Journal of Infrastructure, Policy and Development 2024, 8(11), 7489.
https://doi.org/10.24294/jipd.v8i11.7489

Article

Digital technology for the national sport development: Designing a database model to analyse elite sports data

Alan de Carvalho Dias Ferreira[1,2,*], Pedro Sobreiro[1], Júlio B. Mello[3], Alberto Reinaldo Reppold Filho[4]

[1] Sport Physical Activity and Health Research & Innovation Center, Sports Science School of Rio Maior, Polytechnic Institute of Santarem, 2040-413 Santarém, Portugal
[2] Faculty of Physical Education and Sport, Lusofona University, 1749-024 Lisbon, Portugal
[3] eFiDac Research Group, School of Physical Education, Pontificia Universidad Católica de Valparaíso, Valparaiso 2530388, Chile
[4] Faculty of Physical Education, Physiotherapy and Dance, Federal University of Rio Grande do Sul, Olympic and Paralympic Studies Centre, Porto Alegre 90690-200, Brazil
*** Corresponding author:** Alan de Carvalho Dias Ferreira, alan.ferreira@esdrm.ipsantarem.pt

CITATION

Ferreira AdCD, Sobreiro P, Mello JB, Reppold Filho AR. (2024). Digital technology for the national sport development: Designing a database model to analyse elite sports data. Journal of Infrastructure, Policy and Development. 8(11): 7489.
https://doi.org/10.24294/jipd.v8i11.7489

ARTICLE INFO

Received: 26 June 2024
Accepted: 21 August 2024
Available online: 17 October 2024

Abstract: Relational database models offer a pathway for the storage, standardization, and analysis of factors influencing national sports development. While existing research delves into the factors linked with sporting success, there remains an unexplored avenue for the design of databases that seamlessly integrate quantitative analyses of these factors. This study aims to design a relational database to store and analyse quantitative sport development data by employing information technology tools. The database design was carried out in three phases: (i) exploratory study for context analysis, identification, and delimitation of the data scope; (ii) data extraction from primary sources and cataloguing; (iii) database design to allow an integrated analysis of different dimensions and production of quantitative indicators. An entity-relationship diagram and an entity-relationship model were built to organize and store information relating to sports, organizations, people, investments, venues, facilities, materials, events, and sports results, enabling the sharing of data across tables and avoiding redundancies. This strategy demonstrated potential for future knowledge advancement by including the establishment of perpetual data updates through coding and web scraping. This, in turn, empowers the continuous evaluation and vigilance of organizational performance metrics and sports development policies, aligning seamlessly with the journal's focus on cutting-edge methodologies in the realm of digital technology.

Keywords: sport development; sport management; digital technology; business intelligence; software; sports policy; data analysis

1. Introduction

Systematised information that fulfils sport's managerial function are required for a country to achieve a national sport development strategy (Exel and Dabnichki, 2024; Rajšp and Fister, 2020; Strittmatter et al., 2018). Many initiatives consider data analysis through qualitative methods, not using Data Science methods or information technology tools (Bohlmann and Heerden, 2008; Camps and Pappous, 2016; Ferreira et al., 2018; Ward et al., 2019). Indeed, few studies analyse sports management data approaching the subject through a quantitative perspective (Budovich, 2021; Cossich et al., 2023; Sun et al., 2022; Ward et al., 2019). Link (2018), Sands et al. (2017), and Ofoghi et al. (2013) described the sophisticated informational technology tools and analytical methods such as data mining, statistics, machine learning, pattern recognition, database and data warehouse systems, information retrieval, visualisation, algorithms, and high-performance computing. These proposals are becoming increasingly valuable tools in analysing sports performance and development, and supporting decision-making for the sport success, participation and sport for all policies (Budovich, 2021). Databases are employed as structured information and used as a source of historical data to identify trends and patterns (O'Boyle, 2015).

According to Ward et al. (2019), managing data systems models can support decision-making to improve both the efficiency and effectiveness of decisions and comprises three areas; (1) data collection and organisation, (2) analytic models to drive insight, and (3) interface and communication of information. For example, Qiao (2022) presented an intelligent Big Data framework for the technical design of public management applications in sports. It provides an objective description and analysis of the current situation of public sports management services development and an outlook on its future (Exel and Dabnichki, 2024; Qiao, 2022). Also, Sun et al. (2022) proposed a model for evaluating the competitiveness of the sports and cultural industries based on a fuzzy analytic algorithm and a seven-dimensional sports and cultural industry competitiveness evaluation index system in China. In the same way, large technology companies are developing and adapting their data analysis tools for sports, such as the data analytics tool "Sports One" from the company SAP, used to improve decision-making in German football (SAP—Unveils SAP Sports One Solution for Soccer).

Therefore, there is an emerging interest in developing business intelligence models for elite sports (Cossich et al., 2023; Sands et al., 2017). While quantitative predictive research uses data analysis, this has primarily been utilised for sports betting and result prediction (Camps and Pappous, 2016; Dabnichki et al., 2020). Nevertheless, few initiatives have been developed to construct database models for managing information and continually analysing the multiple quantitative inputs and outputs from elite sports success pillars, critical factors or variables (Ferreira et al., 2018; Link, 2018; Madella et al., 2005; Watanabe et al., 2021). Additionally, relatively few database models provide sports managers with quantitative information and predictions about the sports development determinants (Budovich, 2021; Sun et al., 2022), even though the construction of an analytical data model can provide managing information and continually analyse the multiple quantitative inputs, throughput and outputs in the sports management landscape (Strittmatter et al., 2018; Ward et al., 2019).

Some studies have provided strong evidence that sports development should not just be about outcomes (i.e., the number of Olympic medals) but also other factors that lead to sport success and their relationships with national sport development (De Bosscher et al., 2015, 2016; Grix, 2013). Previous studies have focused on sports funding, the set of organisations that are part of the sports system, the participation of the population in sports, talent identification, the support given to athletes during and after their career, sports facilities and equipment, coach development, national and international competitions, and sports media coverage to analyse the elite sports success and sports development (De Bosscher et al., 2015; Green and Oakley, 2001; Houlihan and Green, 2008).

In this context, a database can be beneficial once it is an organised collection of data stored and accessed electronically, so formally, a "database to analyse sports data" refers to a set of related data and how it is organised. Using information technology tools to access this data is usually provided by a database management system (DBMS) consisting of an integrated set of computer software that allows users to interact with one or more databases and provides access to all its data (Shin, 2022). So, from the architectural information view of previous research on sports success and development (De Bosscher et al., 2015; Houlihan and Green, 2008; Rewilak, 2021), a database model containing multiple quantitative data can provide their ongoing quantitative analysis and an understanding of their interactions and cause-effect relationships (Rewilak, 2021).

In this sense, a sports development database model can benefit the national sport system since, as Budovich (2021) and Wanless (2022) posited, the technology used to analyse information is a vital resource for organisations. The literature highlights the challenge organisations face in the collection and storage of data in a way that facilitates methods of data analysis, producing valuable information (data with meaning) that allow the creation of knowledge and values for organisations from these technologies (Link, 2018), and assessing organisational performance and the sports systems development (Shilbury and Moore, 2006; Winand et al., 2014). According to Houlihan and Green (2008), Ofoghi et al. (2013), and Ferreira et al. (2018), management information and data analysis tools for public or private sectors that provide precise and ongoing information about the development and funding of the sports system should be very useful for national sport development. Therefore, a database model could help researchers and managers answer several quantitative questions: Do sports policies meet the sector's future needs? What type of investment can be more effective? How developed is the network of sports facilities for each sport? What is the ideal talent discovery rate? What are the main quantitative outputs from the elite sports financing input?

