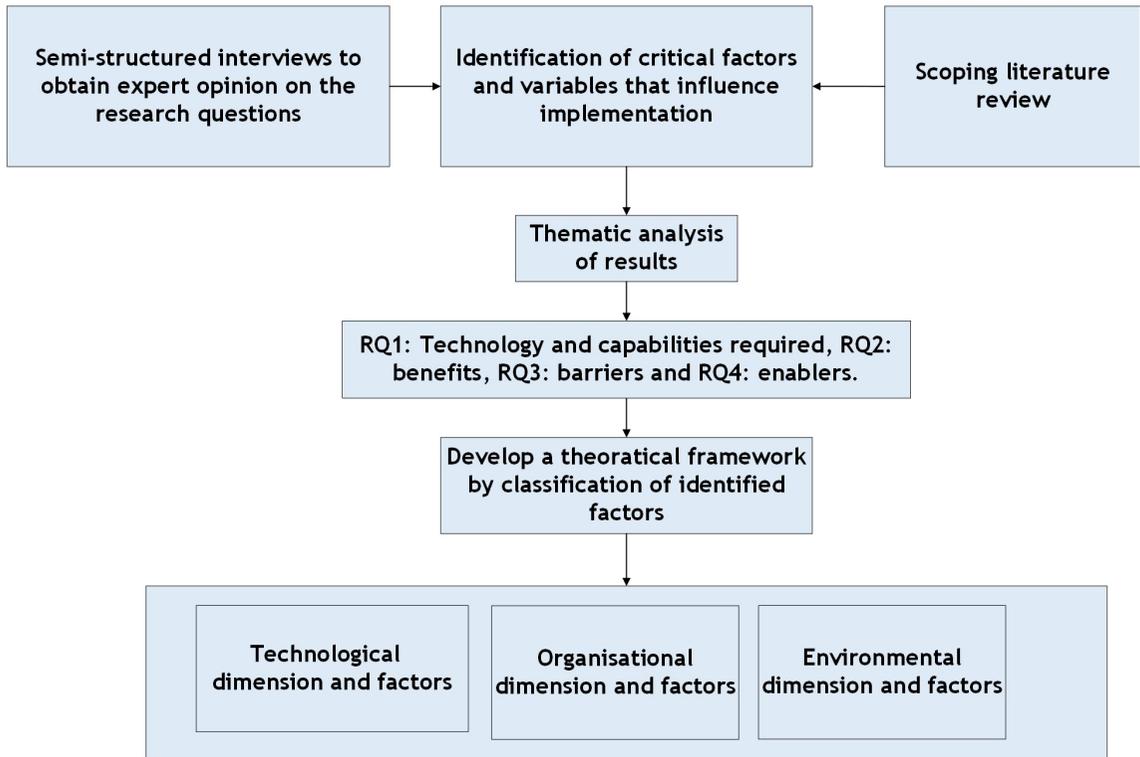


Figure 3: The conceptual framework for developing the action priority matrix and technology roadmap

4. METHODOLOGY

4.1. Research design

The study was conducted as a qualitative case study through the collection of primary data. The research was also conducted under two research paradigms. A positivist paradigm, which has an objectivist epistemological position and an ontology of realism, was used for research questions one, three, and four. This was because the positivist paradigm assumes that reality is independent of people [46]. Research questions two, six, and seven were conducted under a constructive research paradigm, which is based on the ontology that people construct different meanings informed by their experience and knowledge [47].

4.2. Population and sampling

The setting of the research was a railway organisation in South Africa. A judgemental sample [48] was used that included three methods of data collection:

- The semi-structured interviews were sent to fifteen railway experts; eight respondents participated in the interviews.
- The two-round Delphi technique was sent to twenty-five railway experts, and all twenty-five experts - from railway operations, information and communications technology, and maintenance departments - responded. Two elements that promoted variability in perspectives, skills, opinions, and independent sources in the Delphi technique were diversity in demographic characteristics and aspects related to professional experience and expertise [49].
- The final questionnaire was sent to the same twenty-five railway experts and to external experts in machine learning and railway engineering; only three respondents participated in this questionnaire.

5. RESEARCH RESULTS

5.1. Semi-structured interviews

The selected method of analysis was for the thematic analysis of the semi-structured interviews, using the technology, organisation, and environment (TOE) theoretical framework. Thematic analysis, facilitated by an initial coding process of the raw data, allowed the researcher to maintain coherence between the research objectives and results and to ensure validity [50].

The factors that affect implementation and that were identified in international studies through the literature review were compared with the local context of the South African railway industry. New variables were defined that were unique to the context of the study; however, it was found that most of the factors in the benefits, enablers, and barriers applied to the South African context as well.

1. What are the different types of predictive-maintenance, data-driven decision-making, and machine-learning technology that the railway industry could implement?

Key technology systems such as real-time capability, data analytics and machine-learning models, integration of systems, and Internet-of-Things technology and infrastructure were identified as technology that is still required in the organisation to implement data-driven decision-making and machine-learning predictive maintenance. According to Mulders and Haarman's framework, the heavy haul railway industry of South Africa sits on maturity levels 1 (exclusively visual inspections), 2 (combination of visual and instrument measurements), and 3 (continuous real-time monitoring, with alerts given on the basis of pre-established rules) [10].

2. What advantages can be anticipated from implementing various technologies?

Benefits that could be realised were identified in asset management, operations, decision-making, and maintenance activities, which were also identified in the literature. The benefits are not met without barriers to implementation.

3. What are the barriers to implementation?

Many barriers were identified in relation to the technology, compared with those identified in the literature. Other barriers in South Africa's railway industry are national, ethical, economic, socio-economic, and socio-political issues. Theft and vandalism were uniquely identified as existing in the context of the South African railway industry. The reflective analysis done by the researcher acknowledged that these issues were because of the refined context of the research study. The issue of theft had been addressed in the organisation by using machine learning to predict the likelihood of theft in the rail network. With a modification of this application, the organisation would stand to benefit from it. The barriers identified in the organisation included resistance from trade unions and a lack of government support.

