The Impact of Big Data and Sports Analytics on Professional Football: A Systematic Literature Review

Tim A. Herberger and Christoph Litke

1 Introduction

The usage potential of big data analysis is not limited to industrial companies only. In 2001, Billy Beane, the manager of Oakland Athletics, recognized big data analysis and successfully used it in the baseball professional league. Beane revolutionized the baseball game with innovative data-driven management style. He made decisions based only on statistical analysis and achieved a unique winning streak of 20 games won. This success story is described in the book titled *Moneyball: The Art of Winning an Unfair Game* and illustrates how statistics could make sport predictable (Lewis, 2004). "Moneyball" took a decisive turn in high-performance sport and influenced it beyond baseball (Gerrard, 2016: 213). The increasing awareness of a possible success through the combination of increased computing power and better availability of data rapidly increased the proportion of quantitative analyses in sports and led to the widespread use of data analytics in professional sports (McHale & Relton, 2018: 339ff).

The use of data analytics based on innovative technologies is also widely used in professional football, as football companies expect data analytics to improve their competitiveness and thus secure their medium- and long-term sporting and economic success (Möller & Schönefeld, 2011: 1; Gassmann & Perez-Freije, 2011: 394; Werner, 2017: 17ff.). However, to design innovation processes effectively and efficiently and thus ensure the innovation success of a football company, suitable analysis methods are needed, that is, a further clarification of the rather broad term

T. A. Herberger (✉) · C. Litke
Chair of Business Administration, specifically Entrepreneurship, Finance and Digitalization, Andrássy University, Budapest, Hungary
e-mail: tim.herberger@andrassyuni.hu

C. Litke
e-mail: christoph.litke@icloud.com

T. Herberger and J. Dötsch (eds.), *Digitalization, Digital Transformation and Sustainability in the Global Economy*, Springer Proceedings in Business and Economics, https://doi.org/10.1007/978-3-030-77340-3_12

data analytics toward sports analytics (Reichmann et al., 2017). However, the implementation of new technologies and thus new process structures in a company – also in a football company – poses great challenges to the various stakeholders (e.g., owners, top management, athletes, and coaches). This is all the more true for the more disruptive and penetrating innovation. These challenges are reflected not only in the use of financial resources but also in the coverage of personnel capacities and competencies.

The analysis and target-group-specific evaluation and processing of the large data collected during a football match, which are still unspecific and unstructured to a large extent, are only possible with appropriate software tools. IT providers such as SAP and IBM provide specific cloud solutions for big data and sports analytics at football companies. For example, "SAP Sports One" offers Team Management, Training Planning, Fitness, Performance Analysis, Scouting, and Talent Finding modules via a uniform platform (Biermann, 2018a: 40).

However, can these additional instruments (better) answer the main question in football against the background of a balanced cost–benefit ratio: How do you win matches? In this study's context, the following questions will be examined based on a structured literature review (SLR) according to Massaro et al. (2016):

1. *How does the use of big data and sports analytics change professional football?*
2. *Can big data and sports analytics increase the team's competitiveness in football?*
3. *Which chances and risks are to be considered with the use of big data and sports analytics for the different stakeholders in a football company?*

The remainder of the paper is structured as follows: Sect. 2 defines the essential concepts and theoretical framework. Section 3 presents the methodology of the SLR. Section 4 discusses the results from the SLR and transfer on the three research questions. Section 5 concludes the paper.

2 Definitions and Theoretical Framework

In the age of the Internet of Things (IoT), the digitalization of the physical world and its networking with the virtual world continues to increase. The term IoT was first mentioned in 1999 and was used for the networking of clearly identifiable objects using Radio Frequency Identification technology (Ashton, 2009). Today, the definition of IoT goes further. For example, the European Commission defines the future of the IoT in its roadmap as follows:

> *Things having identities and virtual personalities operating in smart spaces using intelligent interfaces to connect and communicate within social, environmental, and user contexts* (IEEE, 2015: 31).

Other definitions go one step further and add people and processes to networking in addition to things (Bardley et al., 2013). The networking of any objects within the framework of the IoT can be used in various areas. Especially, the mobile Internet and the associated increased use of devices and applications offer the possibility of technological reorientation (e.g., wearables) to the sports industry. The most important applications of IoT in the sports world are performance measurement and continuous monitoring. Until a few years ago, quantitative data used in football were hardly considered; if they were, the processes were evaluated qualitatively, mostly subjectively and emotionally. Due to the randomness and the supposed absence of patterns, no direct benefit of quantitative data was discernible in football, as in other sports, such as baseball (Quirling et al., 2017: 324).

A negative example for the quantitative evaluation of certain measured variables would be the 2014 World Cup semi-finals. Due to the "traditional" notation analysis, the Brazilian national team was superior to the German team in almost all categories (including passes into the penalty area, crosses, corners, and tackles won), but the match ended negatively for the Brazilian national team with 1:7 (Biermann, 2018a: 40). The result was due to the exclusion of tally lists in which actions are counted individually in the data analysis. Instead of a notepad and eyewitness reports, algorithms, and video technology should be used to collect and record game data. After all, victory and defeat are not only determined by the meticulous noting of procedures on the pitch but also concretely require sports analytics to predict a reliable correlation (Biermann, 2018b).

According to Memmert and Raabe (2017), both worlds—big data and IoT—combine to form "Game Analysis 4.0," "which uses sensors, applications, and technologies to collect and analyze data from physical objects or other sources. With the correct analysis, for example, basic tactical settings can be calculated from the player position data (Memmert & Raabe, 2017: 7ff.).

The rapid growth of the IoT was favored by numerous IT-supported devices using radio technology, which communicates with each other and generates large amount of data (Fasel & Meier, 2016: 5ff.). Sports companies are investing increasingly in big data in the form of cloud data storage, artificial intelligence (AI), and on-premise-solutions. The basic idea behind big data is the use and analysis of data generated in a wide variety of forms in everyday life (Marr, 2015: 9ff.). The term big data stands for the transition to a new model of dealing with data and the use of huge amounts of data, which are usually combined, made available, and evaluated within the scope of a second use (Weichert, 2013: 133). Nonspecific data are therefore condensed into specific information that is useful for the interest group, depending on the motivation of the respective interest groups in the data.

Moreover, big data is the term used to describe extensive and disproportionately growing data volumes and data stocks that have the following three main characteristics: a comprehensive data volume (or data volume) in the Tera to Zettabyte range (volume); various structured, semi-structured, and unstructured data types (variety); and a high processing speed of data streams (velocity) (Fasel & Meier, 2016: 375; Freiknecht & Rapp, 2018: 7ff).

Many "V" terms play an important role in many definitions of big data and thus seem to confirm Francis Galton's assumption that everything can be measured (Gould, 1996: 107). However, one problem in the past was the nonexistence of the development of useful applications for the evaluation of large amounts of data or databases. Applications for the use of big data are also used in sports, especially in the football industry. For the analysis and interpretation of the video and position data of the players, for example, new ways are needed. Thus, the professional field of the so-called game analysts who used big data and have the skills to evaluate it also emerged (Memmert & Raabe, 2017: 2). Through the IoT, recording important performance data of the players for football is already possible to create performance analyses by match analysts and to support the medical department, the fitness sector, or scouting (Gramlich, 2018: 13ff.). Especially, the stakeholders of a football company such as management, coaches, and players can optimize their strategies and tactics based on big data analyses.