Thus, this study aims to design a database to store and quantitatively analyse elite Brazilian sports data, employing information technology (IT) tools and covering the fundamental pillars and factors for international sporting success. The database construction could provide the information necessary for qualifying the actions of sports managers, which, together with good management, could assist with the policies, given that the required answers, reports, and diagnostics produced from the model could contribute to efficient communication between different organisations and prevent overlapping actions.

2. Materials and methods

2.1. Research design

The design of the database model to quantitatively analyse data from Brazilian elite sports was carried out according to the methods described by Kroenke and Auer (2007); Shin (2022); Mamo (2023) in three phases (**Figure 1**). The data was collected during the 2016 Olympic and Paralympic Games in Rio de Janeiro, Brazil (International Olympic Committee), considering the meso-level elements described by De Bosscher et al. (2006), that is, the factors that sports-related policies can impact.

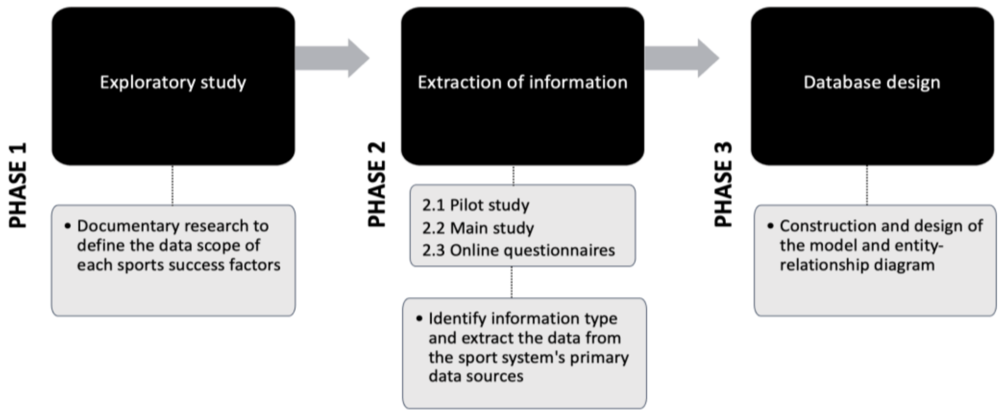

Figure 1. Phases for designing the database model to analyse data from the Brazilian elite sport.

2.2. Exploratory study

For the context analysis, identification and delimitation of the scope of quantitative data that makes up each sporting success factor (De Bosscher et al., 2006), qualitative exploratory research was conducted utilising documentary research. Searches were conducted in the Web of Science, Scopus and SportDiscuss databases, using the keywords: sport success factors, sports policy, sports diagnosis, sports system, sport management, sporting success, high-performance sport, elite sport systems, elite sport development and international sporting success. The Boolean operators OR and AND and truncators and * were used to search for exact phrases and same word variations, respectively. Content published between 2002 and 2022 was considered.

Studies were included that aimed to conduct diagnoses of the sporting situation; investigations of the management structure, administration, and policies of sport; research on the integrated financing, structure, and governance system, talent identification and development, sports facilities, coach training, and participation in elite sports competitions. After excluding studies that did not describe their methods and/or instruments for data collection, cataloguing and analysis, the elements of the nine pillars described by De Bosscher et al. (2015) were considered, in addition to the elements cited by Houlihan and Green (2008). These elements were organised into eight dimensions describing the quantitative data collected in the following database design steps (**Table 1**).

Table 1. Dimensions and information used to collect data on Brazilian elite sport.

Dimension	Data collected to design the database	Primary data source(s)
(1) Funding	Federal funding, state and municipal public financial resources invested in high-performance sports. In addition to private resources coming from the confederations' revenues and sponsorships from private companies	Ministry of Sports; Brazilian Olympic Committee (COB) and Brazilian Paralympic Committee (CPB) accountability reports of the funds received from the Agnelo Piva Law; Federal Revenue's E-CAC Portal; Accounting statement of the competence regime (DRE) of National Sports Administration Entities.
(2) Legislation	Federal, district, state and municipal normative acts that govern the sport in the national territory	Ministry of Sports; State governments, Federal Districts and Brazilian municipalities.
(3) Organisations	Public and private organisations, national, state and municipal that manage sports practice and competitions; that are sources or destinations of financial resources applied to sports; that organise or promote sports events; that produce science and technology related to sports; and that manage sports facilities	Ministry of Sports; Sports Confederations, Federations and Sports Clubs.
(4) Athletes and Sports Professionals	Athletes linked to confederations and federations; those who make up the national teams; athletes who participate in official competitions; professionals hired by confederations, Federations, and sports clubs, with training in different areas of knowledge	Athletes; Ministry of Sports; Confederations, Federations, and Sports Clubs.
(5) Events and Sport Results	Sports events organised by committees, confederations, and federations, on an international and national level, with the participation of athletes, sports clubs and Brazilian teams	International Olympic Committee and International Paralympic Committee; International Sports Federations and Organisations; National Confederations.
(6) Sports Facilities	Every place or space, uni or multipurpose, built, prepared, or used for sports training and competition, which contains the minimum sports materials for training and/or competition	State, Federal District and Brazilian municipal governments; Ministry of Sports; Sports Confederations, federations, and sports clubs.
(7) Sports Materials	Materials used by the athlete for training and/or for competition; materials used exclusively by the organisation of competitions; materials used in the development of physical qualities or for physical evaluations/tests	Ministry of Sports; Confederations and Committees; manufacturing companies.
(8) Science and Technology	Private organisations and public institutions that perform tests, evaluations, training monitoring, scientific research, software development, technological innovation, contract specialised services or specialists or purchase equipment that contribute to sports training. Additionally, there are theses, dissertations and scientific articles related to sport; laboratories and research groups; scientific journals and Physical Education schools, courses and congresses held in the area.	Coordination for the Improvement of Higher Education Personnel (CAPES); Brazilian Institute of Information in Science and Technology (IBICT); National Council for Scientific and Technological Development (CNPq); Public Enterprise Financier of Studies and Projects (FINEP).

2.3. Extraction of information from sports system primary sources

Aiming to identify the type of information and extracting data from the primary data sources of the sports system in question, based on the dimensions and sources established in the first stage (**Table 1**), quantitative and descriptive research was carried out in three subsequent stages, such as:

(i) Pilot study: before the final data extraction, data collection of the eight dimensions (**Table 1**) from two sports (Olympic athletics and Paralympic swimming) was carried out, for calibration of the instruments and procedures, with the elimination or the reduction of the error between the intended data, the database attributes, and the collected contents (Kroenke and Auer, 2007; Mamo, 2023; Shin, 2022).

(ii) Main study: after the adjustments allowed by the pilot study, we proceeded with the data collection from 2008 to 2016 (two Olympic cycles) of the Olympic and Paralympic sports in Brazil, considering the Rio 2016 programme (International Olympic Committee), the primary sources and the scope of the dimensions described in **Table 1**.

