

Best Practices for Big Data Management in the Sports Industry

MART POSTMA, University of Twente, The Netherlands

Abstract – This research paper examines best practices for Big Data Management in the sports industry. With the increasing usage of Big Data Analytics worldwide, sports organisations are now able to collect and analyse enormous amounts of data that can be used to gain valuable insights into topics interesting to sports organisations, such as business activities and athlete performance. However, effective management of this data requires specialised knowledge, skills and infrastructure for Information Technology. Through a semi-structured review of related literature, this paper identifies best practices for Big Data Management in the sports industry. To be specific, these best practices are related to Sports Information Management, Sports Analytics and data-driven decision-making. The paper also aims to identify the challenges that sports organisations have to deal with when implementing these best practices. Furthermore, recommendations are provided for addressing these challenges based on a developed framework to manage Big Data in sports organisations. The findings of this paper intend to serve as an explanatory framework for sports organisations to implement Big Data Management in the near future.

Additional Key Words and Phrases: Big Data, Sports Information Management, Sports Analytics, Sports Science, Big Data Management

1 INTRODUCTION

In recent years, the sports industry has been transformed by the use of Big Data analytics. Wearable devices, such as smartwatches and activity trackers, have rapidly improved and developed through the years [21]. This rapid development allows the collection of vast amounts of data on a frequent basis. These devices collect a wide variety of data types and formats, offering the opportunity to gain valuable insights on, for example, athlete performance through data analysis. However, effective management of this data tends to be a complex and challenging task for sports organisations [24]. To extract knowledge and insights from data, lots of specialised knowledge, skills and Information Technology (IT) infrastructure is required across an interdisciplinary field. This is achieved in sports organisations by employing sports scientists with distinct backgrounds in sports science, such as performance analysts, biostatistics experts and data scientists [27, 43].

Sports science – the application of scientific principles to enrich sports with information and knowledge – has become increasingly common as professional sports organisations seek to gain a performance advantage over their competitors [43]. To achieve this performance advantage, sports organisations have started to develop and adapt to Big Data analytics strategies. Unfortunately, that has become an increasingly challenging issue [45]. As Big Data has evolved in sports, literature has become critical to our understanding of the role of Sports Analytics and its impact on various aspects of the industry [37, 44].

TScIT 39, July 7, 2023, Enschede, The Netherlands

© 2023 University of Twente, Faculty of Electrical Engineering, Mathematics and Computer Science.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Ward et al. argue that next to developing a framework for managing data, a framework to support the decision-making of sports scientists should be developed to "improve both the efficiency and effectiveness of decisions" [43]. As an outcome, they have designed a framework to support this decision-making process. However, the framework of Ward et al. is only designed from a business intelligence perspective. In contrast to the research of Ward et al, we aim to, next to formulating best practices, expand this framework to a multi-perspective view which spans multiple organisations in the sports industry. Sands et al. discuss the development of a framework for the main components of sports "involving the following: analyses of individuals, trend analyses, rules-based analysis, and statistical process control" [34]. This can be related to the framework we propose in this paper, demonstrating the opportunity to develop one framework matching the aspects of Big Data Management in the sports industry.

The identified problems and corresponding research questions are discussed in Section 2. In Section 3, the methodologies for answering the research questions are described. Related work in literature is reviewed in Section 4. The findings of the research are discussed in Section 5. A conclusion is drawn in Section 6 and the research is completed with a discussion in Section 7.

2 PROBLEM STATEMENT & RESEARCH QUESTION(S)

An area of deficiency in many Big Data articles is the lack of theoretical discussion and connection between individual cases of Big Data usage and a greater understanding of Sports Analytics as a tool in the organisational decision-making process [44].

This aspect described by Watanabe et al. is exactly what we aim to focus on in this paper. The problem is that current literature is solely discussing single aspects of Big Data Management but, to the best of our knowledge, there is no literature on how to effectively connect those aspects together for a greater awareness of Big Data Management in the sports industry. Organisations have a lack of understanding of Big Data Strategies, i.e. how to implement, audit, design and maintain their Big Data projects. As a result of reviewing literature, best practices for Big Data Management are formulated which organisations could follow to use Big Data more effectively. Therefore, the following research questions have been formulated:

RQ – What are best practices for Big Data Management and its implementation in the sports industry?

In order to answer the main research question, two sub-research questions have been formulated to serve as a foundation upon which the main research builds.

RQ1 – What are the challenges that sports organisations encounter when implementing Big Data Strategies?

RQ2 – What recommendations can be given to solve the challenges of implementing Big Data Strategies in sports organisations?

3 METHODS OF RESEARCH

The research paper follows a semi-structured literature review to find common themes, form a theoretical foundation and retrieve qualitative information about best practices for Big Data Management in the sports industry. This subsequently involves an analysis of the relevant academic articles, conference papers, journals and other publications. The research mainly focuses on qualitative information gathering, both in theoretical and empirical studies.

The literature is retrieved from multiple sources using different keywords. This is done through multiple search engines such as *Scopus*, *Semantic Scholar*, *Google Scholar* and *Pubmed*. Search queries such as "Big Data Management in Sports", "Data Science/Visualisation in Sports" and "Sports Information Systems" have been used to find relevant literature. This resulted in 26 papers that are English-written, published in journals, peer-reviewed and relevant to the search queries. Subsequently, 16 useful references found in the first selection of papers have been added to the second selection of literature. Additionally, another eight papers were derived from the *Inciteful* tool, which suggests related papers based on the .bib file of the first selection round. This step of the semi-structured literature review resulted in a total of 50 papers, selected on their title, abstract, keywords, introduction and conclusion to ensure relevance to the topic. As a final step, the selected papers have been thoroughly analysed and this resulted in a stricter and final selection of papers. Subsequently, five papers have been removed from the list of relevant literature after this process because of too little information, too few citations or simply due to a lack of relevance. This process is depicted in Figure 1. In the end, the semi-structured literature review is conducted with 45 papers taken into consideration.