The "moneyball principle" has given sports analytics a new dynamic. The expansion and shift of the sports industry toward an increasingly business- and IT-oriented approach led to increased demand. Sports analytics is now a fast-growing industry and is becoming an ever-greater necessity to survive in a competitive environment. Big data and cloud technology played a major role in this growth development (Gowda et al., 2017: 499). The first known use of the prioritization of statistics and data by Billy Bean is also referred to in the literature as Sabermetrics (developed by the statistician Bill James) and is still a central area of sports analytics today, like the search for patterns in sports-related data (Fried & Mumcu, 2017: 57). Sabermetrics (Society for American Baseball Research metrics) has emerged in the literature and is a synonym for the numerical approach to sport (Memmert & Raabe, 2017: 92). Even though no final definition of sports analytics has yet been established, the ultimate goal is to search for, collect, and process sports-related data. These are analyzed with the help of IT systems to realize the user's competitive advantages that are used in two directions: to increase the performance and efficiency of the teams and to improve marketing concerning economic advantages (Link, 2018a: 3). To achieve competitive advantages in the performance area, data-supported decision-making is used to provide action aids for trainers and management. However, both player recruitment and tactical game decision can benefit from it. In addition, key performance indicators (KPIs) are analyzed based on trends and patterns to identify the most relevant KPIs (Gerrard, 2016; Hughes & Barlett, 2002: 739f; O'Donoghue, 2005: 104ff.). The KPIs aim to break down a complex system behavior into individual values for scaling, rating, and ranking systems or system components (Perl & Memmert, 2017: 65). In addition to sports performance and economics, information technology can be regarded as the third topic complex in the field of sports analytics. Sensors generate much information, data volumes are linked via the Internet, and memory chips enable larger data transactions without time restrictions. In this way, data from various sources can be merged in real time and searched for anomalies using new methodological and conceptual approaches from data sciences (Link, 2018a: 3). However, information technology is undergoing a considerable upheaval: the interaction between technical and IT experts and the associated programming of complex systems is being abandoned in favor of machine learning and the development of cognitive systems. The breakthrough of machine learning and thus the paradigm shift was achieved only with the availability of large amounts of data. Today, task solutions can be implemented with the help of computers that did not seem possible with classical programming (BaFin, 2018: 24ff). AI develops very quickly through machine learning. One successful approach is Deep Learning (Lewanczik, 2018). On the basis of existing information and with the help of a neuronal network, the system can link what has been learned with new content repeatedly and thus learns again. In addition, the machine can make its own forecasts and decisions and in turn question them. Once decisions have been made, they are confirmed or changed during the next run (Zoph et al., 2017; Davenport, 2018: 1ff.).

The areas of application of AI in sport are currently still manageable, as the development of new methods for data analysis and the interpretation associated with it is proving to be difficult. In Germany, Fraunhofer IIS makes intensive use of the technical innovations, such as AI and machine learning, for sports and markets the results under the lemma "sports analytics." In the USA, a market worth billions has established itself for the collection, analysis, and marketing of sports data (Research & Markets, 2016). Moreover, a competition for the best data analytics tools has emerged between the major providers (IBM, Intel, Microsoft, SAP, and Google) and their products and services (Link, 2018b: 13).

The beginning of sports analytics in professional football lies in manually recorded and evaluated descriptive statistics, which decomposed football into its numerical components, such as ball contacts, pass quotas, and mileages. Ultimately, however, these were only analytical tools that could support an evaluation with arguments but excluded the possibility of forecasting. In particular, technological innovations enable improved performance measurement and thus analysis and evaluation of individual players or the entire team. Meanwhile, big data and its immediate availability enable the coaching team to make decisions and assumptions about the players and teams' performance potential (Castellano et al., 2014: 701ff.). Monitoring makes it possible to resolve disputed decisions into individual sequences. In addition, thermal images can visualize the sphere of influence or the playing strategies of individual athletes or the entire team and thus allow conclusions to be drawn about effectiveness and efficiency. Furthermore, gaps in defensive formations can be detected by computer-generated grids in still images of the live broadcast.

Real-time training analyses would also be conceivable. Wearable technologies such as Global Positioning System (GPS) integrated into clothing or fitness trackers use software applications to visualize movement and performance profiles. These enable information about the physical constitution of a player and his cognitive abilities, such as anticipation and overview. To conclude the effectiveness of tactical variants, player-tracking programs can be used for tablets and data glasses (Gehrmann, 2017). Wearables are small, lightweight technological devices that can appear in various designs, such as bracelets, belts, or rings. These devices are equipped with sensors that can record a certain type of activity (van der Westhuizen & van der Haar, 2018: 226ff.). The subsequent analysis of the data collected would be conducted either with the same device or with the help of appropriate software programs on a computer belonging to the infrastructure. However, not only the one-sided delivery of collected data via a computer network but also the feeding of data from other data sources (clouds) to the wearable carrier are conceivable. Clouds are the most important development platforms for big data analytics in the environment of 24/7 always-on operations, that is, the constant availability of services or the continuous operation of devices and machines (Kobielus, 2018). This means that huge amounts of data not only can be analyzed and applied in the shortest possible time but are also available everywhere (Pickup, 2018).

Another form of wearable technology is Intelligent Clothing, which can track every movement of a football player. Complete movement sequences including all movement data can be evaluated in real time. Intelligent Clothing assumes special functions for the player at those points where wearables cannot provide adequate data and functions. For example, "smart bandages" measure strain on the joints and make preventive recommendations based on these data. Moreover, activating muscles via low-current stimuli to improve training effects is possible with Intelligent Clothing (Meyer, 2017: 28).

The market for tracking players is highly competitive and diverse (Memmert & Raabe, 2018: 36ff.). Therefore, we only briefly described basic systems for position tracking based on the following: GPS, video, and radio and microwave (Baca, 2015). GPS is the buzzword commonly used for satellite positioning and navigation. The actual GPS receivers in sports are similar in functionality to modern smartphones. Each player is equipped with a transponder (radio communication device). The hardware records position data and fitness parameters such as the athlete's heart rate and respiratory rate. In addition, acceleration sensors, gyro instruments, and compasses can be used to record the players' movements. The subsequent evaluation of the data occurs either in real time or after the game, and it is finally processed during training. However, this technology has two limitations: (1) the need for outdoor use and (2) a tradeoff between the accuracy of measurements and evaluation of data. Video-based systems do not require any additional electronics on the player's body. These are based on the setting of different camera perspectives that can be generated by a large number of different video systems. High-resolution cameras, installed under the stadium roof, the hall roof, or on the training ground, for example, can take at least 25 frames per second. Relative to 22 single players, the amount of data can consist millions of single images. To date, video-based systems have been used in competitions without any problems (Memmert & Raabe, 2018: 38). The systems function rather semi-automatically; for example, if players are concealed during goal celebrations, assigning players is often impossible because their identification no longer functions unambiguously. A manual assignment for recovery is then required (Memmert & Raabe, 2018: 39). In radar- and microwave-based systems, the players carry small transponder units on the body. The positioning is conducted with the help of several receivers at the edge of the pitch of the stadium, the hall, or the training area. The system continuously provides real-time position data. In contrast to GPS tracking, however, the installation effort in the stadium, in the hall, or on the training ground is considerably higher, making it almost impossible to use the system for away games (Laukenmann, 2017).

Two further technologies, namely, augmented reality (AR) and virtual reality (VR), have opened up new possibilities recently. AR is about augmenting the reality and the real environment with additional virtual information, thereby possibly improving the decision-making skills of users. Meanwhile, AR mixes virtual characters with the real world (Azuma, 1997: 2). Different human sensory modalities are addressed by a computer-aided extension of the perception of reality with additional information in different forms (videos, text, graphics, and 3D objects). AR offers the possibility of the user's perception of the virtual elements in his real environment. By contrast, in VR, the user identifies himself completely with a 360° world. VR creates the feeling of being at the scene of the action, whereas AR requires the user to be present to receive further information. The world of VR is supported by VR glasses or cardboard, whereas AR uses smartphones, tablets, or augmented reality glasses (Kipper & Rampolla, 2013: 1; Mangold, 2017).

AR and VR are used in professional football. For example, in the Spanish football league, viewers have access to 360° slow-motion and AR statistics. Moreover, in the National Basketball Association and National Football League, these technologies are standard and enable the user to immerse in other worlds (Zobel et al., 2018: 123ff.).

3 Methodology

Within the framework of the SLR, we applied the procedure of Massaro et al. (2016). The keywords used were "big data" along with "sports," "football," "analytics," and "match analysis." These terms were also searched for their German equivalents. Three databases were initially used to cover the relevant literature. By entering these keywords in (1) Google Scholar, (2) Scopus, and (3) Business Source Premier (EBSCO) (3), we identified the relevant literature.

Whether and to what extent the identified sources fulfilled the following criteria were also examined:

- The work is intended to shed light on the digitization based on measures relating to certain stakeholders and categories in the football field, which is why only those sources that contain explanations on these instruments are used for the SLR.
- The sources can be all types of publications (books, journal articles, research papers, etc.), provided they are officially or scientifically recognized documents.
- The sources must be published in English or German to analyze and compare the results.

Based on these exclusion/inclusion criteria, 102 literature sources in the period 2002 and 2018 were used for the SLR.