(iii) Questionnaire application: to complement the previous phases and obtain the scope of data in its entirety, an electronic questionnaire developed in the Lime Survey platform was also integrated with a MySQL database and database management system (DBMS) that used SQL (Structured Query Language) as an interface (Shin, 2022). The electronic questionnaire was designed according to the premises of Madella et al. (2005), consisting of 20 standardised questions divided into the eight dimensions described in **Table 1**. In short, the questions allowed us to obtain the set of sports organizations for each sport; data on people linked to sports entities (athletes, managers, coaches, etc.); national and international sports results of athletes and teams in sports events; programs and actions to identify sports talents; set of sports facilities and equipment used by each sport; in addition to the sets of resources and people involved in scientific research in each sport. The instrument was applied to a sample of 6423 athletes directly funded by the Brazilian Federal Government, who compete in all Olympic and Paralympic sports. Complete responses received from a representative sample of elite athletes funded by the Brazilian Federal Government, including athletes from all Olympic and Paralympic sports. (3897 elite athletes) were considered for database design. It was also answered by 34 public and private organisations (n = 5) that are part of the Brazilian sports system, including the Ministry of Sports, the Brazilian Olympic Committee; the Brazilian Paralympic Committee; a representative sample of National Confederations (n = 20); State Federations (n = 5), and sports clubs (n = 5).

According to the dimensions described in **Table 1**, the data collected were catalogued in a standardised manner in a database (DB), modelled to allow the articulation between the dimensions and provide elements for an integrated approach to the RAS. A Glossary of Terms and a Data Dictionary were prepared to guide the cataloguing, validate the information's format and content, and register, classify, interpret, and analyse the data. The cataloguing of the data from the Data Dictionary, according to the procedures described by Shin (2022), allowed the understanding of the DB terms and information, from the data collection to the programming phase; the identification of the system's tables, their attributes, existing relationships, and restrictions, facilitating data handling, analysis, and programming.

Therefore, the information of the eight dimensions composed 33 tables of a MySQL database, divided by sport and dimension. When possible, the data from Olympic sports were stratified by event (medal competition). Besides stratification, the Paralympic sports data was subdivided according to athlete functional class. Considering the data from two Olympic/Paralympic cycles (2008 to 2016), 427,395 entries were recorded in the database, a quantity represented mainly by the number of athletes, their sporting results, the respective sporting events, public and private sporting organisations, and their funding. Then, a database system was designed based on the Brazilian Olympic and Paralympic elite sports data via Entity-Relationship Diagram (ERD), relational model, and normalisation of tables implemented in an SQL server (Kroenke and Auer, 2007; Shin, 2022).

2.4. Database design

The design of the database and the Entity-Relationship Diagram (ERD) aimed to build a decision support tool under two premises: (i) the model should allow an integrated analysis of the information and, therefore, of the dimensions or sports success factors, through the ERD; and, (ii) it should allow, through IT tools, permanent updating of the stored information, in addition to the production of reports and visualisation of indicators through Business Intelligence (BI) tools (Ward et al., 2019).

The method for cataloguing and categorising the information followed the concepts and scope of the studies identified in the exploratory study. Normalisation was performed to reduce data redundancy and improve data integrity by making relationships free of insertion, update, and deletion anomalies (3NF relationships). For this, the relationships were established in 1NF (first normal form) by eliminating sub-attributes from the relationship, in 2NF by making the non-primary keys dependent on the primary key, and in 3FN by eliminating transitive dependency between the non-primary keys and the primary key (Codd, 1970). After normalisation, we represented relations between tables with the cardinalities 1-1, 1-N and N-N. The N:N relations were implemented with a third table in the relation.

These associations were made using foreign key and primary key, indexing records that did not repeat and were used as an index for the other fields of the database tables. In this form of processing, a Database Management System (DBMS) and ERD were used so that each piece of information was stored only once. This redundancy control is used to improve data reliability by preventing the same information from being catalogued repeatedly, which ensures data consistency and overall information system performance (Shin, 2022).

Finally, the physical entities were defined as tangibles (e.g., organisations, athletes and sports professionals), and the logical entities were those that occurred as a result of the interaction between or with physical entities, as in the case of the categorisation of financial resources according to their type of application. Furthermore, associative entities result from the need to associate an entity with an existing relationship (Kroenke and Auer, 2007; Shin, 2022). The Entity-Relationship Model (ERM) and the EDR are described below in the Results section.

3. Results

3.1. Entity relationship model (ERM) and entity relationship diagram (ERD)

The entity-relationship model (ERM), as for the information aspects and business domain, was designed considering the eight dimensions identified in the exploratory study (**Table 1**), aiming to contemplate all the pillars and factors of success and sports development. The ERM comprises 55 divided tables, among which 20 are physical entities that represent the quantitative data linked to sports success, and 35 tables represent logical entities. **Figure 2** presents a simplified structure of the complete entity-relationship model. Logical entities are those built during modelling and establish the criteria for data classification and normalisation, such as pre-determining the possible values for database fields like the tables that store the names of continents, countries, cities, sports and their events, the types of applications of financial resources, the types of sports facilities, for example. The ERD is represented in Annex A with all the entities and the relationships between the tables.

The dimensions "Organisations" and "Athletes and sports professionals" are composed of the tables with the highest number of binary and ternary relationships (1:1, 1:N, N:N) with the other tables in the database and are, therefore, fundamental for integrating quantitative data on sports success factors. Each element was coded and stored without redundancy utilising primary keys (i.e., unique identifiers for all the information in each record), giving them uniqueness and producing accurate analyses. In a practical approach, it can be visualised in the ERD that organisations are connected to funding, facilities, materials, athletes and sports professionals, and sporting events. Similarly, athletes connected to organisations are also related to funding and sports outcomes (**Figure 2**).

In other words, an organisation can be identified as the source or the destination of funding; it can be the managing organisation of a sports facility or organising a sporting event at a specific sports facility. Similarly, "Athletes and sports professionals" are connected through the relationship to the organisations stored in the database. Thus, when an athlete/organisation/nation participates in an event, in addition to being able to check their/its results or success rate, it is also possible to analyse the support (financial, infrastructure, human resources) received by the athlete in a given period. The cardinality rates for entities connected to a relationship are explained in the next section.