4. What are the enablers for implementation?

The necessary enablers for implementation were enablers of technology, the organisation, and finances. The technology enablers were mostly those already discussed in the literature study, with an emphasis on the technology for model revision and reassignment to keep benefitting from the technology, and a need to have an integration of systems. Organisational enablers were like those identified in the literature, with an emphasis on the requirement for buy-in from employees at every level, decision-makers, and trade unions; business cases to translate the technical benefits to financial key performance indicators; and a need for a formal programme to carry out the implementation. The financial enablers identified in the literature were found also to apply to the South African railway context, with an emphasis on having a cost-benefit analysis of the technology in the organisation to support its implementation.

5.2. The action priority matrix

The action priority matrix was developed from the prioritised ranking of expert opinions using the Delphi technique. The first step in conducting the Delphi technique was to generate statements and a weighting criterion to allow the participants to rate the ease of implementation and the impact of the implementation of the data-driven decision-making and machine-learning-based predictive maintenance. Statements (see Figure 4) were derived from the literature in a similar study by Gbadamosi *et al.* [30]. A pre-determined weight criterion was used to generate a four-point comparative Likert rating scale, in which respondents were required to rate 'ease of implementation' and 'impact of implementation'. The number of panellists in the study, qualitative analytical skills, and statistical analytical skills provided enough confidence in the stability of the results and in the reliability and validity of the study [51].

5.2.1. Determining consensus from the Delphi technique

The analysis process of the Delphi technique was to identify a consensus on the priority of all the action items [52]. Descriptive statistics of the mean, mode, median, standard deviation, and coefficient of variation were calculated to determine the central tendency and consensus of each Delphi statement (see Figure 4). Three statistical variables were used to generate the consensus of the group: standard deviation, coefficient of variance, and mode frequency.

In the first round of the study, most of the respondents selected the following:

- The digitalisation of maintenance records, integration of alarms into maintenance schedules, and automation of real-time data analytics, all of which was rated as requiring only minimal effort level 1.
- Integration of geographic information system (GIS) data into a common data environment, and using quick response code/radio-frequency identification to detect assets automatically were moved to moderate effort level 2. Tasks that require digitalisation of data from hardware sensors, digital inspection, automatic fault detection, simulation of failure modes, visualisation, and installing smart sensors and integration of them into a common environment were rated as moderate level 2, as in the first round.
- Building information modelling, on-board scanners for continuous monitoring, developing machine-learning algorithms, using drones, having predictive monitoring, and integrating the train design information and embedded sensors into a common data environment were still rated as difficult but achievable at level 3, as in the first round.
- Last, tasks that were considered very difficult and that required maximum effort level 4 were combining sensors and virtual reality to allow real-time asset visualisation, using digital twins, pattern recognition, and having remote expert personnel who would use augmented reality, as in the first round. The results indicated that most of the experts did not change their opinions about the ease of implementation between the two rounds.

In the second round of the study, most of the respondents selected the following:

- Enabling of remote expert assistants or maintenance personnel using augmented reality; integrating wearable technologies with IoT sensors on assets; and augmented reality devices building information modelling were rated as moderate level 2.
- Attaching onboard scanners for continuous data acquisition; using drones for asset flagging; automatic fault detection; developing and using machine-learning algorithms; using QR code or RFID for asset identification; pattern recognition; digitalisation; and integration of train design information and embedded sensors into a common data environment were all rated as having considerable positive impact (level 3), as in the first round.
- Tasks that were rated to have the maximum positive impact (level 4) were implementing digital twins for real-time condition monitoring; combining IoT sensors and virtual reality; integrating alarms into maintenance schedules; digitisation of maintenance records; predictive condition monitoring; and automation of real-time analytics, as in the first round.

Figure 4 and Figure 5 show the change in the standard deviation and coefficient of variance between the two rounds.