Accordingly, a suitable research framework that sets out the various interests and goals of the stakeholders in the digitization process must be developed. The purpose of this research framework is to find an adequate approach for the SLR to analyze and aggregate the relevant information from the selected articles (Vanini & Rieg, 2017: 9). The research framework contains bibliographic information, characteristics of the research design, and first results of the sources used for the SLR (Cooper, 2010: 45; Tranfield et al., 2003: 214).

Our applied research framework includes eight categories (see Table 1). The first category (1.1) organizes the sources used for the SLR according to literary attributes such as the number of sources used, the number of authors, and the period of observation. A constantly increasing development of publications in this research field can be observed. This can of course be explained by the increased processing of the key terms in media and society. The second category (1.2) lists the locations where the authors are predominantly active or conducting research. For a detailed overview, the third category (2.1) classifies the sources according to their type. The fourth category (2.2) shows how many sources are in German or English. In addition,

Table 1 Research design and descriptive (see Vanini & Rieg, 2017: 12)

Category	Variable	Results
1 Bibliographic information		
1.1 Sources, authors, research period	Sources Authors Research period	102 251 2002–2018
1.2 Authors' regions	Africa Asia Australia Europe North America (Canada/USA) South America	3 10 13 163 61 1
2 Research framework		
2.1 Type of sources	Books Book chapter Research papers Journal articles Conference papers	16 13 5 62 6
2.2 Language	German English	13 89
3 Results		
3.1 Stakeholder	Athlete Trainer Medical support Scouting Management	12 25 22 31 23
3.2 Interests of stake holder (in %)	Performance diagnostics, input for training design, and strategy development Injury Prophylaxis, rehabilitation Player rating Contract design, marketing	44% 40% 40% 33%
3.3 Sports analytics	Technique Economy Performance	54 22 46
3.4 Notes on strengths/weaknesses	Yes No	63 39

the fifth category (3.1) shows on which stakeholder groups the explanations of the respective sources are based. The sixth category (3.2) presents the metrics of the sources used that appeared most frequently. The interests and aims of the stakeholders were extracted and aggregated from all sources used for the SLR. Moreover, the seventh category (3.3) refers to the topic complexes of sports analytics. Finally, the eighth category (3.4) shows how many sources considered also discuss the strengths and weaknesses of the categories.

In addition, the citations were analyzed to identify the researchers and sources with the greatest influence. To ensure the validity of the sources, the citations were counted using Google Scholar (Dumay, 2014: 5).

The total number of citations is an indicator of the validity and quality of a source. The higher the number of citations received, the more meaningful or qualitative the source. However, the differences are large as there are seven sources with zero citations received and 26 sources with more than 100 citations received. The large differences can also be seen in the standard deviation. On average, a source receives about 117 citations (see Table 2).

Moreover, because older sources may receive more citations due to their "head start in time," citations per year were used. Table 3 shows the top three sources with the most total citations and the sources with the most citations per year.

The contributions of Davenport and Harris (2007), Di Salvo et al. (2007), and Lewis (2004) can be seen to play a significant role in the framework of our study.

The 102 sources used for the SLR were written by 251 authors. Only 26 sources were written by a single person, whereas most sources were written by two or more authors. Since only 26 authors have written more than one article, book, etc., the defined research field can be characterized as quite fragmented. Moreover, this probably does not lead to the expectation of a "dominant" opinion in the research area.

Table 2 Citation analysis(see Vanini & Rieg, 2017: 13f)

	Min	Max	Average	Median	Standard deviation
Total number of citations	0	1.627	116.73	27	252.49
Citations per year	0	197	18.77	8.07	29.21
Number of authors	1	18	3.24	2	2.67
Number of pages of the sources	1	320	44.67	13	79.34
Number of sources	0	156	40.32	33.5	33.15

Table 3 Citation analysis(see Vanini & Rieg, 2017: 15)

Rank	Source	TNC	Rank	Source	CPY
1	Davenport and Harris (2007)	1.627	1	Gabbett (2016)	197.00
2	Lewis (2004)	1.275	2	Davenport and Harris (2007)	147.91
3	Di Salvo et al. (2007)	1.004	3	Di Salvo et al. (2007)	91.07

The number of pages can also be an indicator of the quality of the sources, as longer articles, books, etc. usually suggest more comprehensive and detailed analyses. The results show that the average of about 45 pages does indicate longer articles, but this depends on the type of source, as not only journal articles but also more extensive books and book sections were used. Therefore, the median should also be considered at this point. Thus, of 102 sources, 37 sources have a page count of less than or equal to 10 pages.

The number of source citations is an indicator of the level of detail and thus the informative value of a source. Here, the range is wide with a minimum of 0 and a maximum of 156. However, this depends on the type of source and its number of pages. Books, for example, have more source citations due to their length.

Our SLR also has some limitations: This concerns the selection and analysis of sources, which is ultimately based on subjective elements despite the systematic and intersubjectively comprehensible procedure. In addition, only German and English language sources were used for the SLR. Studies from other language areas were not considered, which is tantamount to not having a fully comprehensive literature evaluation on this topic, albeit an extensive one.

4 Results and Discussion

Based on the SLR, an increasingly important role that big data play in professional football, one of the most common sports, can be confirmed. This role enables the consolidation of information and knowledge and thus the analysis and interpretation of correlations and patterns. In doing so, the sports uses digital technologies to improve performance in a competitive environment. The proliferation of tracking technology in sports and the associated desire to record and monitor the activities of athletes have led to a significant increase in the volume of data, the variety of data in circulation, and the rapidity of data transmission. (Millington & Millington, 2015: 140).

The use of connected electronic performance and tracking systems components and data analytics is now commonplace in almost every sport. The interplay of innovations in hardware and software is shaping the history of data analytics in sports (Memmert & Raabe, 2018: 26). However, the data obtained are worthless if not accompanied by analysis. Statisticians and analysts are playing an increasingly important role in human resources in the sports industry. However, interest in sports analytics is not yet compelling among many stakeholders in professional sports, even though various data, technologies, new metrics, and analytics exist on the market (Davenport, 2014: 2). One major reason cited for this is that many professional sports teams cannot afford to employ multiple analysts with their specialized skills due to their own economic realities (Davenport, 2014: 13).

In the field of professional football, the situation is different, probably also because of the economic size of this industry. FC Barcelona, Manchester City, FC Liverpool, and FC Bayern are now conducting intensive research into digital match analysis, mostly in a highly secretive manner. They employ renowned scientists or hold seminars on data analysis and AI with experts (Biermann, 2018: 42). Moreover, clubs, such as Chelsea FC and Manchester United, employ data scientists to meticulously track and record every move of players to maximize their chances of winning games (Marr, 2015: 213).

In the majority, however, the performance analysts do not interact directly with individual players (Carling et al., 2018). Rather, it seems to be about a knowledge edge in holistic digital match analysis. The German Football Association (DFB), for example, is planning cutting-edge research on this topic in Frankfurt (Saam, 2017). It's not surprising, then, that the analytics team of the German national team has embarked on a big data project with software company SAP, focusing on how to control space on the pitch and, by extension, the opposition (Biermann, 2018: 42).

For example, FC Barcelona uses an algorithm developed by Daniel Link in Germany. With the help of a large amount of game data, this allows production of statements about how great the goal threat is when the team is on the ball. At any time, the analyst can follow the game and determine how well a team is attacking or the opponent is defending. Link's (2018a: 29ff) concept of "Dangerousity" is used extensively by many professional stakeholders in professional football, and based on our SLR, it appears to be a leading product on the market.

Meanwhile, Benfica Lisbon, a Portuguese football club, is considered a pioneer of professional sports analytics and has become known for its intelligent, data-driven transfer policy (Craig, 2018). The club thus controls 100% of the development of its top performers and subsequently acts as a feeder club for even larger football clubs. Within 6 years, Benfica Lisbon sold 13 top performing players for 270 million UK pounds. By using machine learning and predictive analytics, players can use their personal data and the insights gained to optimize their performance and continuously improve. The basis of the success is mainly based on the fact that the young players can develop without major injuries (Wired, 2018).