In addition to the literature review, we have developed an explanatory framework to demonstrate and summarise the best practices of Big Data Management in the sports industry. It provides a structured approach to guide sports organisations towards the successful implementation of Big Data Management. This explanatory framework serves as a helpful overview for understanding and visualizing the best practices of Big Data Management in the sports industry.

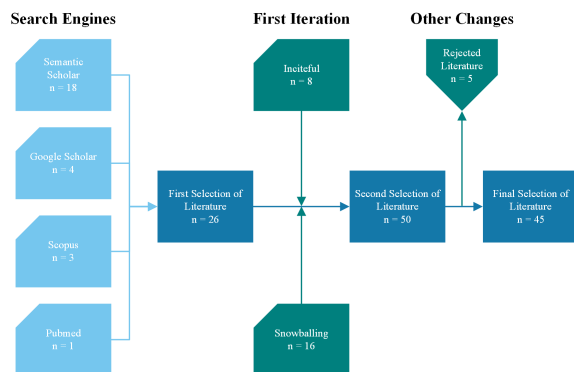


Fig. 1. Selection Process for Literature Review

4 LITERATURE REVIEW

The literature review is the foundation for identifying the best practices in Big Data Management and its recommendations. First, a definition of Big Data and its application in the sports industry is outlined by discussing the characteristics of Big Data defined in literature. Second, information management in sports is explored by discussing the systems used to manage sports information and how sports organisations are using such systems. Last, the applications of Sports Analytics are examined by demonstrating how Big Data is used in Sports Analytics. In summary, the three main research disciplines are Big Data, Sports Information Management and Sports Analytics. A visual orientation of the research streams is shown in Figure 2.

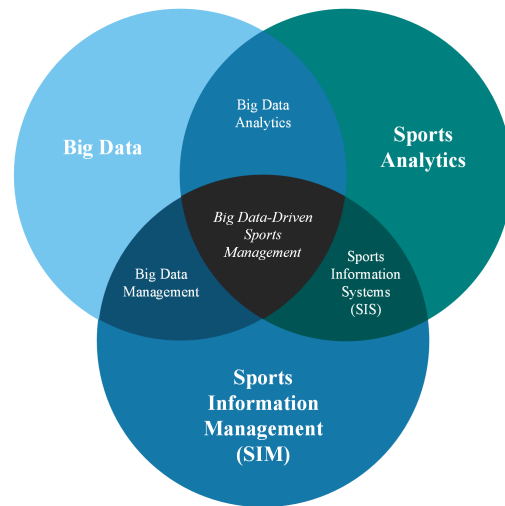


Fig. 2. Orientation of Big Data-Driven Sports Management

4.1 Big Data

As Big Data is a relatively new topic of research in many disciplines, one definition for Big Data is hard to find because it depends on the way it is used and in what context [13]. A definition for Big Data in the sports industry is therefore not yet commonly agreed upon. In order for this research to discuss the definition of Big Data, we aim to formulate a definition of Big Data within the scope of Big Data Management in the sports industry. Thus, as no universally agreed definition of Big Data is available, Big Data is rather described by its characteristics [32].

4.1.1 The Vs of Big Data Characteristics. Big Data is characterized and defined differently by different authors [45]. Often, research refers to the multiple 'Vs', the characteristics of Big Data. Some authors describe Big Data with three characteristics: volume, variety, velocity [11, 13, 31, 32, 44]. Others extend this list of characteristics with two more Vs, namely value and veracity [4, 11, 13, 19, 28, 41]. Additionally, Ranjan & Foropon even describe seven totally different Vs, i.e. variability, veracity, visualization, value, validity, vulnerability, and volatility [31]. The most excessive example is the

description of Big Data characteristics with the ten Vs: venue, variety, volume, variability, value, validity, veracity, velocity, vocabulary, and vagueness [45]. For our research, the application of Big Data within the context of Big Data Management in the sports industry is described. This application in sports is based on the five Vs characteristics of Big Data, with the aim to relate it to a practical definition, as preferred by researchers [13]. The five Vs taken into account for this definition are volume, variety, velocity, value and veracity. That is because the nature of Big Data in sports not only involves volume, variety and velocity but also value and veracity should be included to provide a more comprehensive understanding of Big Data in sports. Moreover, since the application of Big Data is relatively new in sports [44], a description based on these five well-established characteristics in literature is justified. First, volume refers to the size and quantity of the data set [32, 44]. It involves unimaginable amounts of data [45] that are generated each second that is too massive to be processed by traditional technologies [11]. Variety refers to different types and formats of data and data sources [32, 44], such as structured, semi-structured and unstructured data sources [11, 19, 41]. Velocity describes the speed with which novel data is generated [32], delivered and collected [19, 44]. It involves the handling of incoming data sets at varying speeds and activity peaks [45]. Value is the worth carried by data [19]. The advantageous insights are gained through data analysis [11]. Data has intrinsic value but is only discovered when extracted through data analysis [28]. Last, veracity outlines the degree of precision, accuracy and exactness carried by data [19]. Data may be uncertain, untruthful and non-reliable [11]. It questions the sources, methodologies and technologies of the data [28]. The application of Big Data in sports for our research is based on the described characteristics of the five Vs of Big Data.

4.1.2 Big Data in the Sports Industry. First, the aforementioned five Vs need to be defined within the scope of our research. In this section, the characteristics of Big Data are described and related to Big Data Management in the sports industry.