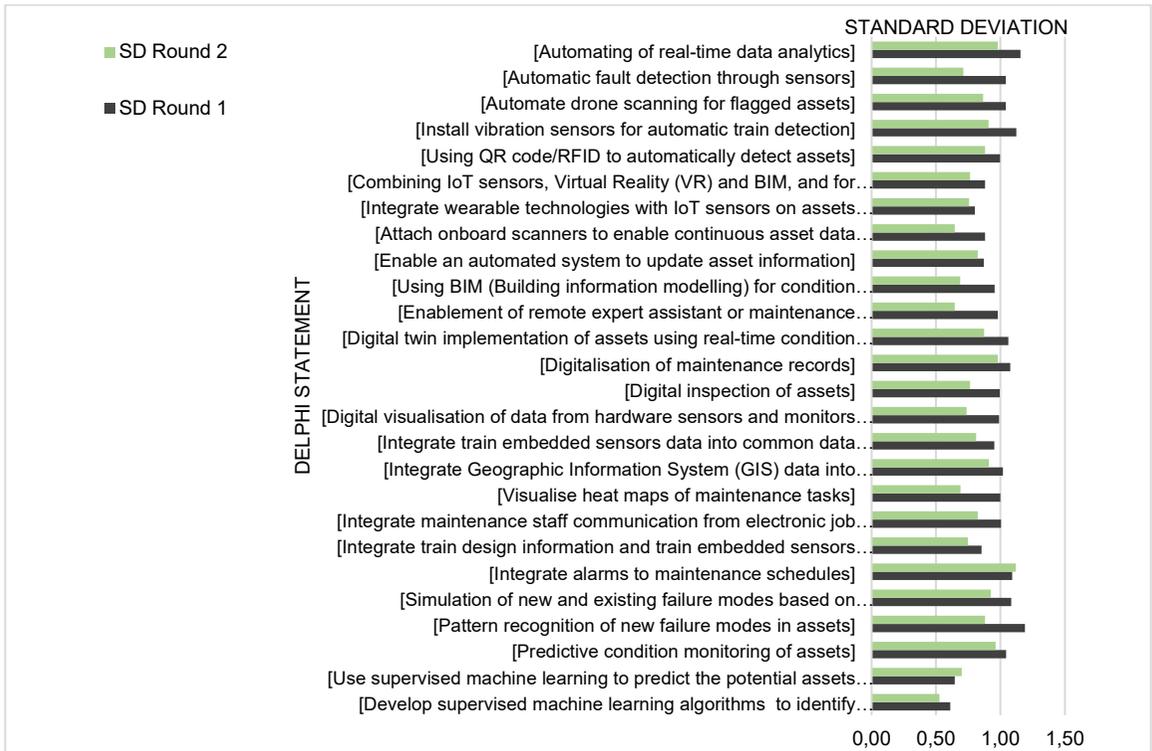


Figure 4: The differences in the standard deviation of Delphi data in rounds one and two as shown in the Delphi statement

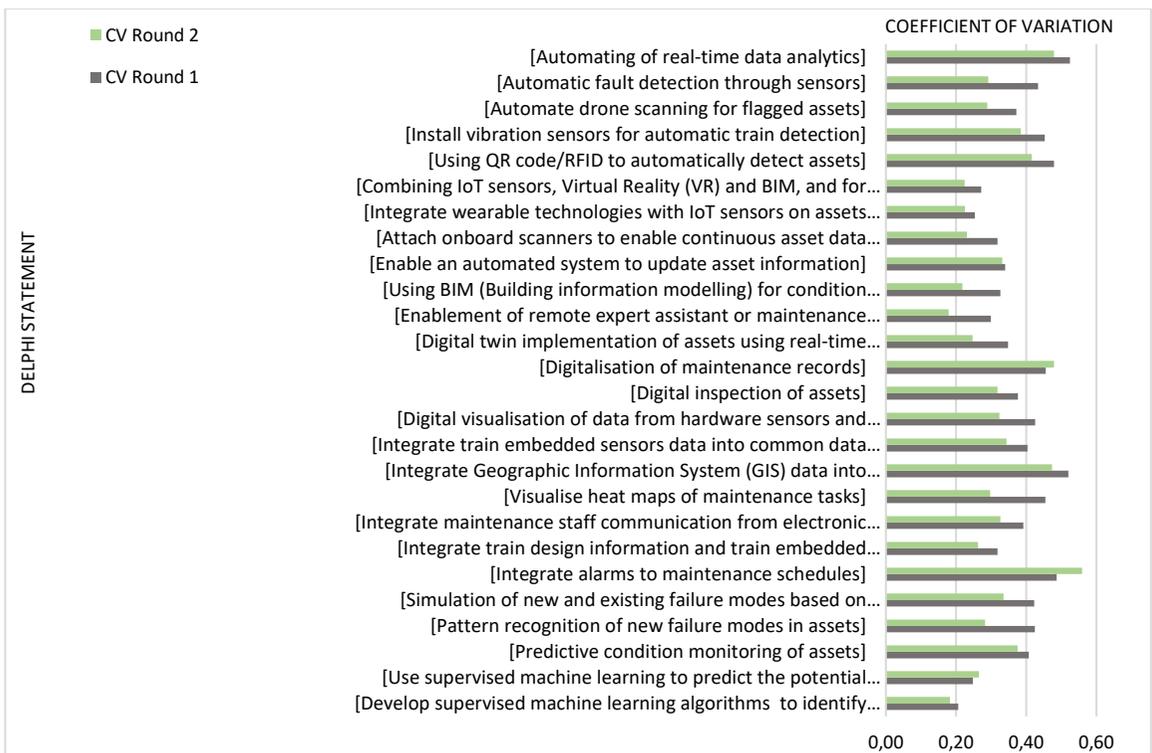


Figure 5: The differences in the coefficient of variation of Delphi data in rounds one and two as shown in the Delphi statement

The decrease in the standard deviation and coefficient of variance shows that the group judgement moved towards consensus in the second round of the Delphi technique. Group judgement showed the least consensus on the task of integrating alarms into the maintenance schedule in the second round of the study.

5.2.2. Transforming the action priority matrix into the technology roadmap

The transformation of activities in the action priority matrix into a technology roadmap was done by listing them on the basis of the quadrant in which they were placed. The matrix assisted in ranking the activities, based on their relative impact and ease of implementation. Figure 6 shows the action priority matrix ranking order. The arrow shows the direction in which the prioritisation ranking of the activities in the action priority matrix was transformed into the technology roadmap.

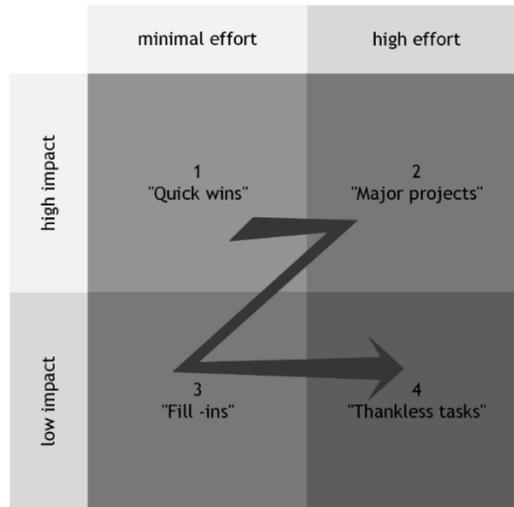


Figure 6: Action priority matrix ranking order

The steps below were followed for the development of the technology roadmap:

1. Step one: Assigning the high-priority activities.
 Quadrant 1 - Quick wins: These are high-impact activities that require minimal effort, and they were prioritised first. They have a greater positive impact on objective achievement, and their benefits can be realised within a short period. They will assist the organisation to move closer to its desired results.
2. Step two: Assigning medium-priority activities.
 Quadrant 2 - Major projects: These are high-impact and the most difficult activities, which should be prioritised second.
3. Step three: Fill-ins: Assigning low-priority activities.
 Quadrant 3 - Fill-ins: These are low-impact activities that require minimal effort, and are therefore prioritised third. The developed action priority matrix did not have activities in this quadrant, which led to moving to step 4.
4. Step four: Assigning very low-priority activities
 Quadrant 4 - Thankless tasks: These activities require a high degree of effort, but have a low impact on the objectives, and can therefore be scheduled last. Activities in this quadrant can be promoted, depending on the design requirements of the objectives' achievement.