The Benifca Lisbon coaching staff can do little to prevent injuries caused by direct contact during a game, that is, traumatic injuries. However, in the case of non-traumatic soft tissue injuries, especially muscle and ligament injuries, big data provide an opportunity to predict and, in the best case, even prevent them altogether. The Benfica Lisbon medical staff supports the coaching team in analyzing the physical stress of athletes with an accurate, predictive injury model. This model provides information about the peak loads an athlete can be exposed to before an injury occurs. The main questions in this model are the kind of injuries that can occur in certain risk situations and the convalescence period in case of an eventual occurrence.

FC Midtjylland has taken a similar but more rigorous approach to Benfica Lisbon (Memmert & Raabe, 2017: 120ff). This club from the 1st Danish Football League is pursuing the objective of eliminating the factor of chance as completely as possible. FC Midtjylland not only wants to be not only a professional football club but also a laboratory for innovative experiments. At the center of all decisions, whether scouting or evaluating games, is analytics. Scientific methods are used to replace irrational, subjective, and emotional decisions. This experimental approach shows that a football club can be successfully managed based on statistical analysis (De Hoog, 2015).

Creating a talent pool, for example, is driven by data and information. The use of scouts for talent management still consists only of recognizing whether the players sighted are fit into the team from a personal and psychological perspective. At this point, at least, efficiency and success are still influenced not only by data but also by the "synthesis of cold analysis and heart" (Thite, 2018: 82). Meanwhile, in the evaluation of games, mathematical models are used almost exclusively (Biermann, 2016). Large football clubs in particular have excellent analysts and scientists with outstanding ideas and innovation potential in the field of data and sports analytics. However, they have no final decision-making impact on strategic decisions; these are ultimately made by the top management. For instance, a personal union exists in Midtjylland: the revolution comes from top management because the chairperson Rasmus Ankersen is himself an analyst (De Hoog, 2015). Therefore, the clear mantra from top management is "we are a club without ears and eyes" (Memmert & Raabe, 2017: 122).

In 1st German Bundesliga, TSG 1899 Hoffenheim, in particular, is often cited as an example of the consistent use of big data to optimize performance. The data from the game and player analysis are used for systematic performance diagnostics. In addition to the athletic performance factors, the psychological performance factors are identified, stored, and used to make statements for training control, individualization, and so on (Görlich & Mayer, 2018: 22).

Moreover, other new technologies are being integrated into the training process to support players' conscious thinking with the help of video games (e.g., software called "Helix") (Söhnlein & Borgmann, 2018: 23ff). The athletes are confronted with VR situations in which their own team and the opposing team meet. It is hoped that this technology will improve the athletes' positional awareness. Another advantage of this software is that the exercises can be adapted and tailored to the positional requirements of different players. For the intuitive aspects of the players, the "Footbonaut" (Fiedler, 2018: 59ff) is additionally used to test the players' reactions to unpredictable technical contexts. Reaction times or technical accuracy are registered to generate a profile of the player and monitor his development. The combination of the conscious thinking approach with tactical and intuitive aspects potentially brings an improvement in a player's decision-making as he perceives game situations more consciously.

Finally, the use of the "SAP Sports One" product as a common platform for the mobile provision of medical, psychological, and performance diagnostic parameters and other live information is an essential component of the football club's big data strategy. The derivation of data from the system for new or optimized applications is available to various team-related groups at any time (Görlich & Mayer, 2018: 22).

The case studies of Benfica Lisbon, FC Midtjylland, and TSG 1899 Hoffenheim, which are presented in the structured literature analysis, show how big data can help football clubs to achieve a competitive advantage in the context of a more efficient usage of their human capital (player squad). The evaluation shows that professional sports analytics especially fixates on human resources, because only selecting and owning the best players at the lowest price makes the difference and represents the most important competitive advantage (Davenport & Harris, 2007: 37). However, a spiral of innovative competition has been initiated among the football clubs, which seems to be spinning much faster and is forcing the clubs to permanently expand their analytical capabilities and to tie up considerable financial resources in the process. (Davenport & Harris, 2007: 67ff).

Results from the SLR reveal the seemingly important role played by big data for the trainer in the field of performance analysis/diagnostics in the future, as many performance indicators are made measurable through data collection and analysis methods. In this context, an important future topic for coaching staff will be prevention, the objective of which is sustainable health promotion with measures to prevent injuries and illnesses. Current trends in the medical field include regeneration, prevention, and individual training (SportHeads, 2018: 33).

The number of games and training sessions and the enormous pressure to perform placed on athletes during a season can have far-reaching effects on the athletic success or failure of the individual athlete and, consequently, on the entire team. If the training load is an important determinant of injury, it must be measured accurately daily over weeks and months (Gabbett, 2016: 273). The goal is clearly the critical analysis and evaluation of a player's physical activities and the associated creation of an optimal training plan (Davenport, 2014: 22). In addition, the medical staff can use the training data to minimize the risk of injury from overuse through preventive treatment (Link, 2018a: 17). The key here is targeted load control of individual players and a possible early warning system for signs of fatigue (Memmert & Raabe, 2018: 72f). McCall et al. (2015) suggested that fatigue is one of the most important factors in football injuries.

The number of kilometers ran, the sprints performed, and the average daily speed achieved by a player are information that is not decisive for the pure medical care of professional football players. Rather, it is a matter of collecting data that can be used to identify specific stress control measures for individual athletes and act as an early warning system for signs of fatigue (Memmert & Raabe, 2017: 134). However, the biggest problems are the indicators themselves: How is fatigue measured in concrete terms? How quickly and when can a player be reintegrated into the daily routine after an injury? The answers cannot yet be answered unambiguously by the scientific community (Memmert & Raabe, 2018: 73). Currently, scientifically based models in practice that enable sensible load control with suitable specifications is still lacking (Memmert & Raabe, 2017: 135). That said, many football organizations today are using existing technology in their training sessions to minimize the risk of injury and prevent long-term negative health effects (van der Westhuizen & van der Haar, 2018: 226ff; Ikram et al., 2015: 1). This was also the main reason why the International Football Association Board allowed wearables in the game, to be able to detect possible health problems at an early stage with targeted stress monitoring (Memmert & Raabe, 2017: 134f).

Another major approach being pursued with big data is the improvement of players' cognitive abilities. In football, considerable importance is attached to understanding the game and decision-making ability, not only for the outcome of individual matches but also, as a consequence, for the position in the table at the end of a season (Frick, 2004: 71). The goal is to develop the players' procedural knowledge and to increase their decision-making speed in game situations. In a football game, the context is constantly changing as players must make decisions based on the position of the ball, the current position of teammates, opponents, their own position on the field, and the estimated distance to the different participants. Because of this variability, each player interprets game situations according to his own experience and preparation for these moments. This effective decision-making is called tactical skill and is a basic requirement for a good football player. In particular, this skill of the experienced player refers to deciding which action to perform at a particular moment of the game. Thus, it is an important feature for performance differentiation and simultaneously serves to distinguish it from purely physical characteristics (Kannekens et al., 2011: 846f).

Today's modern game of football also demands an increase in the speed of the decision-making process among players (Bush et al., 2015: 1ff). An analysis of the differences between the 1966 and 2014 World Cup showed that players were 35% better at targeted passes and speed in 2014 than in 1966. The contact time for a player on the German national team changed from 2.9 s at the 2006 World Cup to 0.9 s at the 2014 World Cup (Mckenna, 2017). However, the fast execution of actions does not mean that the player must run faster than his opponent. Instead, it means that he can be better positioned on the field or play better passes. Thus, the player's tactical skills increase the likelihood of reading the opponent's play and effectively making game decisions based on it (Gerrard, 2016: 214).

Football is highly complex, not least because of the number of players and the size of the field; therefore, the coaching staff must be able to reduce the unpredictability of a game by simulating game situations. This should optimize players' decision-making processes under simulated competition conditions (Di Salvo et al., 2007: 222ff). Today, software solutions can assess the temporal-spatial behavior of players and the movement patterns of teams in terms of their effectiveness and efficiency.

To better prepare for the next game, coaching staff must know exactly which tactical variants should be used depending on the respective opponent. The basis for this sports game analysis is the movements of the athletes, which are recorded by several cameras or body sensors. This enables a reconstruction of the movement sequences. Thus, coaching staff can make an evaluation by visual information. For example, the trainer could suggest a different field position for each player or analyze whether the distances between the players correspond to the strategies defined in training. Even the concrete determination of distances between players can improve the team's positioning in the long term and thus increase performance. It also allows for constructive engagement with players to establish structural patterns of play. Overall, data analysis via games provides the coaching team with the legitimacy of objectified feedback to the players to standardize team strategy (Gudmundsson & Wolle, 2013).