- **Volume** – Big Data is growing increasingly impactful across the sporting landscape [22]. The sports industry generates a massive volume of data, containing players' physical activity, and behavioural actions between athletes allowing time, description, and count data on action to be recorded [20]. This data can help decision-making in sports [12].
- **Variety** – There are many relationships and numerous entities in sports Big Data [4]. This involves physical fitness, physical exercise behaviour, personal information and competition results. The data can be structured (e.g., general statistics), semi-structured (e.g., wearable data, GPS information), or unstructured (e.g., social media) [11]. Sports organisations use this data to assess and enhance the performance of players and teams [25].
- **Velocity** – This can be explained by the growth of sports data [4] which is updated in real-time. For all major sports, analysts may frequently extract enormous numbers of data [20]. This data enters the storage in bursts, from delayed data of post-training and competition analysis to real-time streams during a live match [32].

- **Value** – The extraction of value from sports Big Data is a fundamental objective in the sports industry. It is being used to evaluate the physical, mental and social state of players [1]. Connecting the wide variety of data in sports and using it for analysis assists organisations in becoming more successful [20]. These insights allow informed decision-making within sports organisations.
- **Veracity** – Ensuring data quality is essential in the sports industry. Some data sources may be unreliable, which may result in untrustworthy data analysis outcomes [11, 15]. The accuracy of the analysis depends on the veracity of the source data [19], therefore sensor data, biases in data and validation of data quality are of utmost importance.

In summary, the five Vs of Big Data – volume, variety, velocity, value and veracity – provide a theoretical foundation for understanding the characteristics of Big Data and how to manage and use Big Data in the sports industry. Therefore, within the scope of Big Data in the sports industry, Big Data refers to the vast amount of data of various types rapidly generated by multiple sources to gain valuable insights through informed decision-making while ensuring its veracity.

4.2 Sports Information Management

The sports industry is rapidly evolving due to the constant development of Sports Information Management (SIM) and IT [18]. Jinguo describes that managing sports information "to manage work in the sports field will be an important means for Sports Information Management" [18]. The presence of individual IT systems in sports organisations has expanded in recent years, causing a fragmented situation that is difficult to manage without a centralised Sports Information System (SIS) [6]. SIS can be described as an application of SIM and Sports Analytics, being an application of both disciplines. This section elaborates on SIS within sports organisations, with a precise focus on the organisational aspects and the impact of SIS within these organisations.

4.2.1 Sports Information Systems. As literature is still relatively limited on information systems in the sports industry, multiple authors have different definitions of SIS. An accepted definition of SIS by many scholars is the one from Blobel et al., describing that "Sports Information Systems are information processing systems based on Information Technology which, together with the associated sports-organisational resources, serve to handle data from various sources in sport organisations within one system and to make this information available user-specific and location-independent via user interface" [7]. In addition, it should span across whole organisations by prioritising sport-specific activities, taking into account the need for scalability and integration of other methods [7], such as Big Data [35]. Shuangming et al. point out that due to the rapid development of technology in sports, it is essential for sports organisations to integrate Big Data as an application of SIS [35]. Furthermore, a SIS must not be seen as a closed system but rather as a part of "an integral information system of sports that connects all other information subsystems" [23], such that the entirety of a sports organisation can access information for management purposes.

When talking about the management of such a SIS, scholars refer to SIM. According to Qi & Wang, SIM "can be defined as the management of activities of information resources and its exploitation and utilisation" [30]. Thus, SIM is the application of information management in the sports industry [18]. Managing information systems is of great importance to sports organisations because the existing information systems are often limited in functionality in sports-specific areas due to development and the adaptation of general-purpose software solutions [6]. From a sports organisation perspective, the utilisation and development of resources for SIS and the management hereof is of high value for sports organisations and should therefore be effectively managed accordingly [30].

4.2.2 Management of Sports Information Systems. To succeed in contemporary data-driven landscapes, managing information technologies and their data is essential for sports organisations. IT and its data have become an intrinsic element of every organisation, including sports organisations [26]. This section explores how SIS is used on organisational and strategy-related levels. Furthermore, the role of data in SIS in sports organisations is examined.

Information management in sports is determined by the characteristics of the organisational structure. That is, the way organisations produce, collect, exchange, store, distribute and utilise data and information [23]. The processes, methods and technological basis of information systems and their data exchange in organisations depend on the level of management [14]. Ordinary users of such systems, however, lack the specialist knowledge required to use this data to an advantage [9]. Sports organisations should aim for the extraction of value from data as it aids the growth of the organisation [19]. When talking about Big Data in SIS, organisations do not recognise the strategic value it can offer, together with the fact that managers are usually not willing to implement the required changes at technological, organisational and environmental levels [11, 42].

4.3 Sports Analytics

The third and last research stream of this literature review is Sports Analytics. Passfield & Hopker explain that "gaining knowledge and wisdom from data is challenging, but could spawn a new discipline in the sports sciences, that of Sports Analytics" [27]. For the definition, this paper adopts the definition provided by Alamar, who defines Sports Analytics as "the management of organised historical data, the use of predictive analytical models that employ the data, and the usage of information systems to inform decision-makers and permit them to aid their establishments in the attainment of a competitive benefit on the ground" [3, 28].

Having established a definition, this section delves into the applications of Big Data in sports organisations and how they impact various aspects of sports. By examining sports data analytics and visualisation, data management and data-driven decision-making, this section provides a deeper understanding of how Big Data is used for Big Data-Driven Sports Management.

4.3.1 Visualisation of Sports Analytics. The sports industry has made notable advancements in Sports Analytics, becoming the status quo within the majority of sports organisations. Due to the revolutionary development of technology, significant change can

be seen in the analysis and visualisation of sports data [5]. According to Horkey & Pelka, in terms of data visualisation, interviewed experts argue that data visualisation, compared to other methods such as textual reports, is the best way to explain more complex circumstances [16]. Moreover, there is a great availability of data, both publicly and commercially. The possibility to differentiate and filter on data-sets offers a broad availability of different types of analysis [16]. Perin et al. have derived three types of sports data: box score data, tracking data and meta-data. This separation of data types is needed to determine the required technique to visualise the data. By combining the three types of sports data, extensive and immense possibilities to visualise data will arise [29]. In its turn, data analytics can also assist decision-making processes in sports organisations to realise more benefits [41]. To conclude, data visualisation is necessary to support the decision-making process for sports scientists, given that the wide availability of data for Sports Analytics is properly used and organised.