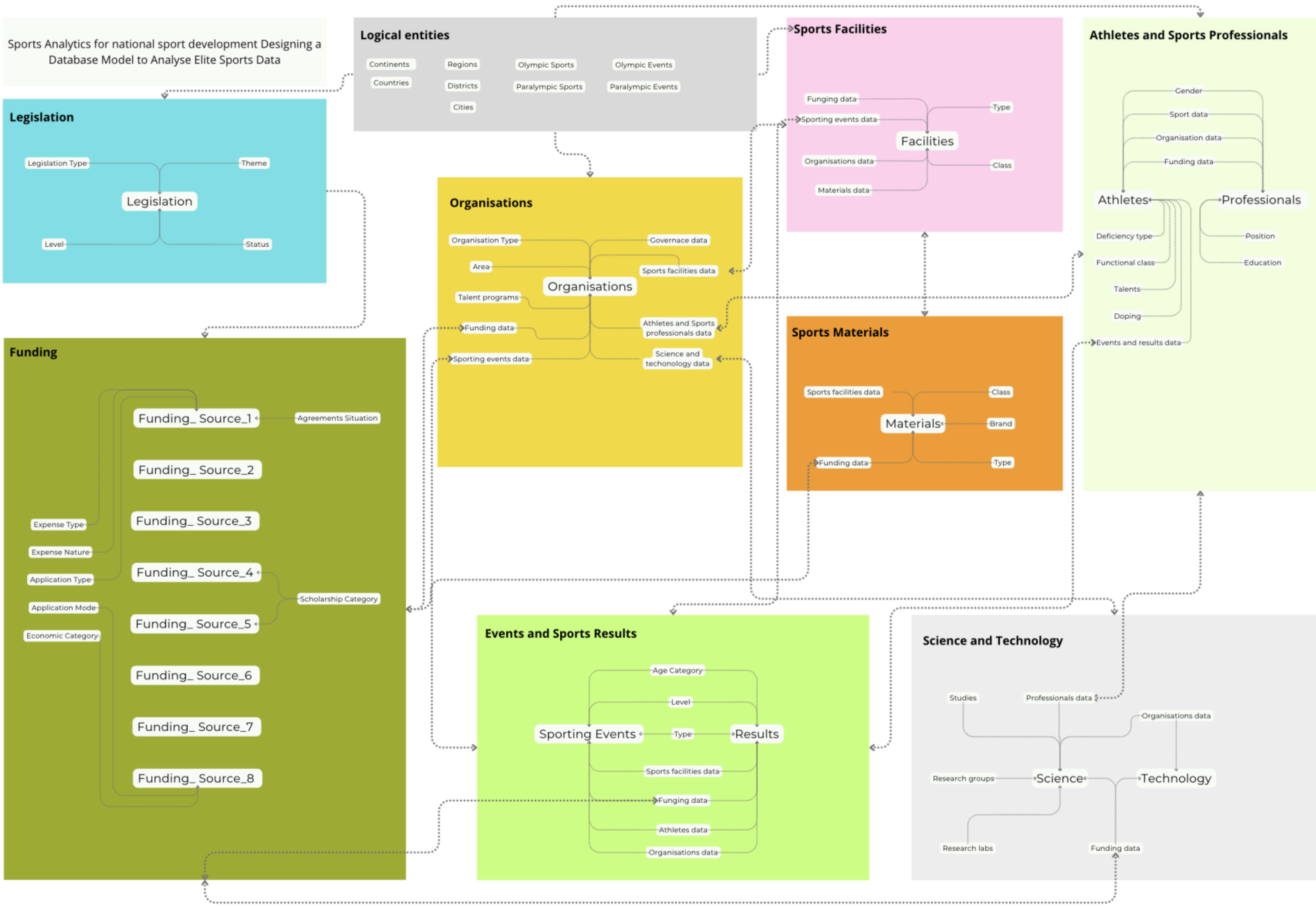

Figure 2. Simplified diagram for entities, dimensions, and database model relationships to analyse Brazilian elite sports.

3.2. Cardinalities

Binary and ternary relationships

In all dimensions of the designed model, the tables have binary relationships. Legislation is related to public funding; athletes and professionals are always related to their respective organisations, just as products of science and technology are related to the organisations that produce them. As seen in a different representation of the ERD (**Figure 3**), funding can occur through one or multiple sources. Each source can allocate funding to one athlete or organisation, or to many, which determines the N:N relationship called the "Funding Source" (enti_id_fund_id, enti_id_athl_id).

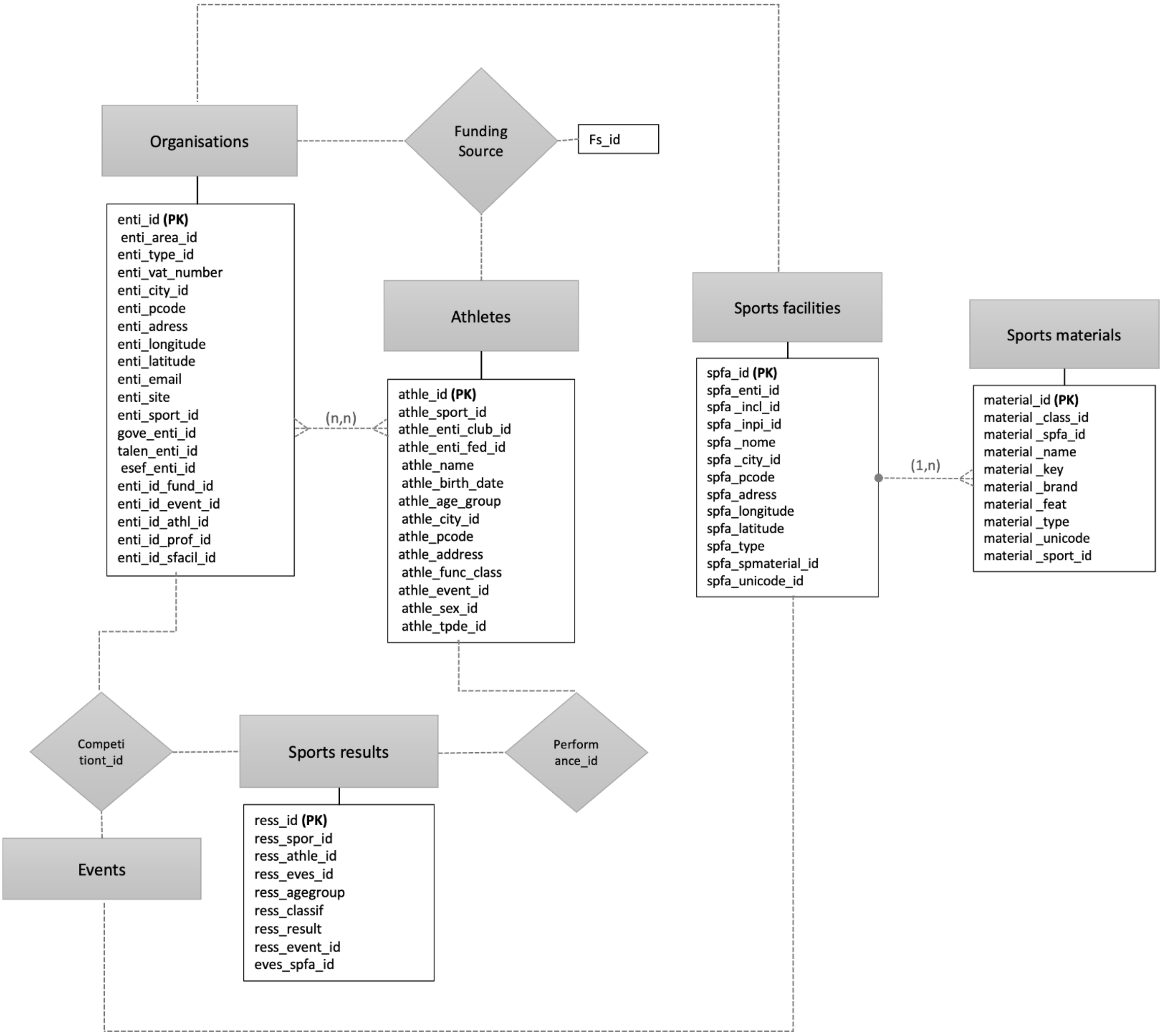

Figure 3. Examples of binary and ternary relationships from the entity-relationship diagram.

In the case of sports facilities, there is a 1:N relationship between facilities and sports materials; that is, each sports facility is connected to several sports materials that are part of that infrastructure (**Figure 3**). The same occurs with sporting events, in which sporting results are connected to the event, like age category and gender of the athlete (1:N), related through binary relations.

In **Figure 3**, ternary relationships can also be seen. In these cases, a logical entity has been created for ternary relationships such as Competition_ID and Performance_ID (see rhombuses). The Competition_ID entity relates to organisations (1:N) and sports results (1:N) (i.e., an event is always associated with a host organisation and simultaneously to multiple sports results of the same event). Similarly, the Performance_ID entity relates athletes and teams to results (N:N) and, the results to events (1:N). Similar logic can be observed in other ternary relationships presented in the ERD in **Figure 2**.