Many coaches see great potential in the cognitive optimization of players. The focus is on players' faster thinking, recognition, anticipation, and decision-making. However, scientific research in the cognitive field is still in its infancy. In the future, imaging techniques such as functional magnetic resonance imaging or electroencephalography will be increasingly used in mobile applications. In addition, an increased individualization in recovery and stress control is expected (SportHeads, 2018: 21).

Based on our SLR, professional sports clubs show a strong tendency to use big data and its analysis to try minimizing the chance of losing in their respective sports. In football, for example, coaching is likely to become even more technology- and computer-assisted in the future. Football analysts will in all likelihood use big data to elicit patterns and to develop new individual algorithms for training control and match analysis. However, the networking of science and sport is only just the beginning (SportHeads, 2018: 21). With the help of big data and its evaluation, the coaching staff can analyze the opponent, even determine the individual game behavior, possibly also the respective game intelligence. Important impulses can be given to determine the optimal team formation in the context of the respective opponent and its strengths and weaknesses. Performance and fitness data provide coaches, physiotherapists, physicians, and psychologists with information for the individual design of training sessions or rehabilitation measures for each player and thus connect to the aforementioned anticipatory measures against signs of fatigue and corresponding injuries (Gehrmann, 2017).

The tactical analysis will play an even greater role in football in the future because continuous-time position data are particularly available in large quantities. Three fundamentally different approaches can be distinguished: trivial methods, system models, and metamodels (Link, 2018a: 19): Trivial methods are cumulative statistics based on player results or measures of running intensity. However, since football is a complex system due to its dynamic, non-linear nature, these data are, in principle, not meaningful enough for strategy development or performance evaluation (Carling et al., 2014: 2ff; Mackenzie & Cushion, 2013: 639ff). System modeling tries interpreting the interaction and coupling behavior of teams, groups of players, and individual players with the help of relative phase or entropy. This should help make statements about the system dynamics a football game is subject to (Rein & Memmert, 2016; Frencken et al., 2012: 1207ff). The network analysis for passing behavior in games and thus for the identification of recurring patterns also provide promising results and belongs to the type of system modeling (Ribeiro et al., 2017: 1689ff). Another possible approach is based on a metamodel that recognizes tactical constructs according to position data (Beetz et al., 2009: 1ff). Moreover, data on ball possession (team ball possession, individual ball possession, or ball control), actions (pass, cross, or one-on-one) and tactics (pressing, synchronization, or playability) are recorded electronically based on raw data such as player position, the ball position, and game status. This allows strategies or playing styles to be defined for various situations (Link, 2018a: 19).

The extraction of tactical structures from raw data is done either by conventional rule-based approaches or by machine learning. When using explicit rules, the knowledge is usually modeled by human experts. However, in machine learning, there are no predefined rules, but they are derived from existing datasets. With the help of artificial neural networks and supervised learning, data analysts determine a so-called ground truth to subsequently use this as a training basis for the network. In unsupervised learning, the ground truth is omitted by the data analysts; that is, the network does not know what it is supposed to recognize. In both cases, algorithms do not need to be defined fixedly because they emerge during the learning process. Therefore, machine learning is mainly about pattern recognition and relationship detection (Kempe et al., 2015: 249ff; Rein & Memmert, 2016).

With conventional models, the results are better if a detailed design is available. However, if the design is too complex, a multiplication of rules and a high error rate can occur. Machine methods are excluded from this problem (Link, 2018a: 20). Supervised learning requires the correct answers to the model constellations. These must be supplied as a unique ground truth. The specification of the ground truth is often associated with a high manual effort in the preparation of the data. Objects cannot always be unambiguously classified, and values cannot always be estimated or predicted. In football, for example, the attacking play of one's own team should be structured in such a way that the constant playability of the athletes is possible. However, playability cannot be fully objectified. Ground truth is lacking, and thus, the net lacks the ability to learn. In unsupervised learning, the advantage is the partially fully automated creation of models. By training the models, the input data are increasingly adapted to the models. The basic challenge in unsupervised learning is to recognize when a model is sufficiently trained based on patterns and relationships. This is the involved verification process (Fraunhofer, 2018: 23ff).

For the 2018 FIFA World Cup, methods based on machine learning were used for the first time. Hence, the so-called random forest method was used for forecasting purposes (Groll et al., 2018: 7ff). A random algorithm gains access to parts of a data set and creates thousands of possible decision trees. This can then be used, for example, to determine expected values for goals. It could be proven that the amount of data plays a decisive role. The larger the number of games to be evaluated, the smaller the percentage deviation between the prediction and the actual result (Schauberger & Groll, 2018).

Predictive analytics is another method from the field of data mining that is used in sports and especially in football after the evaluation of our SLR for data analysis. Here, a statistical or machine learning technique is used to create a quantitative prediction of the future. Often, predictive analytics occurs along with supervised machine learning to predict a probability (e.g., how likely is this player to delay a shot on goal?) or a future value (e.g., how long can a player be used before needing a recovery period?) (Burns, 2020). Although machine learning techniques are improving forecasts, even in the unlikely event that a model could take all relevant variables into account, the emergence of "predictable" football is unlikely and thus no certainty about the actual match outcome. With big data as a basis, almost everything in professional football can be measured: ball possession, passing rates, or running distances of individual players. For a long time, data such as ball possession and running performance were considered factors in winning a game. However, successful teams like FC Barcelona or Bayern Munich often have lower running performance and are still successful (Memmert & Raabe, 2018,: 256).

The collection of biometric player data is highly visible as a useful information tool for training. According to the literature, intelligent technologies offer players great opportunities, but also considerable risks. The potential benefits lie in the optimization of players' physical performance. As a result, many professional sports clubs have become "laboratories" dedicated to injury prevention and performance enhancement (King & Robeson, 2007, 2013: 13f). Football companies in particular monitor their players in private by observing their diet and sleep patterns. In addition, emotional well-being is recorded via social networks (Marr, 2015). The interest in comprehensive information about the health and performance of players on the part of club managers is understandable, especially since a player's absence represents a considerable cost factor. However, the boundary of the relationship between employee and employer in terms of privacy and monitoring can quickly be crossed. Therefore, biometric data also raise questions about players' rights and data protection, as it carries the risk of compromising players' privacy and autonomy by no longer guaranteeing the confidentiality of the data. Moreover, this may negatively influence contract negotiations with the consequence of career damage or curtailment (Karkazis & Fishman, 2017: 45–50).

A digital divide can be seen between athletes who have access to their own biometric data and the knowledge of how to use them and those who have neither access nor the necessary knowledge of what to do with the provided data. For athletes, the logical question is whether this digital divide can be bridged and a degree of data sovereignty can be ensured (Baerg, 2017: 9). With the help of the so-called "quantified self-movement," self-tracking can be performed, and the players can generate their own analysis (Nafus & Sherman, 2014: 1785). At first glance, digital self-measurement and data analysis in training work can be viewed positively in terms of data sovereignty for the players, but in practice, this approach raises the question of the effectiveness of data collection and data evaluation.

Another risk from player's perspective is that other parties (e.g., advertising partners and third-party ownership investors[1]) are interested in the data of individual players and may acquiring them for sale. Third-party use of biometric data raises ethical concerns: a) validity and interpretation of data; b) increased surveillance and privacy threats; c) confidentiality risks and data security concerns; d) conflicts of interest; and e) resulting dependence and coercion (Karkazis & Fishman, 2017: 48). On a positive note, all players' data that are measured and recognized can be presented with complete transparency (Memmert & Raabe, 2018: 147). However, from a negative viewpoint, there is a risk of the "glass player." A buyer of these detailed statistics could not only develop strategies to exploit the weaknesses but also impart insider market knowledge when buying and selling (Marr, 2015). In addition, data analytics procedures can be used to de-anonymize data once it has been anonymized; that is, "harmless" data are eliminated. Personal data can be used for correlations that may reveal, for example, political orientation or sexual orientation of a player (Thür, 2015: 5). This would then have a concrete impact on the player's relevant advertising marketing. Apart from the sporting attributes of the athlete, this factor also has a major influence on the evaluation of a player.

[1] For further information about third-party ownership arrangements see Herberger et al. (2018) and Herberger et al. (2019).