4.3.2 Management of Big Data for Sports Analytics. A conclusion drawn by Favaretto et al. is that for the management of Big Data, "most of the researchers preferred a more practical definition, linking it to processes such as data collection and data processing" [13]. As indicated by Rein & Memmert, due to the improvement of advanced wearables the volume of data available is becoming increasingly challenging to manage [32]. The collection and processing of data is therefore important for managing Big Data projects. However, the capability to assemble information has not been matched with the capability of processing it in a purposeful manner [19]. To produce useful data insights, raw data must be converted into information, information into knowledge, and knowledge into wisdom [33]. In order to perform valid analyses, the quality of data must be carefully considered, hence ensuring its veracity as described in Section 4.1.1. In sports organisations, low data usage, unscientific data classification, significant data redundancy and wide variations in data formats are provoking low-quality data [35]. High-quality data is essential for gaining accurate and reliable insights and informed decision-making. However, due to the heterogeneity and unstructured nature of data, managing Big Data is a challenging process [36]. According to Bai & Bai, managing Big Data in sports "is crucial to the success of the national sports industry, teams, and individuals" [4]. In fact, the exploitation of potential value by the enhancement of data quality and accessibility of data for decision-making processes is the primary goal of sports Big Data Management [4]. Furthermore, Bai & Bai provide insight into the three aspects of sports Big Data Management, namely (1) Big Data acquisition, (2) Big Data labelling and (3) the improvement of existing data. The acquisition of Big Data involves the secure collection of sports Big Data which is used for Big Data Analytics. Alongside data acquisition, data labelling is done for every piece of information that is extracted to build an accurate database of high quality. Last, enhancement of the quality of existing labels to get high-quality data labels and reduce noisy data with incorrect labels is a crucial step in improving the overall data quality [4]. By managing the steps described by Bai & Bai, a foundation is formed for Big Data-driven decision-making with accurate and reliable data that is improved over time.

4.3.3 Big Data-Driven Sports Management. To gain valuable insights into problems, the application of Big Data is an advancing topic which helps to make more informed and fundamentally strong decisions [5]. Making informed decisions in professional sports should be supported by (1) data collection and organisation, (2) analytical models to gain insights and (3) an interface for the communication of information [43]. Ward et al. explain that there are two types of analytical models in sports organisations, i.e. fast and slow analysis. Fast analysis implies a decision-making situation which should be resolved short-term, on a high-pressure environmental level. Slow analysis requires scientific examination to achieve a deeper understanding of the problem encountered. Furthermore, Ward et al. suggest that combining the three areas in data into a communal platform which supports decision-making on all levels in a professional sports organisation, while also taking the nature of uncertainty into consideration, could result in better decision-making in this complex, fast-moving environment [43]. Ranjan & Foropon have defined value-creation through the means of decision-making as Competitive Intelligence (CI). Hence CI can be defined as "the collection, analysis, interpretation, and dissemination of strategic information at the right time for use in the decision-making process" from an organisational perspective [2, 31]. Multiple scholars concluded that Big Data is often regarded as a valuable asset for making strategic decisions in organisations [31].

We used the literature reviewed in this section for the formulation of best practices, hence answering the research questions based on analysis of existing literature. The found common themes, theoretical foundation and qualitative information of the reviewed literature, in combination with the formulated best practices, are thereafter used to develop an explanatory framework.

5 FINDINGS

In this section, the identified best practices for Big Data Management in the sports industry are described. These best practices are inspired by the literature review in Section 4. It is crucial to discover and apply these best practices to effectively manage Big Data in the sports industry, given that its volume and complexity keep increasing. By evaluating the current landscape of Big Data Management in sports, this section aims to summarise the challenges of Big Data for sports organisations and provide recommendations to maximise the full potential of their data resources. The identified best practices could serve as practical guidelines for professionals and stakeholders in the sports industry who are involved in Big Data Management. The implementation and understanding of these best practices could enhance sports organisations' Big Data Management strategies and ultimately gain an advantage over their competitors.

5.1 Challenges of Implementing Big Data Strategies

As claimed by Abeza et al., the adoption and utilisation of Big Data in the sports industry have numerous challenges [10, 39], such as a lack of experienced employees [8, 9], data integration, poor data warehouse architecture, data privacy and security concerns and enormous data synchronisation [38]. Additionally, Cabrera-Sánchez & Villarejo-Ramos argue that "further implementation challenges arise from a lack of analysis techniques, resistance to

change, fear and anxiety, and technology limitations" [10]. Blobel & Lames have determined two different types of obstacles within the sports industry when it comes to implementing SIS, namely strategic and organisational problems. The strategic challenges include the underestimation of the demands of the task of implementing SIS, miscommunication due to the different stakeholders involved and as a result, there is a lack of integration of information systems in sports [6]. The organisational problems relate to the non-adjustment of organisational aspects to the new circumstances of SIS implementation [17]. In a study on the effective adoption of Big Data, Surbakti et al. have established 13 organisational aspects that are of influence on this non-adjustment process, described by Jayal et al.: organisational cultural competence, talent management, change management, strategy alignment, project management, performance management, organisational structure and size, interdepartmental collaboration, communication, top management support, environmental effect, clear goals, and focus on innovation [40]. Moreover, Blobel & Lames address that often SIS are not integrated actively within the different departments and their systems, such that the systems run concurrently [6]. Another identified organisational problem is that poor information management is usually caused by the deployment of SIS without obligations being made explicit: all stakeholders should work with, supply data to and complete the necessary duties of SIS implementation [6]. Last, the environment of an organisation influences the course of the Information System, meaning that the fast-changing environment of sports must constantly be aligned with the development, maintenance and configuration of SIS [23].