3.3. Relational Model

The entity-relationship diagram shown in **Figure 2** can be fully converted into an entity-relationship model (ERM), as in **Table 2**. The ERM defines the criteria for categorising the data and the relationships in each dimension and each model table. The main attribute(s) of a relationship, the primary keys of the tables, are those in bold, while attributes representing foreign keys are indicated as "fk" (foreign key). The Appendix presents the names of all variables for each dimension included in the entity-relationship model (see **Table A1**).

Table 2. Dimensions, main attributes, primary keys, foreign keys and relationships of the entity-relationship model.

Dimensions	Attributes, primary keys, foreign keys	Relationships
(1) Funding	fund_id, fund_value, fund_dateInitial, fund_dateFinal, fund_year, sport_id (fk), city_id (fk)	fund_athle_id, fund_tpap_id, fund_legtype_id, fund_enti_id
(2) Legislation	legi_id, legi_date, legi_summary, legi_number, legi_link, city_id (fk),	legi_theme_id, legi_type_id, leg_level_id, legi_id_fund_id
(3) Organisations	enti_id, enti_vat_number, enti_pcode, enti_adress, enti_longitude, enti_latitude, enti_email, enti_site, city_id (fk), sport_id (fk)	enti_area_id, enti_type_id, talen_enti_id, stech_enti_id, enti_id_fund_id, enti_id_event_id, enti_id_athle_id, enti_id_prof_id, enti_id_spfa_id
(4) Athletes and Sports Professionals	athle_id, athle_name, athle_birth_date, sport_id (fk), city_id (fk), athle_pcode, athle_adress, sex_id (fk), age_group_id (fk), func_class_id (fk)*, deftype_id (fk)*, prof_id, prof_name, prof_ birth_date, sex_id (fk), city_id, prof_pcode, prof_adress	athle_enti_club_id, athle_enti_fed_id, athle_event_id, prof_enti_id, prof_area_id, prof_posit_id
(5) Events and Sports Results	eves_id, eves_name, eves_year, eves_date, city_id (fk), sport_id (fk), ress_id, sport_id (fk), event_id (fk), age_group_id (fk), func_class_id (fk)* ress_classif, ress_result	eves_type_id, event_sex_id, eves_freque_id, eves_level_id, eves_enti_id, eves_spfa_id, ress_athle_id, ress_eves_id
(6) Sports Facilities	spfa_id, spfa_name, city_id (fk), spfa_pcode, spfa_adress, spfa_longitude, spfa_latitude, sport_id (fk)	spfa_class_id, spfa_type_id, spfa_enti_id, spfa_material_id, eves_spfa_id
(7) Sports Materials	material_id, spmaterial_key, spmaterial_name, spmaterial_brand, spmaterial_feat, sport_id (fk)	material _type_id, material _class_id, material _spfa_id
(8) Science and Technology	stud_id, stud_year, stud_author, sport_id (fk), lab_id, lab_name, rgroup_id, rgroup_name, conf_id, conf_name	stud_theme_id, stud_area_id, stud_type_id, lab_theme_id, lab_area_id, lab_enti_id, rgroup_theme_id, rgroup_area_id, stech_enti_id, conf_theme_id, conf_area_id, conf_enti_id

(1) Funding: this dimension identifies the sources of resources (e.g., fund_1; fund_2...), their allocation to organisations, sports, individuals and cities (e.g., fund_ent_id; fund_athl_id, fund_sport_id; fund_city_id), the amount invested

(fund_value), the funding period (fund_dateInitial, fund_dateFinal), in addition to the classification of the application form according to the object funded (fund_tpap_id) and according to the legislation (fund_legtype). After extracting data from primary sources (i.e., the Brazilian sports system) using the Glossary of Terms and the Data Dictionary, and considering total funding invested in high-performance sports between 2008 and 2016, we identified eight sources of funding for elite sport (agreements, transference contracts, athlete-scholarship, Brazil Medals Plan, decentralisations, Sports Incentive Law, sponsorships and private resources), and 12 different ways of applying such financial resources (athlete, human resources, training, competition, sports material, sports facilities, scientific research, training, technology, administrative expenses, marketing, health and performance evaluation).

(2) Legislation: the sports legislation was organised into a library of normative acts and can be consulted by type and level of legal diploma (legi_type_id; leg_level_id), classified into 26 different themes (legi_theme_id), in an articulated manner through a relationship with the financing dimension (legi_id.fund_id).

(3) Organisations: organisations (private non-profit), public administration entities, state companies, higher education institutions, and private for-profit companies (enti_id, enti_area_id, enti_type_id) were stored in the database. In the database, the organisations are related to logical entities that categorise them according to projects to identify sporting talent (talen_enti_id); to the financing of high-performance sport (enti_id_fund_id); with the organisation of sports events (enti_id_event_id); with the management of sports facilities (enti_id_sfacil_id); with the management of athletes and sports professionals (enti_id_athl_id, enti_id_prof_id); and, with the production of science and technology for sport (esef_enti_id).

(4) Athletes and sports professionals: in this dimension, attributes such as *func_class and deftype_id categorise athletes with disabilities according to their functional class and type of disability, a stratification also used in paralympic competitions. Athletes and professionals represent a single registry in the database that differentiates them between athletes (athle_id) and non-athletes (prof_id). While athletes can be queried according to their sport (sport_id) and their organisation (athle_enti_club_id, athle_enti_fed_id), professionals in areas such as management, refereeing, training, health, and marketing (prof_area_id) can be queried stratified into 139 different positions (prof_posit_id).

(5) Events and sports results: During data extraction from the primary sources, 11,303 sports events were registered in the database. These events and the sports results of two Olympic cycles (2008 to 2012 and 2013 to 2016; 69,496 sports results) were categorised into seven levels (eves_level_id): Olympic/Paralympic, world, continental, international, national, regional, local), for Olympic and Paralympic sports (eves_type). The developed model also allows the visualisation and analysis of events and results stratified by sex and athlete (sex_id; ress_athle_id), age category (agegroup_id), event (sport_id, event_id), functional class (for Paralympic sports events). In addition to classifying athletes and teams (1st, 2nd, 3rd), the mark achieved in the disputed events is also registered (e.g., time, height, weight, distance, score, points, victory).

(6) Sports facilities: from 11,234 sports facilities registered in the database,

logical entities were built to categorise the infrastructures into four classes (spfa _class_id) (basic, sets, complexes, and training centres) and 32 different types (spfa_type_id). In this modelling, sports and facilities were related from a matrix to identify each sport's ideal and alternative infrastructure(s). Thus, the sports in the database were linked only to the infrastructure type where their practice is effectively performed.

(7) Sports materials: this dimension is complementary to "Sports facilities" since, in addition to registering the materials used for training and competition according to their type, class and characteristics (material_class_id, material_brand, material_feat, material_type_id), it also registers the materials used in infrastructures where the sports are practised (material _spfa_id, material _sport_id).

(8) Science and technology: the last dimension of the model registers the scientific production through studies (stud_id) (theses, dissertations and articles), groups and research labs in the area of sport (lab_id, rgroup_id). Additionally, it relates the organisations (stech_enti_id) registered in the database with the distribution of science (for example, universities that hold courses, training and conferences) (conf_id) and of technologies (for example, companies that manufacture technological equipment used in sport) (stech_enti_id).

Through the foreign keys, the modelling built from elite sports data presents 151 relationships among the eight dimensions of the database (**Figure 2**). With this, it is possible to perform analyses, produce reports, and visualise indicators in an intra- or inter-dimensional, uni-sport or multi-sport way integrating several or all dimensions and sports included in the model.