5.2.3. The action priority matrix

Figure 7 shows the action priority matrix for implementing data-driven decision-making and machine learning-based predictive maintenance.

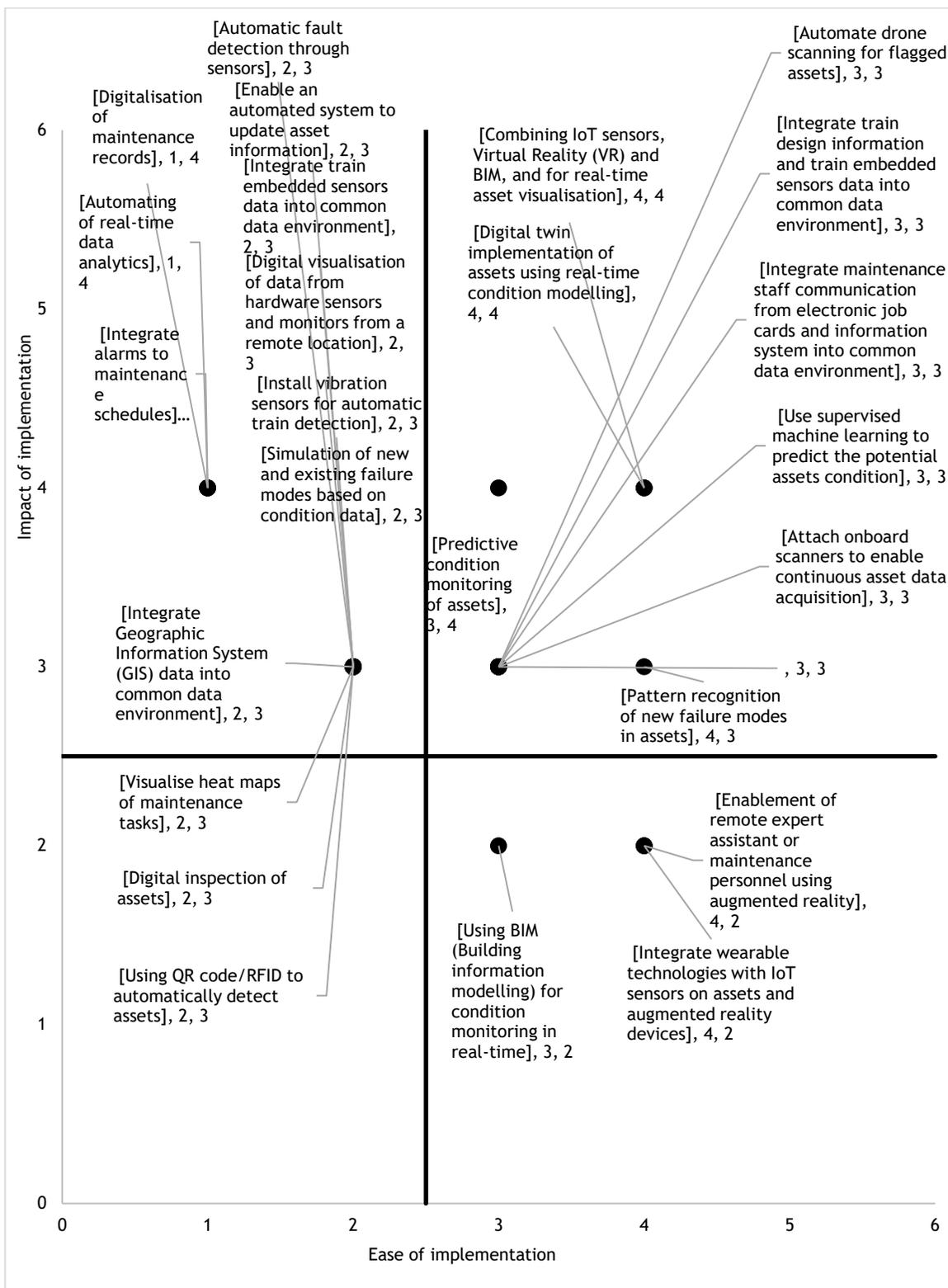


Figure 7: Action priority matrix for implementing data-driven decision-making and machine learning-based predictive maintenance

5.2.4. The technology roadmap

Figure 8 shows the technology roadmap for implementing data-driven decision-making and machine-learning-based predictive maintenance in the South African railway industry. A company perspective [53] was used to develop the technology roadmap through a process of setting a vision, identifying gaps between the current reality and the vision, and identifying actions to bridge these gaps. The features of the roadmap were described according to the study by Mang Hang and Phaal [54].