One potential drawback of data analytics in sports is the ever-increasing number of data sources, and consequently, coaches and management must handle and process this large amount of information. Almost all factors can be collected to measure performance; thus, the challenge is to distill the most relevant information from the data. Doing so creates the risk of digging too deep into the data, thereby increasing the problem of incorrect conclusions based on statistically significant results that lack causality. AI and machine learning can assist in this problem by developing algorithms that independently support and significantly improve the decision-making process; however, the fundamental problem cannot be eliminated (SportTechie, 2018). In individual cases, this can also lead to incorrect conclusions by decision-makers.

The evaluation of the SLR shows that the influence of chance will be further minimized in the coming years. The supervision, training, and coaching of players will be supported even more by technology and computers. Analysts will increasingly try exploiting opportunities in the field of big data by searching for patterns and creating individual algorithms for training control and game analysis. However, the potential from the networking of science and sport has not yet been fully developed (SportHeads, 2018: 21).

Sports data analysts must also be able to effectively communicate discoveries to decision-makers, typically various stakeholders within the professional football club. Currently, this communication can be a bottleneck, as executives often lack the ability and familiarity to translate quantitative data into qualitative thinking. Consequently, the demand from decision-makers within the professional football club for sports analytics is significantly less than the potential supply of data, technology, new metrics, and analytics (Davenport, 2014: 2; Biermann, 2016).

Professional football is undergoing sustainable and strong changes, mainly due to the influence of big data. Moreover, professional football clubs are increasingly investing in deep learning, predictive analytics, and the entire sports analytics sector. Given the large investments in transfer rights trading, foregoing sports analytics in player signings to reduce uncertainty would also be negligent.

However, investments in big data may not have the positive effects hoped for if professional football clubs are unable to make adequate decisions with the information they obtain from data analytics. The purchase of analysis tools does not automatically develop the competence to produce meaningful arithmetically developed results and especially to draw the appropriate conclusions from them. As long as a professional football club has not learned to use data and analytics to support its decisions, it will not be able to benefit from big data and sports analytics (Ross et al., 2013).

5 Conclusion

This study aims to show the changes caused by the use of technologies in the context of big data and sports analytics based on an SLR concretely in football. We also aim to analyze in this context to what extent the use of big data and sports analytics changes the strategies of professional football clubs and their stakeholders.

Based on our SLR, for the first research question, the emergence of digital methods and technologies in football is apparent. Modern computer-aided systems can be used to record individual players or the entire team to analyze tactical behavior. Moreover, kinematic characteristics, such as running paths, can be determined. The use of automatic tracking methods leads to a considerable reduction of the acquisition effort. Since modeling the course of a game proves to be complex, promising methods of data mining, such as the random forest method or predictive analytics, are conducted with the help of machine learning to support strategic planning. In practice, data mining has only been used in a few areas so far, as the trainers ultimately want to make the decisions themselves (Schoop & Brauchle, 2016: 20).

In the future, data analysts will become indispensable consultants. Big data and digital analysis tools will support the stakeholders of a professional football club. Moreover, software will help merge an increasing amount of measured and predicted data. The analysis of every game movement will serve as a primary factor for evaluating the game and recognizing the opponent's strategy. At the same time, it will serve as a basis for making informed decisions and calculating chances of winning.

For the player as human capital, digitalization means that the head coach can use objective data to justify his decisions in the debriefing of a match, resulting in better traceability and clear error analysis of the player. This transparency provides the club with the so-called "transparent player." However, the ethical component must not be forgotten. Many active football players are now asking about the whereabouts of the data collected. In particular, attention should be paid to the global scouting companies, which monitor and measure players at a young age to provide current market values to potential buyers.

Concerning the second research question on the benefits of big data and sports analytics for increasing competitiveness, the main purpose is to transform raw data into meaningful, value-adding, and actionable information. Data analysts make their strategic decisions based on these data, which leads to improved performance and generates a measurable and sustainable competitive sports advantage, ultimately resulting in a concurrent economic competitive advantage (Bukstein, 2016: 25ff).

The quantifiable performance indicators controlled by advanced tracking tools allow coaches to try achieving optimal load and maximum performance of their players. Big data and sports analytics are a great help to develop adequate strategies based on interdisciplinary synergies and maximize value creation from the potential of coaches and players. However, results on the field are not only determined by predictable facts.

Due to technological growth, real-time data can now be extracted and made immediately available to trainers on smart devices. This enables decisions to be made in a fraction of a second (van der Westhuizen & van der Haar, 2018: 226ff). To ensure this, coaches are increasingly reliant on data scientists and their analytics. Therefore, big data and sports analytics can be used to effectively improve performance in training and competition.

Our third research question addressed the chances and risks to be considered in using big data and sports analytics for the different stakeholders in a professional football club. Big data applications allow supporting, accelerating, and critically evaluating the decision-making ability of players based on predefined game situations or performance indicators. In addition, sports-specific software will be used to improve players' perceptual abilities and thus train conscious and unconscious thinking, which can lead to competitive advantages and should therefore be viewed positively from the perspective of stakeholders within a professional football club. However, the collection and analysis of a large amount of highly sensitive personal data can lead to transparency but has negative effects, especially for the player himself.

The football of the future is also facing a digital transformation, but this does not mean that big data and sports analytics will decode it tomorrow. In the next decade, football will be more modern and advanced from an analytical perspective. However, how will players, coaching and medical staff, management, and fans react to the different methods, approaches, analytics, and other innovations? Is it even possible to finally analyze the interpersonal factors in such a complex game like football?

Sports analytics will most likely change football on the field, but chance will continue to be a determining factor (Lucey et al., 2013: 2706ff). With sports analytics, a wide variety of data types are generated during data storage and use. An integrated database or data warehouse should be available for this purpose. However, not only the physical integration of data but also stronger cross-functional coordination of analysts must be present, for example, between performance analysts (qualitative video analysts and quantitative data analysts) and sports scientists (Gerrard, 2016: 216).

Acknowledgments We do not receive any grants or funds for our paper. All remaining errors are our own.

References

Ashton, K. (2009). That 'Internet of Things' Thing: In the real world, things matter more than ideas. Retrieved February 28, 2021, http://www.rfidjournal.com/articles/view?4986

Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, *6*(4), 355–385.

Baca, A. (2015). *Computer science in sport. Research and practice*. Routledge.

Baerg, A. (2017). Big data, sport, and the digital divide: Theorizing how athletes might respond to big data monitoring. *Journal of Sport and Social Issues, 41*(1), 3–20.

BaFin. (2018). Big Data trifft auf Künstliche Intelligenz. Bonn, Frankfurt am Main: Bundesanstalt für Finanzdienstleistungsaufsicht.

Bardley, J., Reberger, C., Dixit, A., & Gubta, V. (2013). Internet of everything: A $4.6 trillion public-sector opportunity: More relevant, valuable connections will boost productivity, revenue, and citizen experience, while cutting costs. Retrieved February 28, 2021, https://www.cisco.com/c/dam/en_us/about/business-insights/docs/ioe-public-sector-vas-white-paper.pdf

Beetz, M., Hoyningen-Huene, N., Kirchlechner, B., Gedikli, S., Siles, F., Durus, M., & Lames, M. (2009). ASPOGAMO: Automated sports game analysis models. *International Journal of Computer Science in Sport, 8*(1), 1–21.

Biermann, C. (2016). Moneyball im Niemandsland. *11 Freunde*. Retrieved February 28, 2021, https://www.11freunde.de/artikel/midtjyllands-revolution

Biermann, C. (2018a). Ohne Mathe geht es nicht. Wie die digitale Datenanalyse den Fußball verändert. *Rotary Magazin*, *2018*(6).

Biermann, C. (2018b). Matchplan: Die Neue Fußballmatrix. Köln: Kiepenheuer & Witch.

Bukstein, S. (2016). Evolution and impact of business analytics in sport. *Sport Business Analytics*, 25–46.

Burns, E. (2020). Predictive analytics. Retrieved February 02, 2021, https://www.searchenterprisesoftware.de/definition/Predictive-Analytics

Bush, M., Barnes, C., Archer, D. T., Hogg, B., & Bradley, P. S. (2015). Evolution of match performance parameters for various playing positions in the English Premier League. *Human Movement Science, 39*, 1–11.