3.4. Implementation in SQL server and system usage

A system was developed with the programming language PHP (free interpreted language), used in a specific domain for the web development field to store and analyse the data from the tables described in **Figure 2**. In conjunction with PHP, the CodeIgniter framework was used to enable faster project development. CodeIgniter was developed on the object-oriented programming paradigm under the MVC (Model-View-Controller) software architecture pattern. Also, on top of CodeIgniter, an open-source library called Grocery CRUD was used to facilitate insert, view, edit and delete operations. Apache (v. 2.3 open-source) was used as the application server, together with the MySQL database, which has open-source code and ease of programming and learning.

On this server, you can store data easily, retrieve data quickly and execute queries conveniently by SQL language. It is also possible to create relationships by code and manually. From these features of the designed database, some interesting queries can be imposed, including:

a) Which types of funding are associated with the sports with the most wins? What is the success rate of each funding type when considering the results achieved by athletes and teams? What funding is available per elite athlete?

b) Which sport(s) or athlete(s) has the most representative volume of support (financial, facilities, human resources)? How much support can determine medals or improved results? What financial resources are being mainly applied in what types of support?

c) Are the sports facilities available for practice proportional to the number of organizations and athletes in a given sport?

4. Discussion

This study aimed to design a database to store and quantitatively analyse elite Brazilian sports data, employing information technology (IT) tools and covering the fundamental pillars and factors for international sporting success. In this way, the proposed large-scale database initiative in sports has the power to significantly enhance a country's sports development by providing data-driven insights for informed decision-making, optimizing talent identification and training, guiding evidence-based policy formulation, tracking and improving athlete performance, efficiently allocating resources, fostering research and innovation, promoting transparency and accountability, facilitating strategic partnerships, and ultimately encouraging broader participation in sports through accessible and transparent data (Bai and Bai, 2021).

There is an emerging search for intelligent models that collaborate with decision-making in sport management and sport development policies (Budovich, 2021; Link, 2018; Sun et al., 2022; Ward et al., 2019). According to Camps and Pappous (2016), quantifying the financial resources, sports facilities, sports practice, and market potential of sports can be essential for sports organisations and governments to build effective strategic plans for sports development and to adapt their decisions to future goals. In this sense, the implementation of databases, with the use of structured information, through historical data sources, is essential to identify trends and patterns and collaborate with data analysis routines (Ren and Liu, 2021; Ward et al., 2019).

The database model design presented in this study identified the main actors involved with elite sport in Brazil - public and private organisations, athletes, coaches, multidisciplinary teams, managers in addition to identifying the pillars, factors and variables that interfere with success and sports development (De Bosscher et al., 2015; Green and Oakley, 2001; Houlihan and Green, 2008). The method that associated the deep study of scientific literature and the extraction of data from primary sources with the use of technology tools for information systematisation (Rajšp and Fister, 2020), collaborates with the construction of standardised repositories that can allow the permanent updating of data (Bunker and Thabtah, 2019), monitoring of quantitative indicators (Madella et al., 2005; O'Boyle, 2015) and, consequently, increased knowledge for the sports sector (Budovich, 2021; Caya and Bourdon, 2016).

Following the results presented by the Sports Policy Leading to International Sporting Success (SPLISS) consortium, it is almost a consensus among researchers that meson-level factors (De Bosscher et al., 2006), those that interfere with sports policies, are good predictors of elite sport development performance (De Bosscher et al., 2015). This result is explained by the fact that countries have become more in offering support to athletes, relying less on uncontrolled variables and focusing their efforts on management and governance (Houlihan and Green, 2008; Strittmatter et al., 2018). From this perspective, that the model designed in this study allows the elements linked to elite sport development, identified in previous studies (De Bosscher et al., 2015; Houlihan and Green, 2008), to be quantified and analysed in an integrated manner through computerised systems.

As Houlihan and Green (2008) argue, information is a fundamental input for knowing a sports system's reality and building development strategies. In the model presented, based on the Brazilian context, quantitative and integrated analysis can occur due to relationships between the tables and data recorded in each dimension. This approach permits data manipulation according to user needs (researcher, manager, policymaker) and detailed analysis of sports organisations (events, infrastructure, athletes, professionals and financial resources); on investment efficiency (results obtained from investments in each of the success factors); on sport facilities and other factors that interfere in sports success.

In the designed database, data from different sources were standardised and stratified to enable comparisons across sports, dimensions, and time series (Dabnichki et al., 2020). Such standardisation corroborates the principles of Business Intelligence, through which one can collect data from various sources, organise, analyse, and share them (Caya and Bourdon, 2016; Ward et al., 2019). Considering funding as the primary input for elite sport development (De Bosscher et al., 2015), the model allows the identification and study of the sources of financial resources, their destinations (people and organisations), and classifying the application of resources into 13 different types. Since the database built contemplates public and private investment sources, quantitative answers can be consulted about the funding sources (compiled or individual), the amounts invested per sport, per type of support, and chronologically and regionally.

A similar strategy for stratifying sports management data was used by Qiao (2022) to analyse the weaknesses of research and public management of the Chinese sports system. of the authors employed Big Data tools to examine the theoretical development of sport management by including public sport management data such as data from sport industry organisations, human sports resources, club and sports event management, and the scientific literature produced in that country in the framework and summarise the most significant scientific research issues.

The methods of building the database presented in our study also meet the "Framework for Understanding the Performance of National Governing Bodies of Sport (Bayle and Robinson, 2007), since the dimensions of the model and the categorisation of information favour monitoring organisational performance under the financial and non-financial perspectives (participation in events and results obtained; human resources available, scientific production in the field of sport; use of sports infrastructure). In other words, the database model can collaborate with the production of knowledge about funding intensity, the need for coach training, facility expansion, the number of revenue-generating events, access to equipment, and the cost of the sport (Bayle and Robinson, 2007; Ren and Liu, 2021; Sun et al., 2022).

Regarding the scope of data stored and related in the designed database, the eight dimensions and 55 tables described meet the nine pillars and 96 critical success factors identified in the studies by De Bosscher et al. (2015), as well as the notable dimensions for international sport success described by Houlihan and Green (2008). In this sense, the use of tools and methods from the field of data science can be seen as a complement to the methods used by previous studies, as it can collaborate with the collection and organisation of data from sports and high-performance sports, with the construction of analytical models to generate new insights that articulate the essential dimensions and success factors (Lopez, 2020; Qiao, 2022; Ward et al., 2019).

5. Conclusion

This study uses data science methods to design a relational database to store and analyse quantitative information that pertains to national sport development. Taking the literature on the dimensions and factors of sports success as an object of investigation, the design of the entity-relationship model required the identification of the primary sources of data of the sport system, whether they come from public or private entities, sport or not; and the determination of the procedures for extraction of data available in the digital environment and their complementation by research instruments, also digital. Database modelling methods allowed the standardised storage and logical categorisation of data for each model dimension, considering the specificity of each Olympic and Paralympic sport. Also, because of the modelling, the database allows informative and quantitative, chronological and territorial queries, considering one or several database elements, whether they are organisations, athletes, facilities, materials, sports results or funding, which can allow the production of information and analysis for organisations and sports management.

The model presented in this study can be implemented in an organization, a city, or even a country. The database acts as a manager and central repository for data collected from various sources. Public administration entities, sports organizations, and private companies generate data that are standardized by the database to then produce reports and sports development indicators, considering international sporting success factors. In this way, the model serves as an interface between data sources and the generation of reports, diagnostics, and indicators.