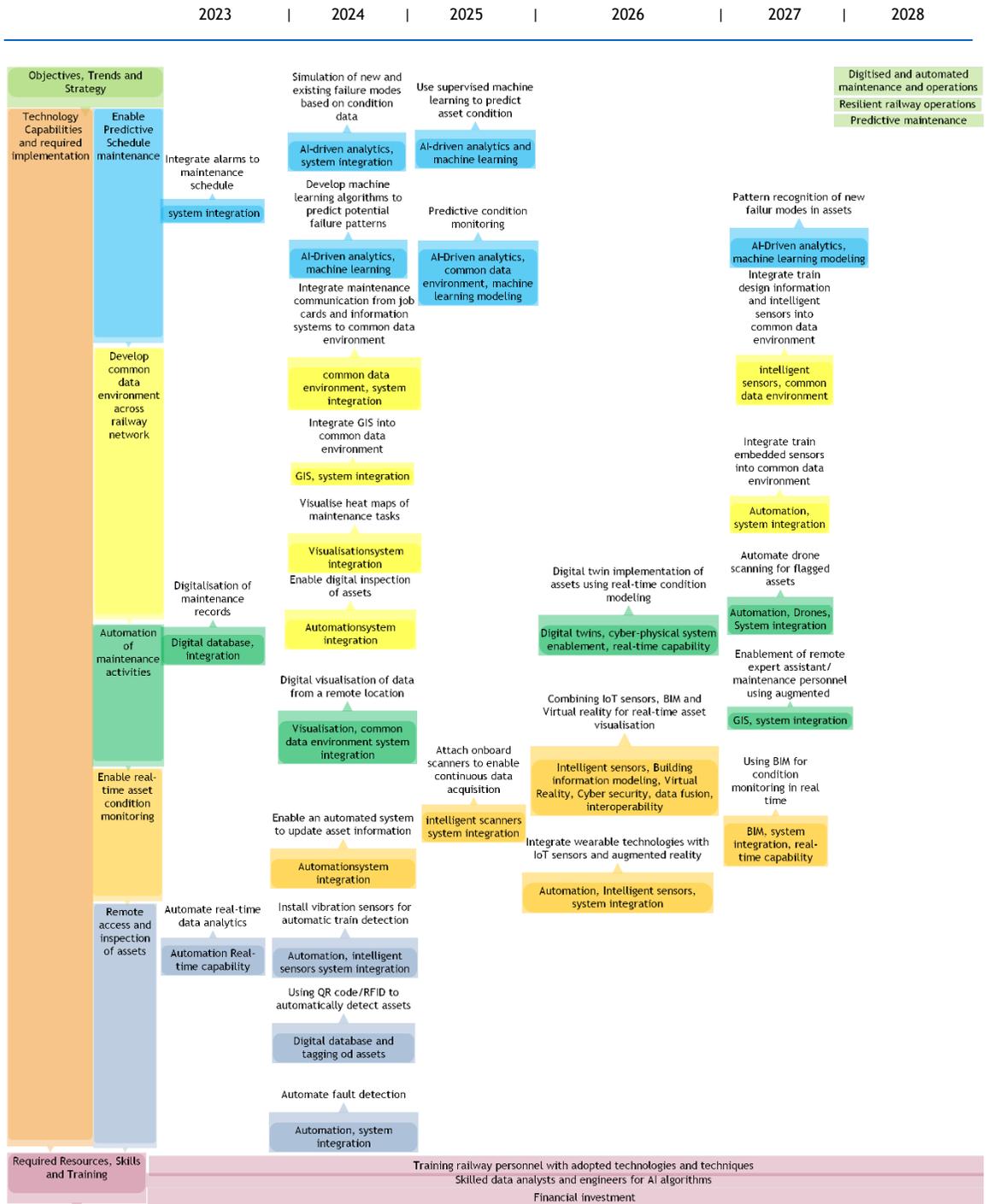


Figure 8: The technology roadmap for the implementation of data-driven and machine-learning-based predictive maintenance in the South African railway industry

The Delphi technique revealed a consensus of opinions about the prioritisation of the activities required for implementation. Automation of maintenance activities, enabling predictive scheduled maintenance, and remote access and inspection of assets were identified as the top priority activities for implementation. The prioritisation of these activities was found to be consistent with studies by Gbadamosi *et al.* [30] of the implementation strategy of Internet-of-Things-enabled predictive assets monitoring and maintenance for the United Kingdom's railway.

5.2.5. Research question six: What benefits can be realised by using the developed technology roadmap for the organisation?

The identified benefit of using the technology roadmap is that it promotes alignment within different teams in the organisation, as everyone is aware of the main objectives in developing data-driven decision-making and machine-learning-based predictive maintenance. This identified benefit matches the benefits in studies by Satoglu and Erdogan, who wrote that technology roadmaps assist in visualising outputs: goals, processes, and signs of progress in guiding and leading project teams [55][43]. The technology roadmap addresses the shortcomings of the current use and application of monitoring data. The last benefit addressed the importance of organisations having a plan, which was also identified by Satoglu and Erdogan, who wrote that technology roadmaps provide a high-level integrated plan and a common framework [43][55].

5.2.6. Research question seven: What are the realised shortcomings and limitations of using the developed technology roadmap for the organisation?

The limitations of the developed technology roadmap lie in specifying the timelines of the tasks in the roadmap. Geisler discussed this limitation, and argued that the design could reflect an underestimation of the time and effort required for adoption and implementation [56]. The technology roadmap is also limited in forecasting the ideal sequence and length of events. The amount of data required to create meaningful models is significantly underestimated in the technology roadmap, because integrating all of the sensors that gather the data is a lengthy process. The roadmap is also limited in specifying data-related tools that could emerge in this technology area [55][43].

6. DISCUSSION

The research suggests that the benefits of using data-driven decision-making and machine learning for predictive maintenance are not without barriers. The identified barriers were related to the required technology, the organisation, and ethics. For example, an investigation of the existing South African railway industry technology revealed that real-time capability, data analytics, machine-learning models, integration of systems, Internet-of-Things technology, and infrastructure would be required to steer the organisation towards data-driven decision-making and machine-learning-based predictive maintenance.

The local survey confirmed the findings of international studies on the benefits, barriers, and enablers of implementation. In addition to technological and organisational barriers, others that were identified locally were national, economic, socio-economic, and socio-political barriers that would strongly affect the implementation. The issue of theft and vandalism, which was uniquely identified and emphasised, also places the industry at the disadvantage of not being able to implement the proposed technology, decision-making strategy, and maintenance strategy. Organisational-level enablers that were identified included buy-in from employees at every level, decision-makers, and trade unions; business cases to translate the technical benefits into financial key performance indicators; cost-benefit analysis; and a need for a formal programme to carry out the implementation.