Based on the comprehensive review of literature, the challenges of Big Data Management and implementing Big Data Strategies can be summarised. First, a SIS is often not actively integrated into sports organisations but runs alongside the organisation and its processes. Second, there is usually a lack of expertise and experience in Big Data Management that can be divided into two subcategories: (1) the technological part relates to poor data integration & collection, inferior data warehouse architecture, lack of data analysis techniques and other limitations of technology for handling Big Data. In addition, the (2) managerial part relates to the non-employment of experts, not authorising employees to make important decisions regarding the implementation of Big Data projects. Third, the absence of top-management support in an organisation for Big Data Management results in resistance to change because of fear and anxiety about the unknown. Again, this can be divided into two subcategories: (1) strategic problems relate to the underestimation of implementing Big Data Strategies and miscommunication leading to poor integration of Big Data Management in sports organisations. In addition, the (2) organisational problems relate to sports organisations that often do not adjust to the new work environment involving Big Data Strategies. Moreover, there are no clear task distinctions defined when it comes to managing Big Data. Last, sports is an ever-changing environment and therefore Big Data Strategies must be continually coordinated to ensure legitimate development and progress.

These challenges underline the importance of considering the technological, managerial, strategic, and organisational aspects of Big Data Management to effectively implement and leverage

Big Data Strategies within sports organisations. The framework depicted in Figure 3 takes these challenges into account and provides recommendations to address them in Section 5.2.

5.2 Recommendations for Implementing Big Data Strategies

The aforementioned challenges of implementing Big Data Strategies demonstrate the need for recommendations to effectively overcome and address these obstacles in sports organisations. This section, therefore, aims to provide recommendations to help sports organisations tackle the identified challenges.

Firstly, it is essential to develop a clear understanding of the goals and objectives regarding Big Data Management of a sports organisation. This also involves a proper understanding of Big Data in general, to ensure that all involved stakeholders have the same perception of the new strategy that will be followed within the organisation. In addition, organisations should promote a data-driven culture amongst their employees at all levels. That can be achieved by training programs to supply employees with the necessary skills for Big Data Management. Moreover, interdisciplinary collaboration between data scientists, sports analysts, athletes, trainers, coaches and management should be encouraged to utilise Big Data to its full potential.

In line with the first recommendations, sports organisations should define their strategic priorities, which should align with the Big Data strategies and their impact on the organisation. The strategic priorities should outline roles, responsibilities, and policies to create an environment in which Big Data is effectively managed. Additionally, organisations should establish a comprehensive Big Data governance policy to guide the management and utilisation of Big Data. This policy should include guidelines and procedures for Big Data collection, storage, integration, management and decision-making.

Thirdly, the integration and the assurance of the quality of Big Data are fundamental to maximise its value. The implementation of integration processes for diverse Big Data sources and controlling quality will enhance the accuracy and reliability of data-driven insights derived from Big Data. These processes should be monitored regularly to maintain high standards. In addition, investments should be made in the technology infrastructure to guarantee that these processes can be executed effectively. This infrastructure should be scalable and reliable to handle large volumes of data and complex analysis.

Finally, Big Data Management should be continuously monitored and evaluated to determine its impact and effectiveness. This could be done by determining key performance indicators and other metrics to assess the success of the implemented strategies, as well as to identify areas of improvement. The approaches taken for Big Data Management should be able to adapt to sudden changes in sports organisations to ensure the ongoing optimisation of best practices for managing Big Data. That is to maximise and unlock the full potential of Big Data and drive innovation and development in the dynamic landscape of sports.

5.3 Explanatory Framework

In this section, the comprehensive framework is introduced which aims to provide a structured approach to Big Data Management for stakeholders in the sports industry. This framework serves as a visual representation of the research streams in the literature review, key findings, challenges, and recommendations identified in this research. The framework is illustrated in Figure 3.

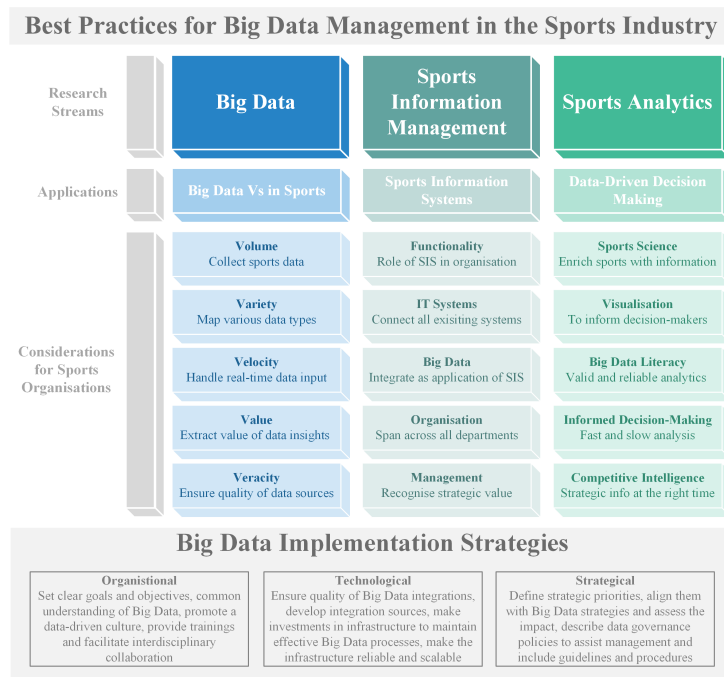


Fig. 3. Framework for best practices of Big Data Management.