Utilising the methods adopted, we detected that the data from organisations, athletes and sports professionals are fundamental for the integration, in the database, of the other elements that interfere in the success of Brazilian elite sports. This result indicates that the relational database stores information on sports, organisations, people, investments, places, facilities, materials, events, sports results, and documents such as normative acts and scientific productions related to sports, sharing data among the tables without allowing redundancies. This approach can collaborate with the future development of methods for permanent data updates through coding and web scraping, facilitating constant investigation and monitoring of performance metrics of organisations and sports systems.

6. Limitations and future research

Despite the significant advances in the application of digital technologies for data analysis in the development of elite sports, this investigation highlights several limitations that need to be overcome. One of these limitations lies in the difficulty of standardizing data collection and analysis methodologies, a crucial element for building databases and generating new knowledge from data analysis. The lack of uniformity can result in inconsistencies in results and hinder comparisons across different studies or implementations, as also observed in this research. Moreover, the integration of multiple data sources, such as wearable sensors, artificial intelligence algorithms, and data visualization tools, although offering new opportunities, faces significant challenges related to interoperability and the accuracy of data captured in dynamic environments, as described in the database model proposed in this study.

Although technological tools provide valuable insights for the performance analysis of sports systems, their large-scale implementation may be constrained by the demands for robust infrastructure and specialized technical knowledge. This creates a significant barrier to the adoption of these technologies in sports environments with more limited resources, restricting the benefits to a smaller number of athletes and organizations that have access to such technologies.

For future research, the most promising directions include the development of data collection methods that involve technological tools and can be implemented at different levels of a sports system, within an organization, a city, or a country. Additionally, the adoption of data science models that integrate predictive and personalized analyses is also essential, allowing managers and policymakers to make informed decisions based on a combination of historical data as well as real-time information. In this regard, future research should focus on creating more robust frameworks that ensure interoperability between different technological platforms, facilitating the implementation of integrated approaches that can encompass various dimensions of sports development, as demonstrated in the model presented in this study.

Supplementary materials: A: Annex A presents Entities, dimensions, and database model relationships to analyse Brazilian elite sports.

Author contributions: Conceptualization, AdCDF and ARRF; methodology, AdCDF, PS and ARRF; software, AdCDF and PS; validation, AdCDF, ARRF and JBM; formal analysis, AdCDF and JBM; investigation, AdCDF and ARRF; resources, ARRF; data curation, PS; writing—original draft preparation, AdCDF; writing—review and editing, JBM, ARRF and AdCDF; supervision, ARRF; project administration, AdCDF. All authors have read and agreed to the published version of the manuscript.

Acknowledgments: We would like to thank Brazilian Ministry of Sports and “References Research project” members (Federal University of Rio Grande do Sul) for the data collection support.

Conflict of interest: The authors declare no conflict of interest.

References

Baca, A., Dabnichki, P., Hu, C.-W., et al. (2022). Ubiquitous Computing in Sports and Physical Activity—Recent Trends and Developments. Sensors, 22(21), 8370. https://doi.org/10.3390/s22218370

Bai, Z., & Bai, X. (2021). Sports Big Data: Management, Analysis, Applications, and Challenges. Complexity, 2021(1). Portico. https://doi.org/10.1155/2021/6676297

Bayle, E., & Robinson, L. (2007). A Framework for Understanding the Performance of National Governing Bodies of Sport. European Sport Management Quarterly, 7(3), 249–268. https://doi.org/10.1080/16184740701511037

Bohlmann, H. R., & Heerden, J. H. V. (2008). Predicting the economic impact of the 2010 FIFA World Cup on South Africa. International Journal of Sport Management and Marketing, 3(4), 383. https://doi.org/10.1504/ijsmm.2008.017214

Budovich, L. S. (2021). Business management of the sport industry by considering the digitalization. Journal of Human Sport and Exercise. https://doi.org/10.14198/jhse.2021.16.proc4.07

Bunker, R. P., & Thabtah, F. (2019). A machine learning framework for sport result prediction. Applied Computing and Informatics, 15(1), 27–33. https://doi.org/10.1016/j.aci.2017.09.005

Camps, A., & Pappous, A. S. (2016). Predicting the Evolution of Sports Federation Membership: AnImportant Tool to Asses National Governing Bodies' Strategic Planning. Journal of Sports Science, 4(2). https://doi.org/10.17265/2332-7839/2016.02.001

Caya, O., & Bourdon, A. (2016). A Framework of Value Creation from Business Intelligence and Analytics in Competitive Sports. In: Proceedings of the 2016 49th Hawaii International Conference on System Sciences (HICSS). https://doi.org/10.1109/hicss.2016.136

Codd, E. F. (1970). A relational model of data for large shared data banks. Communications of the ACM, 13(6), 377–387. https://doi.org/10.1145/362384.362685

Cossich, V. R. A., Carlgren, D., Holash, R. J., et al. (2023). Technological Breakthroughs in Sport: Current Practice and Future Potential of Artificial Intelligence, Virtual Reality, Augmented Reality, and Modern Data Visualization in Performance Analysis. Applied Sciences, 13(23), 12965. https://doi.org/10.3390/app132312965

Dabnichki, P., Hu, C.-W., Kornfeind, P., & Exel, J. (2022). Ubiquitous computing in sports and physical activity—Recent trends and developments. Sensors, 22(21), 8370. https://doi.org/10.3390/s22218370

De Bosscher, V., De Knop, P., Van Bottenburg, M., et al. (2006). A Conceptual Framework for Analysing Sports Policy Factors Leading to International Sporting Success. European Sport Management Quarterly, 6(2), 185–215. https://doi.org/10.1080/16184740600955087

De Bosscher, V., Shibli, S., Westerbeek, H., & Van Bottenburg, M. (2015). Successful elite sport policies: an international comparison of the sports policy factors leading to international sporting success (SPLISS 2.0) in 15 nations. Meyer & Meyer.

De Bosscher, V., Shibli, S., Westerbeek, H., et al. (2016). Convergence and Divergence of Elite Sport Policies: Is There a One-Size-Fits-All Model to Develop International Sporting Success? Journal of Global Sport Management, 1(3–4), 70–89. https://doi.org/10.1080/24704067.2016.1237203

Exel, J., & Dabnichki, P. (2024). Precision Sports Science: What Is Next for Data Analytics for Athlete Performance and Well-Being Optimization? Applied Sciences, 14(8), 3361. https://doi.org/10.3390/app14083361

Ferreira, A. D. C. D., Vitor, K. P., Haiachi, M. D. C., et al. (2018). Financing paralympic sport in Brazil: agreements (Portugal). Cadernos de Educação Tecnologia e Sociedade, 11(1), 22. https://doi.org/10.14571/brajets.v11.n1.22-36

Green, M., & Oakley, B. (2001). Elite sport development systems and playing to win: uniformity and diversity in international approaches. Leisure Studies, 20(4), 247–267. https://doi.org/10.1080/02614360110103598

Grix, J. (2013). Sport Politics and the Olympics. Political Studies Review, 11(1), 15–25. https://doi.org/10.1111/1478-9302.12001

Houlihan, B., & Green, M. (2008). Comparative Elite Sport Policy: Systems, Structures, and Public Policy. 1st ed. Elsevier, pp. 316. https://doi.org/10.4324/9780080554426

International Olympic Committee. (n.d.). Available online: Https://Olympics.Com/Ioc/Documents/Olympic-Games/Rio-2016-Olympic-Games (accessed on 2 June 2023).