7. CONCLUSIONS

Railway systems around the world are being improved by implementing Internet-of-Things sensor technology, real-time condition monitoring, remote connectivity, remote inspection, data mining of collected asset data, diagnostic and prognostic tools, cloud-based storage, and machine vision. The South African railway industry lags behind in adopting and implementing the technology systems that would be required to promote data-driven and machine learning for predictive maintenance practices. Wayside condition monitoring components do not have a continuous real-time monitoring capability, and therefore do not provide the performance that would be necessary to implement data-driven decision-making and predictive maintenance [57].

The research results contribute to the sparse research on the adoption of Fourth Industrial Revolution-driven technology and practices in the South African railway context. The technology roadmap could be used as a framework to guide the implementation planning for using machine learning in the predictive maintenance of railway assets and infrastructure. The benefits of data-driven decision-making and machine-learning-based predictive maintenance are faced with many specified barriers to implementation in the technological, organisational, and environmental contexts, and reveal a major gap in technology adoption and readiness.

This research has contributed to the importance of an organisation considering its existing technology and identifying gaps and opportunities in the required capabilities to steer it towards a specific objective. It highlighted the highest-ranked capabilities that were required, which ought to be prioritised in planning to implement data-driven methods, machine learning, and predictive maintenance. The research results also indicate the need for managers to consider the lessons learned from international studies and from the local context when looking into adopting technology, and ensuring that the business objectives align with the required system capabilities. Managers need to address the local barriers to the technological, organisational, and external environment when it comes to implementing any technology that will provide value for the organisation.

8. RECOMMENDATIONS AND FUTURE STUDY

This research study was conducted on a South African railway organisation. Future research could investigate the implementation of the studied technology, decision-making strategy, and maintenance strategy in different industries. In addition, future research could address the specific areas of concern and enablers that have been highlighted as determinants of adopting and implementing the technology, decision-making strategy, and maintenance strategy discussed in this study. These could address the factors identified in the external environment of the organisation and the identified organisational-level enablers.

REFERENCES

- [1] R. Berntsson Svensson and M. Taghavianfar, "Toward becoming a data-driven organization: challenges and benefits," in *Research challenges in information science*, 2020, pp. 3-19. doi: 10.1007/978-3-030-50316-1_1
- [2] M. van Leijen, "Majority of rail freight industry plans to invest in automation - RailFreight," 2021. Available: <https://www.railfreight.com/technology/2021/09/01/majority-of-rail-freight-industry-plans-to-invest-in-automation/?gdnr=accept>. [Accessed Mar. 13, 2022].
- [3] Z. A. Bukhsh, A. Saaed, and I. Stipanovic, "A machine learning approach for maintenance prediction of railway assets," *Transport Research Arena TRA*, no. August 2019, 2018. doi: 10.5281/zenodo.3381949
- [4] L. Nethamba and S. Grobbelaar, "Considerations for the development of a data-driven and machine learning-based predictive maintenance implementation plan for the South African railway industry," *Proceedings of the International Annual Conference of the American Society for Engineering Management*, pp. 1-10, 2022 [Online]. Available: <https://www.proquest.com/conference-papers-proceedings/considerations-development-data-driven-machine/docview/2771104534/session?accountid=14717>
- [5] L. Cao, C. Zhang, P. S. Yu, and Y. Zhao, "D³ M Methodology," in *Domain Driven Data Mining*, pp. 27-47, 2010. doi: 10.1007/978-1-4419-5737-5_2
- [6] A. K. Raz, J. Llinas, R. Mittu, and W. Lawless, "Engineering for emergence in information fusion systems: A review of some challenges," *FUSION 2019 - 22nd Int. Conf. Inf. Fusion*, pp. 1-8, 2019. doi: 10.1016/b978-0-12-820543-3.00012-2
- [7] M. Mohammed, M. B. Khan, and E. B. M. Bashie, *Machine learning: Algorithms and applications*, Boca Raton: CRC Press, 2016.
- [8] S. Arena, E. Florian, I. Zennaro, P. F. Orrù, and F. Sgarbossa, "A novel decision support system for managing predictive maintenance strategies based on machine learning approaches," *Safety Science*, vol. 146, 105529, 2022. doi: 10.1016/j.ssci.2021.105529
- [9] D. Adebajo, T. Laosirihongthong, P. Samaranayake, and P. L. Teh, "Key enablers of Industry 4.0 development at firm level: Findings from an emerging economy," *IEEE Transactions on Engineering Management*, vol. 70, no. 2, pp. 400-416, 2023. doi: 10.1109/TEM.2020.3046764
- [10] M. Mulders and M. Haarman, "Predictive maintenance 4.0: Predict the unpredictable," *PwC*, no. June, 2017. Available: <https://www.pwc.be/en/documents/20171016-predictive-maintenance-4-0.pdf>. [Accessed Mar. 13, 2022].