6 CONCLUSION

In conclusion, this research provides valuable insights into the management of Big Data in the sports industry. Through a comprehensive literature review, key findings have been identified, highlighting best practices for Big Data Management and the challenges faced by sports organisations in implementing Big Data strategies. The literature review covered the analysis of various research streams such as Big Data, Sports Information Management and Sports Analytics. In addition to the addressed challenges, recommendations have been provided to overcome these challenges. The challenges involve technological, managerial, strategic, and organisational aspects, emphasising the need for an integrated approach to address these obstacles.

The recommendations offered in this study empower sports organisations to overcome the challenges and effectively leverage Big Data to gain an advantage over their competitors. By promoting a data-driven culture, defining strategic priorities, ensuring data governance policies, investing in appropriate technological infrastructure, adopting Big Data techniques, and integrating tools to monitor and

evaluate these processes, sports organisations can maximise the potential of their Big Data resources.

The framework presented, based on the challenges and recommendations, in this research provides a comprehensive overview for understanding the management of Big Data in the sports industry. It visually represents the findings, including the identified challenges, recommendations and best practices for effective implementation.

This research serves as a practical guide for professionals and stakeholders involved in Big Data Management in the sports industry, enabling them to enhance their Big Data strategies and be competitive in an ever-changing data-driven landscape. Overall, the findings and recommendations presented in this research contribute to the advancement of Big Data Management practices in the sports industry.

7 DISCUSSION

7.1 Limitations

This research has provided valuable insights into the best practices for Big Data Management in the sports industry. However, it is important to acknowledge certain limitations that may have impacted the research. Due to the limited timeframe in which this research was conducted, only a selection of literature has been analysed based on the set criteria. Other research disciplines and methodologies could have been explored and included in this research, however, this was difficult to carry out given the timeframe.

7.2 Future Work

While this research has shed light on various aspects of Big Data Management in the sports industry, there are several possibilities for extending this research. We recommend other scholars to further investigate and develop the following areas.

As Big Data and its application in Sports Analytics is still in its infancy, one could delve deeper into other data analysis techniques such as artificial intelligence (AI), predictive modelling and other machine learning algorithms. This can provide better insights into, for example, player performance, injury prevention and tactical game strategies.

Another opportunity is to investigate and perform case studies in sports organisations. This can be extended with comparative studies between different sports organisations. A case study can contribute to a better understanding of how sports organisations actually work in regard to Big Data Management, rather than taking conclusions from literature for granted. Applying the proposed framework to sports organisations in case studies may validate the effectiveness of such a framework.

As discussed before, sports and technology are ever-changing environments that have rapidly evolved over time. Therefore, the identified best practices and proposed framework may not be up-to-date anymore in the future. Repeating this research in the future could result in new valuable insights such as new best practices or other progress in this field of study over time.

ACKNOWLEDGMENTS

I would like to thank my supervisor, Patricia Rogetzer, for her valuable guidance, insightful discussions, and constructive feedback throughout the process of conducting this research.

REFERENCES

- [1] Gashaw Abeza, Norm O'Reilly, John Nadeau, and Yann Abdourazakou. 2022. Big data in professional sport: the perspective of practitioners in the NFL, MLB, NBA, and NHL. *Journal of Strategic Marketing* 0, 0 (Aug. 2022), 1–21. <https://doi.org/10.1080/0965254X.2022.2108881> Publisher: Routledge _eprint: <https://doi.org/10.1080/0965254X.2022.2108881>.
- [2] Abhilash Acharya, Sanjay Kumar Singh, Vijay Pereira, and Poonam Singh. 2018. Big data, knowledge co-creation and decision making in fashion industry. *International Journal of Information Management* 42 (Oct. 2018), 90–101. <https://doi.org/10.1016/j.ijinfomgt.2018.06.008>
- [3] Benjamin C. Alamar. 2013. *Sports Analytics: A Guide for Coaches, Managers, and Other Decision Makers*. Columbia University Press. Pages: 152 Pages.
- [4] Zhongbo Bai and Xiaomei Bai. 2021. Sports Big Data: Management, Analysis, Applications, and Challenges. *Complexity* 2021 (Jan. 2021), e6676297. <https://doi.org/10.1155/2021/6676297> Publisher: Hindawi.
- [5] Vishvak Bhatt, Udgam Aggarwal, and Vinoth Kumar. 2022. Sports Data Visualization and Betting. In *2022 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)*. 1–6. <https://doi.org/10.1109/SMARTGENCON56628.2022.10083831>
- [6] Thomas Blobel and Martin Lames. 2020. A Concept for Club Information Systems (CIS) - An Example for Applied Sports Informatics. *International Journal of Computer Science in Sport* 19, 1 (2020), 102–122. <https://doi.org/doi:10.2478/ijcss-2020-0006>
- [7] Thomas Blobel, Martin Rumo, and Martin Lames. 2021. Sports Information Systems: A systematic review. *International Journal of Computer Science in Sport* 20, 1 (2021), 1–22. <https://doi.org/doi:10.2478/ijcss-2021-0001>
- [8] John Bone, Chukwuemeka David Emele, Adeniyi Abdul, George Coghill, and Wei Pang. 2016. The social sciences and the web: From 'Lurking' to interdisciplinary 'Big Data' research. *Methodological Innovations* 9 (Jan. 2016), 2059799116630665. <https://doi.org/10.1177/2059799116630665> Publisher: SAGE Publications Ltd.
- [9] Richard Brunauer, Wolfgang Kremser, and Thomas Stöggel. 2020. From Sensor Data to Coaching in Alpine Skiing – A Software Design to Facilitate Immediate Feedback in Sports. In *Proceedings of the 12th International Symposium on Computer Science in Sport (IACSS 2019) (Advances in Intelligent Systems and Computing)*, Martin Lames, Alexander Danilov, Egor Timme, and Yuri Vassilevski (Eds.). Springer International Publishing, Cham, 86–95. https://doi.org/10.1007/978-3-030-35048-2_11
- [10] Juan-Pedro Cabrera-Sánchez and Ángel F. Villarejo-Ramos. 2020. Acceptance and use of big data techniques in services companies. *Journal of Retailing and Consumer Services* 52 (Jan. 2020), 101888. <https://doi.org/10.1016/j.jretconser.2019.101888>
- [11] Gianluca Elia, Gloria Polimeno, Gianluca Solazzo, and Giuseppina Passiante. 2020. A multi-dimension framework for value creation through big data. *Industrial Marketing Management* 90 (Oct. 2020), 508–522. <https://doi.org/10.1016/j.indmarman.2019.08.004>
- [12] Rob Euman and José Abdelnour-Nocera. 2013. Data Visualisation, User Experience and Context: A Case Study from Fantasy Sport. In *Human-Computer Interaction. Users and Contexts of Use (Lecture Notes in Computer Science)*, Masaaki Kurosu (Ed.). Springer, Berlin, Heidelberg, 146–155. https://doi.org/10.1007/978-3-642-39265-8_16
- [13] Maddalena Favaretto, Eva De Clercq, Christophe Olivier Schneble, and Bernice Simone Elger. 2020. What is your definition of Big Data? Researchers' understanding of the phenomenon of the decade. *PLOS ONE* 15, 2 (Feb. 2020), e0228987. <https://doi.org/10.1371/journal.pone.0228987>
- [14] Vlado Galicic and Ljubica Pilepic. 2007. The role of logistics information system in the business-decision process. *Tourism and hospitality management* 13, 3 (2007), 571–582. <https://doi.org/10.20867/thm.13.3.4>
- [15] Amir Gandomi and Murtaza Haider. 2015. Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management* 35, 2 (April 2015), 137–144. <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>
- [16] Thomas Horkey and Philipp Pelka. 2017. Data Visualisation in Sports Journalism. *Digital Journalism* 5, 5 (May 2017), 587–606. <https://doi.org/10.1080/21670811.2016.1254053> Publisher: Routledge _eprint: <https://doi.org/10.1080/21670811.2016.1254053>.
- [17] Ambikesh Jayal, Allistair McRobert, Giles Oatley, and Peter O'Donoghue. 2018. *Sports Analytics: Analysis, Visualisation and Decision Making in Sports Performance*. Routledge, London. <https://doi.org/10.4324/9781315222783>
- [18] Dong Jinguo. 2020. The Theoretical Construction and Application System Development Study of Sports Information Management. In *2020 IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS)*. 526–529. <https://doi.org/10.1109/ICPICS49253.2020.9371111>