Kroenke, D. M., & Auer, J. D. (2007). Database Concepts, 3rd ed. Prentice.

Link, D. (2018). Sports Analytics. German Journal of Exercise and Sport Research, 48(1), 13–25. https://doi.org/10.1007/s12662-017-0487-7

Lopez, M. J. (2020). Bigger data, better questions, and a return to fourth down behavior: an introduction to a special issue on tracking datain the National football League. Journal of Quantitative Analysis in Sports, 16(2), 73–79. https://doi.org/10.1515/jqas-2020-0057

Madella, A., Bayle, E., & Tome, J. (2005). The organisational performance of national swimming federations in Mediterranean countries: A comparative approach. European Journal of Sport Science, 5(4), 207–220. Portico. https://doi.org/10.1080/17461390500344644

Mamo, Y. Z. (2023). Big Data and Innovative Research Methods. International Journal of Sport Communication, 16(3), 352–360. https://doi.org/10.1123/ijsc.2023-0109

O'Boyle, I. (2015). Developing a performance management framework for a national sport organisation. Sport Management Review, 18(2), 308–316. https://doi.org/10.1016/j.smr.2014.06.006

Ofoghi, B., Zeleznikow, J., MacMahon, C., et al. (2013). Data Mining in Elite Sports: A Review and a Framework. Measurement in Physical Education and Exercise Science, 17(3), 171–186. https://doi.org/10.1080/1091367x.2013.805137

Qiao, J. (2022). Intelligent Big Data Framework for the Technical Design of Public Management Applications in Sports. Mathematical Problems in Engineering, 2022, 1–9. https://doi.org/10.1155/2022/1900548

Rajšp, A., & Fister, I. (2020). A Systematic Literature Review of Intelligent Data Analysis Methods for Smart Sport Training. Applied Sciences, 10(9), 3013. https://doi.org/10.3390/app10093013

Ren, P., & Liu, Z. (2021). Efficiency Evaluation of China's Public Sports Services: A Three-Stage DEA Model. International Journal of Environmental Research and Public Health, 18(20), 10597. https://doi.org/10.3390/ijerph182010597

Rewilak, J. (2021). The (non) determinants of Olympic success. Journal of Sports Economics, 22(5), 546–570. https://doi.org/10.1177/1527002521992833

Sands, W. A., Kavanaugh, A. A., Murray, S. R., et al. (2017). Modern Techniques and Technologies Applied to Training and Performance Monitoring. International Journal of Sports Physiology and Performance, 12(s2), S2-63-S2-72. https://doi.org/10.1123/ijspp.2016-0405

SAP. (n.d.) Unveils SAP sports one solution for soccer. Available online: http://News.Sap.Com/Sap-Unveils-Sap-Sports-One-Solution-for-Soccer/ (accessed on 2 June 2023).

Shilbury, D., & Moore, K. A. (2006). A Study of Organizational Effectiveness for National Olympic Sporting Organizations. Nonprofit and Voluntary Sector Quarterly, 35(1), 5–38. https://doi.org/10.1177/0899764005279512

Shin, S.-S. (2022). Teaching Method for Entity–Relationship Models Based on Semantic Network Theory. IEEE Access, 10, 94908–94923. https://doi.org/10.1109/access.2022.3206028

Strittmatter, A.-M., Stenling, C., Fahlén, J., et al. (2018). Sport policy analysis revisited: the sport policy process as an interlinked chain of legitimating acts. International Journal of Sport Policy and Politics, 10(4), 621–635. https://doi.org/10.1080/19406940.2018.1522657

Sun, G., Zhang, X., & Lin, Y. (2022). Evaluation Model of Sports Culture Industry Competitiveness Based on Fuzzy Analysis Algorithm. Mathematical Problems in Engineering, 2022, 1–8. https://doi.org/10.1155/2022/9708646

Wanless, L. (2022). Progressive Analytics Diffusions: Rewiring our Software. Journal of Applied Sport Management. https://doi.org/10.7290/jasm14mxjx

Ward, P., Windt, J., & Kempton, T. (2019). Business Intelligence: How Sport Scientists Can Support Organization Decision Making in Professional Sport. International Journal of Sports Physiology and Performance, 14(4), 544–546. https://doi.org/10.1123/ijspp.2018-0903

Watanabe, N. M., Shapiro, S., & Drayer, J. (2021). Big Data and Analytics in Sport Management. Journal of Sport Management, 35(3), 197–202. https://doi.org/10.1123/jsm.2021-0067

Winand, M., Vos, S., Claessens, M., et al. (2014). A unified model of non-profit sport organizations performance: perspectives from the literature. Managing Leisure, 19(2), 121–150. https://doi.org/10.1080/13606719.2013.859460

Appendix

Table A1. Names for all variables included in the entity-relationship model.

Dimensions	Variables names	Attributes, primary keys, foreign keys
(1) Funding	Funding identification number; name of funding source; year; start and end date; amount; type of funding; funding destination; funded sport; city	**fund_id**, fund_value, fund_dateInitial, fund_dateFinal, fund_year, sport_id (fk), city_id (fk)
(2) Legislation	Legislation identification number; Name of legislation; legislation number; type and level of coverage; date; summary; main subject; legislation status; location data (city, state, region)	**legi_id,** legi_date, legi_summary, legi_number, legi_link**,** city_id (fk),
(3) Organisations	Organization identification number; name of entity; VAT number; location data (full address, latitude and longitude); email address; website; classification and type of entity, sports managed by the organization	**enti_id**, enti_vat_number, enti_pcode, enti_adress, enti_longitude, enti_latitude, enti_email, enti_site, city_id (fk), sport_id (fk)
(4) Athletes and Sports Professionals	Athlete identification number; athlete's name; gender; date of birth; location data (full address); age category; athlete's sport and event; name of the athlete's confederation, federation and club; type of disability and functional class (if Paralympic athlete); professional identification number; professional's name; date of birth; gender; profession; location data (full address).	**athle_id**, athle_name, athle_birth_date, sport_id (fk), city_id (fk), athle_pcode, athle_adress, sex_id (fk), age_group_id (fk), func_class_id (fk)*, deftype_id (fk)*, **prof_id**, prof_name, prof_ birth_date, sex_id (fk), city_id, prof_pcode, prof_adress
(5) Events and Sports Results	Sporting event identification number; event name; event year; event date; event city; event sports; event result type identifier; event age category; event functional class (if Paralympic event); event athletes; event results.	**eves_id**, eves_name, eves_year, eves_date, city_id (fk), sport_id (fk), **ress_id**, sport_id (fk), event_id (fk), age_group_id (fk), func_class_id (fk)* ress_classif, ress_result
(6) Sports Facilities	Sports facility identifier number; facility name; location data (full address, latitude and longitude); facility type; sports practiced.	**spfa_id**, spfa_name, city_id (fk), spfa_pcode, spfa_adress, spfa_longitude, spfa_latitude, sport_id (fk)
(7) Sports Materials	Sports equipment identifier number; equipment name; material brand; material characteristics; sports using the equipment.	**material_id**, spmaterial_key, spmaterial_name, spmaterial_brand, spmaterial_feat, sport_id (fk)
(8) Science and Technology	Study, research group or laboratory identifier number; organization name; study year; study author; main study subject; laboratory name; research group name; research group main subject.	**stud_id**, stud_year, stud_author, sport_id (fk), **lab_id**, lab_name, **rgroup_id,** rgroup_name, **conf_id,** conf_name