Carling, C., Wells, S., & Lawlor, J. (2018). Performance analysis in the professional football club environment. In W. Gregson & Littlewood (Eds.), *Science in soccer: Translating theory into practice*. Bloomsbury Publishing PLC.

Carling, C., Wright, C., Nelson, L. J., & Bradley, P. S. (2014). Comment on performance analysis in football: A critical review and implications for future research. *Journal of Sports Sciences, 32*(1), 2–7.

Castellano, J., Alvarez-Pastor, D., & Bradley, P. S. (2014). Evaluation of research using computerised tracking systems (Amisco® and Prozone®) to analyse physical performance in elite soccer: A systematic review. *Sports Medicine, 44*(5), 701–712.

Cooper, H. M. (2010). *Research synthesis and meta-analysis: A step-by-step approach*. SAGE.

Craig, L. (2018). Sports analytics: How data gives teams the edge. Retrieved February 02, 2021, https://www.techerati.com/features-hub/opinions/sports-analytics-how-data-gives-teams-the-edge/

Davenport, T. H. (2018). *The AI advantage. How to put artificial intelligence revolution to work*. MIT Press.

Davenport, T. H. (2014). Analytics in sports: The new science of winning 2. International Institute for Analytics, 1–28.

Davenport, T. H., & Harris, J. G. (2007). Competing on Analytics: The New Science of Winning. Harvard Business Press.

De Hoog, M. (2015). How data, not people, call the shots in Denmark. Retrieved February 02, 2021, https://thecorrespondent.com/2607/how-data-not-people-call-the-shots-in-denmark/230219386155-d2948861

Di Salvo, V., Baron, R., Tschan, H., Calderon Montero, F. J., Bachl, N., & Pigozzi, F. (2007). Performance characteristics according to playing position in elite soccer. *International Journal of Sports Medicine, 28*(3), 222–227.

Dumay, J. (2014). 15 years of the Journal of Intellectual Capital and counting–A manifesto for transformational IC research. *Journal of Intellectual Capital, 15*(1), 2–37.

Fasel, D., & Meier, A. (2016). Was versteht man unter Big Data und NoSQL? In D. Fasel & A. Meier (Eds.), Big Data. Grundlagen, Systeme und Nutzungspotenziale. Springer Fachmedien.

Fiedler, H. (2018). Der Einsatz von KI in der Personalauswahl - Erkenntnisse aus dem "Footbonaut". In R. Lanwehr & J. Mayer (Eds.), People Analytics im Profifußball. Springer Fachmedien GmbH.

Fraunhofer. (2018). Maschinelles Lernen - Kompetenzen Anwendungen und Forschungsbedarf. Retrieved February 02, 2021, https://www.bigdata.fraunhofer.de/content/dam/bigdata/de/documents/Publikationen/BMBF_Fraunhofer_ML-Ergebnisbericht_Gesamt.pdf

Freiknecht, J., & Rapp, S. (2018). *Big Data in der Praxis. Lösungen mit Hadoop, Spark, HBase und Hive. Daten Speichern, Aufbereiten, Visualisieren* (2nd ed.). Carl Hanser Verlag.

Frencken, W., Poel, H., Visscher, C., & Lemmink, K. (2012). Variability of inter-team distances associated with match events in elite-standard soccer. *Journal of Sports Sciences, 30*(12), 1207–1213.

Fried, G., & Mumcu, C. (2017). *Sport Analytics. A Data-Driven Approach to Sport Business and Management*. Routledge.

Frick, B. (2004). Die Voraussetzungen sportlichen und wirtschaftlichen Erfolges in der Fußball-Bundesliga. In M. Bieling, M. Eschweiler, & J. Hardenacke (Eds.), *Business-To-Business-Marketing im Profifußball* (pp. 71–93). Deutscher Universitätsverlag.

Gabbett, T. J. (2016). The training—injury prevention paradox: Should athletes be training smarter and harder? *British Journal of Sports Medicine, 50*(5), 273–280.

Gassmann, O., & Perez-Freije, J. (2011). Eingangs-, Prozess- und Ausgangskennzahlen im Innovationscontrolling. *Controlling & Management Review., 55*(6), 394–396.

Gehrmann, S. (2017). Die Vermessung Des Sports. Retrieved March 03, 2021, https://www.bi-scout.com/die-vermessung-des-sports

Gerrard, B. (2016). Understanding sports analytics. Retrieved February 28, 2021, https://winningwithanalytics.com/2016/06/22/first-blog-post/.

Görlich, P., & Mayer, J. (2018). Falldarstellung: TSG 1899 Hoffenheim-Herkunft und Strategie. In R. Lanwehr & J Mayer (Eds.), People Analytics im Profifußball. Springer Fachmedien GmbH.

Gould, S. J. (1996). *Das Missverständnis Des Menschen.* WW Norton & Company.

Gowda, M., Dhekne, A., Shen, S., Choudhury, R. R., Yang, X., Yang, L., Golwalker, S., & Essanian, A. (2017). Bringing IoT to sports analytics. In *14th USENIX Symposium on Networked Systems Design and Implementation.* Boston, MA, USA.

Gramlich, D. (2018). Technische Grundlagen der Datensammlung und -analyse und ihre Auswirkungen auf den Fußball. Big Data im Fußball. Studylab 2018. Norderstedt: Books on Demand GmbH.

Groll, A., Ley, C., Schauberger, G., & van Eetvelde, H. (2018). Prediction of the FIFA World Cup. A random forest approach with an emphasis on estimated team ability parameters. Retrieved February 28, 2021, https://arxiv.org/pdf/1806.03208.pdf2018

Gudmundsson, J., & Wolle, T. (2013). Computers, Environment and Urban Systems Football analysis using spatio-temporal tools. *Computers, Environment and Urban Systems.* Retrieved February 28, 2021, https://thomaswolle.net/article-resources/GudmundssonWolle_FootballAnalysisUsingSpatioTemporalToolsEA.pdf

Herberger, T. A., Oehler, A., & Wedlich, F. (2018). Third-Party-Ownership-Konstruktionen im Profifußball: Finanzwirtschaftliche Einordnung, kritische Bestandsaufnahme und Implikationen. In T. Herberger (Ed.), *Sportökonomie im Kontext Von Governance & Gesellschaft* (pp. 129–158). Verlag Dr. Kovač.

Herberger, T. A., Oehler, A., & Wedlich, F. (2019). Third party ownership arrangements: Is a ban in football really appropriate? *Journal of Governance & Regulation, 8*(3), 47–57.

Hughes, M. D., & Bartlett, R. M. (2002). The use of performance indicators in performance analysis. *Journal of Sports Sciences, 20*(10), 739–754.

IEEE. (2015). Towards a definition of the Internet of Things (IoT). Retrieved February 28, 2021, https://iot.ieee.org/images/files/pdf/IEEE_IoT_Towards_Definition_Internet_of_Things_Issue1_14MAY15.pdf(1)

Ikram, M. A., Alshehri, M. D., & Hussain, F. K. (2015). Architecture of an IoT-based system for football supervision (IoT Football) 2nd World Forum on Internet of Things (WF-IoT) 2015 (pp. 69–74). IEEE.

Karkazis, K., & Fishman, J. R. (2017). Tracking US professional athletes: The ethics of biometric technologies. *American Journal of Bioethics, 17*(1), 45–60.

Kannekens, R., Elferink, T., & Visscher, C. (2011). Positioning and deciding: Key factors for talent development in soccer. *Scandinavian Journal of Medicine and Science in Sports, 21*(6), 846–852.

Kempe, M., Grunz, A., & Memmert, D. (2015). Detecting tactical patterns in basketball. Comparison of merge self-organising maps and dynamic controlled neural networks. *European Journal of Sport Science*, *15*(4), 249–255.

King, N. M. P., & Robeson, R. (2007). Athlete or guinea pig? Sports and enhancement research. *Studies in Ethics, Law, and Technology*, *1*(1).

King, N. M., & Robeson, R. (2013). Athletes are guinea pigs. *American Journal of Bioethics, 13*(10), 13–14.

Kipper, G., & Rampolla, J. (2013). *Augmented reality*. Elsevier Inc.