- [//doi.org/10.1109/ICPICS50287.2020.9202135](https://doi.org/10.1109/ICPICS50287.2020.9202135)
- [19] Gagandeep Kaur and Gagandeep Jagdev. 2020. Analyzing and Exploring the Impact of Big Data Analytics in Sports Science. In *2020 Indo – Taiwan 2nd International Conference on Computing, Analytics and Networks (Indo-Taiwan ICAN)*. 218–224. <https://doi.org/10.1109/Indo-TaiwanICAN48429.2020.9181320>
- [20] Aijun Liu, Rajendra Prasad Mahapatra, and A. V. R. Mayuri. 2021. Hybrid design for sports data visualization using AI and big data analytics. *Complex & Intelligent Systems* (Oct. 2021). <https://doi.org/10.1007/s40747-021-00557-w>
- [21] Leonardo Jose Mataruna-Dos-Santos, Alessio Faccia, Hussein Muñoz Helú, and Mohammed Sayeed Khan. 2020. Big Data Analyses and New Technology Applications in Sport Management, an Overview. *Proceedings of the 2020 International Conference on Big Data in Management (May 2020)*, 17–22. <https://doi.org/10.1145/3437075.3437085> Conference Name: ICBDM 2020: 2020 International Conference on Big Data in Management ISBN: 9781450375061 Place: Manchester United Kingdom Publisher: ACM.
- [22] Brad Millington and Rob Millington. 2015. ‘The Datafication of Everything’: Toward a Sociology of Sport and Big Data. *Sociology of Sport Journal* 32, 2 (June 2015), 140–160. <https://doi.org/10.1123/ssj.2014-0069> Publisher: Human Kinetics, Inc. Section: Sociology of Sport Journal.
- [23] J. Miočić, Lj. Zekanić-Korona, and Boris Bosančić. 2019. Information Systems in Sports Organizations: Case Study of the Sports Association of the City of Zadar. In *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*. 1362–1367. <https://doi.org/10.23919/MIPRO.2019.8756923> ISSN: 2623-8764.
- [24] Michael Mondello and Christopher Kamke. 2014. The Introduction and Application of Sports Analytics in Professional Sport Organizations. *Journal of Applied Sport Management* 6, 2 (Jan. 2014). <https://trace.tennessee.edu/jasm/vol6/iss2/11>
- [25] Elia Morgulev, Ofer H. Azar, and Ronnie Lidor. 2018. Sports analytics and the big-data era. *International Journal of Data Science and Analytics* 5, 4 (June 2018), 213–222. <https://doi.org/10.1007/s41060-017-0093-7>
- [26] Tanty Oktavia, Ford Lumban Gaol, Takaaki Hosoda, and Arsyian Syahir. 2020. Sport Science Model to Support the Professional Sports Organization Decision Making. In *2020 International Conference on Information Management and Technology (ICIMTech)*. 599–604. <https://doi.org/10.1109/ICIMTech50083.2020.9211238>
- [27] Louis Passfield and James Hopker. 2016. A Mine of Information: Can Sports Analytics Provide Wisdom From Your Data? *International Journal of Sports Physiology and Performance* 12 (Dec. 2016), 1–17. <https://doi.org/10.1123/ijssp.2016-0644>
- [28] Devansh Patel, Dhwanil Shah, and Manan Shah. 2020. The Intertwine of Brain and Body: A Quantitative Analysis on How Big Data Influences the System of Sports. *Annals of Data Science* 7, 1 (March 2020), 1–16. <https://doi.org/10.1007/s40745-019-00239-y>
- [29] C. Perin, R. Vuillemot, C. D. Stolper, J. T. Stasko, J. Wood, and S. Carpendale. 2018. State of the Art of Sports Data Visualization. *Computer Graphics Forum* 37, 3 (2018), 663–686. <https://doi.org/10.1111/cgf.13447> <https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.13447>
- [30] Ning Qi and Le Wang. 2014. Constructing of sports information management theories and developing research of application system. In *2014 IEEE Workshop on Advanced Research and Technology in Industry Applications (WARTIA)*. 715–717. <https://doi.org/10.1109/WARTIA.2014.6976366>
- [31] Jayanthi Ranjan and Cyril Foroapon. 2021. Big Data Analytics in Building the Competitive Intelligence of Organizations. *International Journal of Information Management* 56 (Feb. 2021), 102231. <https://doi.org/10.1016/j.ijinfomgt.2020.102231>