- [11] Y. Kovaleva, O. Iliashenko, V. Iliashenko, and R. Anselm, "Implementation of predictive maintenance in a railway logistics company," *Proceedings of the International Scientific Conference - Digital Transformation on Manufacturing, Infrastructure and Service*, 2020. doi: 10.1145/3446434.3446507
- [12] C. Yang, S. Letourneau, J. Liu, Q. Cheng, and Y. Yang, "Machine learning-based methods for TTF estimation with application to APU prognostics," *Applied Intelligence*, vol. 46, no. 1, pp. 227-239, 2017. doi: 10.1007/s10489-016-0829-4
- [13] Z. Munn, M. D. J. Peters, C. Stern, C. Tufanaru, A. McArthur, and E. Aromataris, "Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach," *BMC Medical Research Methodology*, vol. 18, no. 1, 143, 2018, doi: 10.1186/s12874-018-0611-x
- [14] M. D. J. Peters, C. M. Godfrey, P. McInerney, C. B. Soares, H. Khalil, and D. Parker, *The Joanna Briggs Institute reviewers' manual 2015: Methodology for JBI scoping reviews*. Adelaide: Joanna Briggs Institute, University of South Australia, 2015.
- [15] I. Kalathas and M. Papoutsidakis, "Predictive maintenance using machine learning and data mining: A pioneer method implemented to Greek railways," *Designs*, vol. 5, no. 1, pp. 1-18, 2021. doi: 10.3390/designs5010005
- [16] I. A. Daniyan, K. Mpofo, and A. O. Adeodu, "Development of a diagnostic and prognostic tool for predictive maintenance in the railcar industry," in *Procedia CIRP*, 2020, vol. 90, pp. 109-114. doi: 10.1016/j.procir.2020.02.001
- [17] R. Prytz, *Machine learning methods for vehicle predictive maintenance using off-board and on-board data*, Volume 9 of Halmstad University Dissertations. Halmstad, Sweden: Halmstad University, 2014.
- [18] K. Han, M. Sun, X. Zhou, G. Zhang, H. Dang, and Z. Liu, "A new method in wheel hub surface defect detection: Object detection algorithm based on deep learning," in *International Conference on Advanced Mechatronic Systems, ICAMechS*, 2018, vol. 2017-Decem, pp. 335-338. doi: 10.1109/ICAMechS.2017.8316494
- [19] P. Strauß, M. Schmitz, R. Wöstmann, and J. Deuse, "Enabling of predictive maintenance in the Brownfield through low-cost sensors, an IIoT-architecture and machine learning," in *Proceedings – 2018 IEEE International Conference on Big Data, Big Data 2018*, 2019, pp. 1474-1483. doi: 10.1109/BigData.2018.8622076
- [20] Z. Allah Bukhsh, A. Saeed, I. Stipanovic, and A. G. Doree, "Predictive maintenance using tree-based classification techniques: A case of railway switches," *Transportation Research Part C: Emerging Technologies*, vol. 101, pp. 35-54, 2019. doi: 10.1016/j.trc.2019.02.001
- [21] C. Yang, Y. Sun, C. Ladubec, and Y. Liu, "Article developing machine learning-based models for railway inspection," *Applied Science*, vol. 11, no. 1, 13, 2021 doi: 10.3390/app11010013
- [22] G. A. Susto, A. Schirru, S. Pampuri, S. McLoone, and A. Beghi, "Machine learning for predictive maintenance: A multiple classifiers approach," *IEEE Transactions on Industrial Informatics*, vol. 11, pp. 812-820, 2015.
- [23] B. Baasch, J. Heusel, M. Roth, and T. Neumann, "Train wheel condition monitoring via cepstral analysis of axle box accelerations," *Applied Sciences*, vol. 11, no. 4, 1432, 2021. doi: 10.3390/app11041432
- [24] P. Killeen, B. Ding, I. Kiringa, and T. Yeap, "IoT-based predictive maintenance for fleet management," in *Procedia Computer Science*, 2019, vol. 151, no. 2018, pp. 607-613. doi: 10.1016/j.procs.2019.04.184
- [25] M. Baptista, S. Sankararaman, I. P. de Medeiros, C. Nascimento, H. Prendinger, and E. M. P. Henriques, "Forecasting fault events for predictive maintenance using data-driven techniques and ARMA modeling," *Computers & Industrial Engineering*, vol. 115, pp. 41-53, 2018. doi: 10.1016/j.cie.2017.10.033
- [26] A. Lasisi and N. Attoh-Okine, "Principal components analysis and track quality index: A machine learning approach," *Transportation Research Part C: Emerging Technologies*, vol. 91, pp. 230-248, 2018. doi: 10.1016/j.trc.2018.04.001
- [27] M. Calabrese *et al.*, "SOPHIA: An event-based IoT and machine learning architecture for predictive maintenance in Industry 4.0," *Information.*, vol. 11, no. 4, pp. 1-17, 2020. doi: 10.3390/INFO11040202
- [28] S. P. Gayialis, E. P. Kechagias, G. D. Konstantakopoulos, and G. A. Papadopoulos, "A predictive maintenance system for reverse supply chain operations," *Logistics*, vol. 6, no. 1, 4, 2022. doi: 10.3390/logistics6010004
- [29] H. Hadj-Mabrouk, "Analysis and prediction of railway accident risks using machine learning," *AIMS Electronics and Electrical Engineering*, vol. 4, no. 1. pp. 19-46, 2019. doi: 10.3934/ElectrEng.2020.1.19