Kobielus, J. (2018). Wikibon's 2018 big data analytics trends and forecast. Retrieved February 28, 2021, https://wikibon.com/wikibons-2018-big-data-analytics-trends-forecast/

Laukenmann, J. (2017). Forscher *entschlüsseln* den Erfolg im Fußball. Retrieved February 28, 2021, https://www.tagesanzeiger.ch/wissen/technik/entschluesselung-des-fussballs-dank-big-data/story/30186152?track

Lewanczik, N. (2018)*:* Was wird wirklich relevant, was bleibt Zukunftsmusik?. Retrieved February 28, 2021, https://onlinemarketing.de/news/technologie-trends-2018-relevant-zukunft. Technologie. Trends.

Lewis, M. (2004). *Moneyball: The art of winning an unfair game*. W. W. Norton, & Company Inc.

Link, D. (2018a). Data analytics in professional soccer. Performance analysis based on spatio-temporal tracking data. Springer Fachmedien GmbH.

Link, D. (2018b). Sports analytics: Wie aus (kommerziellen) Sportdaten neue Möglichkeiten für die Sportwissenschaft entstehen. *German Journal of Exercise and Sport Research*. Deutschland: Springer-Verlag GmbH, *48*(1), 13–25.

Lucey, P., Bialkowski, A., Carr, P., Morgan, S., Matthews, I., & Sheikh, Y. (2013). Representing and discovering adversarial team behaviors using player roles. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2706–2713).

Mackenzie, R., & Cushion, C. (2013). Performance analysis in football. A critical review and implications for future research. *Journal of Sports Sciences*, *31*(6), 639–676.

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

McCall, A., Davison, M., Andersen, T. E., Beasley, I., Bizzini, M., Dupont, G., Duffield, R., Carling, C., & Dvorak, J. (2015). Injury prevention strategies at the FIFA 2014 World Cup: Perceptions and practices of the physicians from the 32 participating national teams. *British Journal of Sports Medicine, 49*(9), 603–608.

McHale, I. G., & Relton, S. D. (2018). Identifying key players in soccer teams using network analysis and pass difficulty. *European Journal of Operational Research, 268*(1), 339–347.

McKenna, B. (2017). T*SG* 1899 Hoffenheim gets "faster in the Head" with SAP analytics. Retrieved February 28, 2021, https://www.computerweekly.com/news/450423973/TSG-1899-Hoffenheim-gets-faster-in-the-head-with-SAP-analytics

Mangold, M. (2017). Unterschied virtual reality (VR) und augmented reality (AR). Retrieved February 28, 2021, https://magic-holo.com/unterschied-virtual-reality-vr-und-augmented-reality-ar/

Massaro, M., Dumay, J. C., & Guthrie, J. (2016). On the shoulders of giants: Undertaking a structured literature review in accounting. *Accounting, Auditing and Accountability Journal, 29*(5), 767–801.

Memmert, D., & Raabe, D. (2017). Revolution im Profifußball. Mit Big Data zur Spielanalyse 4.0. Deutschland: Springer-Verlag GmbH.

Memmert, D., & Raabe, D. (2018). *Data analytics in football: Positional data collection, modelling and analysis*. Routledge.

Meyer, J.-U. (2017). 2.0. Die Digitalisierung der Sportbranche. E. *Sport*-1st Aufl. Göttingen: BusinessVillage GmbH.

Millington, B., & Millington, R. (2015). The datafication of everything: Toward a sociology of sport and big data. *Sociology of Sport Journal*, *32*(2), 140–160, Sects. 140–160.

Möller, K., & Schönefeld, C. (2011). Innovation Performance Measurement Framework - Ein Referenzmodell zur Analyse der Innovationssteuerung. *Controlling & Management, 55*(6), 367–371.

Nafus, D., & Sherman, J. (2014). Big data, big questions|This one does not go up to 11: The quantified self-movement as an alternative big data practice. *International Journal of Communications, 8*, 1784–1794.

O'Donoghue, P. (2005). Normative profiles of sports performance. *International Journal of Performance Analysis in Sport, 5*(1), 104–119.

Perl, J., & Memmert, D. (2017). A Pilot study on offensive success in soccer based on space and ball control-key performance indicators and key to understand game dynamics. *International Journal of Computer Science in Sport, 16*(1), 65–75.

Pickup, O. (2018). Three major ways cloud is transforming sport. Retrieved February 28, 2021, https://www.raconteur.net/technology/three-major-ways-cloud-transforming-sport

Quirling, C., Kainz, F., & Haupt, T. (2017). *Sportmanagement*. Verlag Franz Vahlen GmbH.

Reichmann, T., Kißler, M., & Baumöl, U. (2017). *Controlling mit Kennzahlen. Die systemgestützte Controlling-Konzeption* (9th ed.). Verlag Franz-Vahlen GmbH.

Rein, R., & Memmert, D. (2016). Big data and tactical analysis in elite soccer: Future challenges and opportunities for sports science. *Springerplus, 5*(1), 1410.

Research and Markets. (2016). Worldwide sports analytics market (2016–2022). Retrieved February 28, 2021, https://www.researchandmarkets.com/research/gpr3fw/worldwide_sports

Ribeiro, J., Silva, P., Duarte, R., Davids, K., & Garganta, J. (2017). Team sports performance analysed through the lens of social network theory. Implications for research and practice. *Sports Medicine*, *47*(9), 1689–1696.

Ross, W. J., Beath, C. M., & Quaadgras, A. (2013). You may not need Big Data after all. Retrieved February 28, 2021, https://hbr.org/2013/12/you-may-not-need-big-data-after-all

Saam, S. (2017). DFB-Akademie: 'Silicon Valley' des Fußballs. Retrieved February 28, 2021, https://www.dw.com/de/dfb-akademie-silicon-valley-des-fu%C3%9Fballs/a-41669374

Schauberger, G., & Groll, A. (2018). Predicting matches in international football tournaments with random forests. *Statistical Modelling*, *18*(5–6), 460–482. Retrieved February 28, 2021, https://journals.sagepub.com/doi/full/10.1177/1471082X18799934

Schoop, M., & Brauchle, A. (2016). Anwendung von Data-Mining-Technologien zu Statistischen Auswertungen und Vorhersagen Im Fußball. Universität Hohenheim.

Söhnlein, K., & Borgmann, S. (2018). Diagnostik von Exekutivfunktionen im Fußball. In R. Lanwehr & J. Mayer (Hrsg.). People Analytics im Profifußball. Springer Fachmedien GmbH.

SportHeads. (2018). Sport professionals 2018. Eine Studie zum Arbeitsumfeld des Teams Hinter dem Team. München.

SportTechie. (2018). Techie and N3XT sports state of soccer technology & innovation. *Sport*. Retrieved February 28, 2021, https://www.sporttechie.com/sporttechie-and-n3xt-sports-state-of-soccer-technology-innovation/

Thite, M. (2018). *e-HRM: Digital approaches, directions & applications*. Taylor & Francis Group.

Thür, H. (2015). Die Privatsphäre im Zeitalter von big Data. Jusletter IT, 21.

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence—Informed management knowledge by means of systematic review. *British Journal of Management, 14*(3), 207–222.

van der Westhuizen, E. J., & van der Haar, D. T. (2018). A wearable device-based framework for determining player effectiveness on the football pitch. *ACM International Conference Proceeding Series*, 226–231.

Vanini, U., & Rieg, R. (2017). *Effects of voluntary intellectual capital disclosure—a structured literature review*. University of Applied Sciences.

Weichert, T. (2013). Big data. Eine Herausforderung für den Datenschutz. In: Geiselberger, H., & Moorstedt, T. (Eds.). Big Data. Das Neue Versprechen der Allwissenheit. Suhrkamp Verlag.

Werner, H. (2017). Performance-Messung in Forschung und Entwicklung. Controlling & management review. 3/2017, 16–26.

Wired (2018). The unlikely secret behind Benfica's fourth consecutive Primeira Liga title. Retrieved February 28, 2021, https://www.wired.co.uk/article/bc/microsoft-sl-benfica

Zobel, B., Werning, S., Metzger, D., & Thomas, O. (2018). Augmented und Virtual Reality: Stand der Technik, Nutzenpotenziale und Einsatzgebiete. In: de Witt, C., & Gloerfeld, C. (Hrsg.). Handbuch Mobile Learning. Wiesbaden: Springer Fachmedien GmbH.

Zoph, B., Vasudevan, J., Shlens, J., & Le, Q. (2017). AutoML for large scale image classification and object detection. Retrieved February 28, 2021, https://ai.googleblog.com/2017/11/automl-for-large-scale-image.html