- [32] Robert Rein and Daniel Memmert. 2016. Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science. *SpringerPlus* 5, 1 (Dec. 2016), 1410. <https://doi.org/10.1186/s40064-016-3108-2>
- [33] Jennifer Rowley. 2007. The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science* 33, 2 (April 2007), 163–180. <https://doi.org/10.1177/0165551506070706> Publisher: SAGE Publications Ltd.
- [34] William A. Sands, Ashley A. Kavanaugh, Steven R. Murray, Jeni R. McNeal, and Monem Jenni. 2017. Modern Techniques and Technologies Applied to Training and Performance Monitoring. *International Journal of Sports Physiology and Performance* 12, s2 (April 2017), S2–72. <https://doi.org/10.1123/ijssp.2016-0405> Publisher: Human Kinetics, Inc. Section: International Journal of Sports Physiology and Performance.
- [35] Wang Shuangming, Shi Jianwei, Fu Shiqiu, and Xu Wanlin. 2018. Research on the Construction of Sports Resources Information Platform Based on Big Data. *International Journal of Advanced Network, Monitoring and Controls* 3, 2 (2018), 28–32. <https://doi.org/doi:10.21307/ijanmc-2018-036>
- [36] Aisha Siddiq, Ibrahim Abaker Targio Hashem, Ibrar Yaqoob, Mohsen Marjani, Shahabuddin Shamshirband, Abdullah Gani, and Fariza Nasaruddin. 2016. A survey of big data management: Taxonomy and state-of-the-art. *Journal of Network and Computer Applications* 71 (Aug. 2016), 151–166. <https://doi.org/10.1016/j.jnca.2016.04.008>
- [37] Nitin Singh. 2020. Sport Analytics: A Review. *The International Technology Management Review* 9, 1 (Sept. 2020), 64–69. <https://doi.org/10.2991/itm.r.k.200831.001> Publisher: Atlantis Press.
- [38] Uthayasankar Sivarajah, Muhammad Mustafa Kamal, Zahir Irani, and Vishanth Weerakkody. 2017. Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research* 70 (Jan. 2017), 263–286. <https://doi.org/10.1016/j.jbusres.2016.08.001>
- [39] Ramón Spaaij and Ansgar Thiel. 2017. Big data: critical questions for sport and society. *European Journal for Sport and Society* 14, 1 (Jan. 2017), 1–4. <https://doi.org/10.1080/16138171.2017.1288374> Publisher: Routledge [_eprint: https://doi.org/10.1080/16138171.2017.1288374](https://doi.org/10.1080/16138171.2017.1288374)
- [40] Feliks P. Sejahtera Surbakti, Wei Wang, Marta Indulska, and Shazia Sadiq. 2020. Factors influencing effective use of big data: A research framework. *Information & Management* 57, 1 (Jan. 2020), 103146. <https://doi.org/10.1016/j.im.2019.02.001>
- [41] Poojan Thakkar and Manan Shah. 2021. An Assessment of Football Through the Lens of Data Science. *Annals of Data Science* 8, 4 (Dec. 2021), 823–836. <https://doi.org/10.1007/s40745-021-00323-2>
- [42] S. Verma and S.S. Bhattacharyya. 2017. Perceived strategic value-based adoption of Big Data Analytics in emerging economy: A qualitative approach for Indian firms. *Journal of Enterprise Information Management* 30, 3 (2017), 354–382. <https://doi.org/10.1108/JEIM-10-2015-0099>
- [43] Patrick Ward, Johann Windt, and Thomas Kempton. 2019. Business Intelligence: How Sport Scientists Can Support Organization Decision Making in Professional Sport. *International Journal of Sports Physiology and Performance* 14, 4 (April 2019), 544–546. <https://doi.org/10.1123/ijssp.2018-0903> Publisher: Human Kinetics Section: International Journal of Sports Physiology and Performance.
- [44] Nicholas M. Watanabe, Stephen Shapiro, and Joris Drayer. 2021. Big Data and Analytics in Sport Management. *Journal of Sport Management* 35, 3 (April 2021), 197–202. <https://doi.org/10.1123/jsm.2021-0067> Publisher: Human Kinetics Section: Journal of Sport Management.
- [45] Ana Šuštaršič, Mateja Videmšek, Damir Karpljuk, Ivan Miloloža, and Maja Meško. 2022. Big Data in Sports: A Bibliometric and Topic Study. *Business Systems Research Journal* 13, 1 (June 2022), 19–34. <https://doi.org/10.2478/bsrj-2022-0002>