- [30] A. Q. Gbadamosi *et al.*, “IoT for predictive assets monitoring and maintenance: An implementation strategy for the UK rail industry,” *Automation in Construction*, vol. 122, 103486, 2021. doi: 10.1016/j.autcon.2020.103486
- [31] Y. K. Teoh, S. S. Gill, and A. K. Parlikad, “IoT and fog computing based predictive maintenance model for effective asset management in Industry 4.0 using machine learning,” *IEEE Internet of Things Journal*, vol. 10, no. 3, pp. 2087-2094, 2021. doi: 10.1109/JIOT.2021.3050441
- [32] G. Kruppenacher, C. S. Ong, S. Koller, S. Kobayashi, and J. M. Buhmann, “Wheel defect detection with machine learning,” *IEEE Transactions on Intelligent Transport Systems*, vol. 19, no. 4, pp. 1176-1187, 2018. doi: 10.1109/TITS.2017.2720721
- [33] Z. M. Çınar, A. A. Nuhu, Q. Zeeshan, O. Korhan, M. Asmael, and B. Safaei, “Machine learning in predictive maintenance towards sustainable smart manufacturing in Industry 4.0,” *Sustainability*, vol. 12, no. 19, 8211, 2020. doi: 10.3390/su12198211
- [34] A. Theissler, J. Pérez-Velázquez, M. Kettelgerdes, and G. Elger, “Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry,” *Reliability Engineering and System Safety*, vol. 215, 107864, 2021. doi: 10.1016/j.res.2021.107864
- [35] C. Chauhan, A. Singh, and S. Luthra, “Barriers to Industry 4.0 adoption and its performance implications: An empirical investigation of emerging economy,” *Journal of Cleaner Production*, vol. 285, 124809, 2021. doi: 10.1016/j.jclepro.2020.124809
- [36] C. Reis, P. Ruivo, T. Oliveira, and P. Faroleiro, “Assessing the drivers of machine learning business value,” *Journal of Business Research*, vol. 117, pp. 232-243, 2020. doi: 10.1016/j.jbusres.2020.05.053
- [37] R. Welte, M. Estler, and D. Lucke, “A method for implementation of machine learning solutions for predictive maintenance in small and medium sized enterprises,” *Procedia CIRP*, vol. 93, pp. 909-914, 2020. doi: 10.1016/j.procir.2020.04.052
- [38] K. Wang, “Intelligent predictive maintenance (IPdM) system - Industry 4.0 scenario,” *WIT Transactions on Engineering Sciences*, vol. 113, pp. 260-267, 2016. doi: 10.2495/IWAMA150301
- [39] M. Compare, P. Baraldi, and E. Zio, “Challenges to IoT-enabled predictive maintenance for Industry 4.0,” *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 4585-4597, 2020. doi: 10.1109/JIOT.2019.2957029
- [40] J. Watson *et al.*, “Overcoming barriers to the adoption and implementation of predictive modeling and machine learning in clinical care: What can we learn from US academic medical centers?” *JAMIA Open*, vol. 3, no. 2, pp. 167-172, 2020. doi: 10.1093/jamiaopen/ooz046
- [41] B. Rakova, J. Yang, H. Cramer, and R. Chowdhury, “Where responsible AI meets reality: Practitioner perspectives on enablers for shifting organizational practices,” *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW1, pp. 1-23, 2021. doi: 10.1145/3449081
- [42] C. Zhang, X. Xie, and X. Guo, “Scheme design of railway predictive maintenance based on IOT and AI technology,” in *Proceedings of the 7th Annual International Conference on Social Science and Contemporary Humanity Development (SSCHD 2021)*, 2021, vol. 610, no. Sschr, pp. 337-342. doi: 10.2991/assehr.k.211215.062
- [43] Z. Erdogan, “Technology roadmap for Kuveyt Turk mobile banking,” T.C. Dogus University, 2014. Available: <https://openaccess.dogus.edu.tr/xmlui/bitstream/handle/11376/1045/376048.pdf?sequence=1&isAllowed=y>. [Accessed Mar. 20, 2022].
- [44] L. G. Tornatzky, M. Fleischer, and A. K. Chakrabarti, *Processes of technological innovation*. Lexington KY: Lexington Books, 1990.
- [45] J. N. Jere and N. Ngidi, “A technology, organisation and environment framework analysis of information and communication technology adoption by small and medium enterprises in Pietermaritzburg,” *SA Journal of Information Management*, vol. 22, no. 1, a1166, 2020. doi: 10.4102/sajim.v22i1.1166
- [46] A. A. Rehman and K. Alharthi, “An introduction to research paradigms,” *International Journal of Educational Investigations*, vol. 3, no. 8, pp. 51-59, 2016.
- [47] J. Cleland and S. Durning, *Researching medical education*. Hoboken NJ: John Wiley & Sons, 2015.
- [48] C. Page and D. Meyer, *Applied research design for business and management*. New York NY: McGraw-Hill Higher Education, 1999.
- [49] M. Barrios, G. Guilera, L. Nuño, and J. Gómez-Benito, “Consensus in the Delphi method: What makes a decision change?” *Technological Forecasting and Social Change*, vol. 163, 120484, 2021. doi: 10.1016/j.techfore.2020.120484
- [50] M. Skjott Linneberg and S. Korsgaard, “Coding qualitative data: A synthesis guiding the novice,” *Qualitative Research Journal*, vol. 19, no. 3, pp. 259-270, 2019. doi: 10.1108/QRJ-12-2018-0012
- [51] R. Brown, “Interviews, focus groups and Delphi techniques,” in *Advanced Research Methods for Applied Psychology*, pp. 95-10, 2018. doi: 10.4324/9781315517971

- [52] S. Drumm, C. Bradley, and F. Moriarty, “‘More of an art than a science’? The development, design and mechanics of the Delphi technique,” *Research in Social and Administrative Pharmacy*, vol. 18, no. 1, pp. 2230-2236, 2022. doi: 10.1016/j.sapharm.2021.06.027
- [53] R. Phaal, C. J. P. Farrukh, and D. R. Probert, “Technology roadmapping – A planning framework for evolution and revolution,” *Technological Forecasting and Social Change*, vol. 71, no. 1-2, pp. 5-26, 2004. doi: 10.1016/S0040-1625(03)00072-6
- [54] M. H. Yip and R. Phaal, “Roadmap feature analysis: Viewing ‘roadmaps’ as maps,” *2019 Portland International Conference on Management of Engineering and Technology (PICMET)*, Portland OR, 2019.
- [55] A. Ustundag and E. Cevikcan, *Industry 4.0: Managing the digital transformation*, Springer series in advanced manufacturing, Springer International, 2018.
- [56] E. Geisler, “What do we know about: R & D metrics in technology-driven organizations,” 2002. Available: https://www.researchgate.net/publication/228871609_What_do_we_know_about_RD_metrics_in_technology-driven_organizations. [Accessed Mar. 20, 2022].
- [57] L. Tobias, “Rail in South Africa: Breakthrough technologies,” *Transnet*, October 2019. Available: <https://www.transportevolution.com/wp-content/uploads/2019/10/Lloyd-Tobias>. [Accessed Oct. 10, 2021].