BadmintONTO: A Badminton Domain Ontology

Huanyu Li[1,2,*], Patrick Lambrix[1,2]

[1]Department of Computer and Information Science, Linköping University, Linköping, Sweden
[2]Swedish e-Science Research Centre, Linköping, Sweden

Abstract

In different sports fields, collecting play-by-play data has become significant for data analysis, as seen in sports like baseball, basketball, ice hockey and football. Badminton remains a relatively new domain in terms of systematically collecting game play data (i.e., play-by-play or shot-by-shot) with only some recently published datasets online. While the game play data provides some detailed information about matches, it lacks rich semantics for complex information retrieval and data analysis. Consequently, the data cannot be used for applications where semantics are needed. This paper introduces a badminton domain ontology - BadmintONTO, along with its illustrative usages to showcase the capability of this ontology to represent basic domain knowledge and to annotate play-by-play data.

Keywords

Ontology, Badminton Domain, Badminton Game Play Data, Knowledge Representation

1. Introduction

Sports analytics involves using data from sports events to gain insights about a sport and its surroundings. The advancements in computer science (e.g., in machine learning, deep learning, computer vision) have significantly facilitated the development of many methods and tools for sports analytics. Video understanding techniques have been used in assistant referee systems (e.g., video assistant referees (VAR) in football, video review in ice hockey, instant replay in baseball and basketball, and the Hawk-Eye system in badminton and tennis). In various sports, there is a growing focus on interpreting play-by-play data generated through video understanding techniques. Play-by-play data holds value for analyzing player performance or investigating techniques and tactics in different sports such as prior studies for basketball [1], baseball [2], football [3], ice hockey [4], and badminton [5]. Unlike the baseball, basketball, ice hockey and football fields which have provided and utilized play-by-play data for quite a long time, the badminton field recently obtained play-by-play datasets, BadmintonDB [6] (with 9 single matches), and ShuttleSet [7] (with 44 single matches) which were published in 2022 and 2023, respectively. ShuttleSet models more information than BadmintonDB such as locations of players and shuttles in terms of court areas, and hitting styles (i.e., forehand or backhand). Similar to the challenges faced in ice hockey [8], the badminton domain encounters

FOIS 2024 Ontology Showcase
*Corresponding author.
huanyu.li@liu.se (H. Li); patrick.lambrix@liu.se (P. Lambrix)
http://huanyuli.se (H. Li); https://www.ida.liu.se/~patla00/ (P. Lambrix)
0000-0003-1881-3969 (H. Li); 0000-0002-9084-0470 (P. Lambrix)

difficulties in leveraging play-by-play data for complex query requirements. Ontologies can alleviate this problem as they provide benefit to information retrieval, data integration, and FAIR (Findable, Accessible, Interoperable and Reusable) [9] data sharing, by formally representing domain knowledge. However, there is a lack of ontology-related work focusing on badminton.

To address the aforementioned challenges, we develop the BadmintON onTOlogy (BadmintONTO). In Section 2, we present related work. Then in Section 3, we present the development and content of BadmintONTO. BadmintONTO is developed and maintained in a GitHub repository,[1] with a published SPARQL server.[2] Moreover, we publish the ontology by employing a permanent URI[3] as an identifier through the w3id service. Section 4 outlines the example usages and evaluation of BadmintONTO. Finally, we discuss the result and outline several directions for future work in Section 5, and conclude in Section 6.

2. Related Work

This section introduces relevant sports domain ontologies and general ontologies.

2.1. Sports Domain Ontologies

Several existing works focus on modeling domain knowledge in sports using ontologies. For instance, the International Press Telecommunications Council (IPTC) established a standard for exchanging news data known as SportsML.[4] Additionally, IPTC developed several domain-specific ontologies for various sports (e.g., football, baseball and volleyball) as well as a general ontology called Sport Schema.[5] However, these domain ontologies primarily concentrate on representing sports statistics and lack concepts or relationships to represent domain knowledge, particularly at the level of techniques and tactics. Sport Schema aims to model certain concepts and relationships applicable to various sports for data sharing purposes, such as event and participation concepts. Similarly, the British Broadcasting Corporation (BBC) developed a sport ontology[6] focusing on representing data related to news.

There is academic work focusing on developing ontologies for various sports, including ice hockey, tennis, football and baseball. For example, the ice hockey ontology introduced in [8] aims to interpret play-by-play data. In [10], a shot taxonomy for tennis is presented, derived from an analysis over trajectory data provided by the Hawk-Eye system which considers information such as the ball's shape and trajectory and speed of balls. Additionally, in the domains of football [11] and baseball [12], there is work focusing on commentary generation or match summarization, which involve natural language text generation tasks based on ontologies. For basketball, a generic and unified classification with semantic annotations of basketball-related terms is proposed in [13].

[1] BadmintONTO GitHub repository: https://github.com/huanyu-li/BadmintONTO
[2] SPARQL server showcase: https://huanyu-li.github.io/BadmintONTO/demo/
[3] BadmintONTO Identifier: http://w3id.org/BadmintONTO
[4] SportsML: https://iptc.org/standards/sportsml-g2/
[5] Sport Schema: https://sportschema.org
[6] BBC sport ontology: https://www.bbc.co.uk/ontologies/sport-ontology/

2.2. General Ontologies for Sports Events Modeling

Other relevant work is about ontologies focusing on modeling event-related semantics and spatio-temporal extent. In [14], a survey on event-related ontologies is conducted based on six dimensions of events, which are *What* (changes and actions of events), *Where* (possible places of events), *Who* (active/passive entities of events), *Why* and *How* (possible reasons of events) [15]. In addition, an event pattern is used in both [16] and [17] where concepts and relationships such as roles and spatio-temporal extent are modeled. Similarly, the Event Ontology[7] defines relationships between events and temporal, spatial entities, as well as a relationship between events and agents. An event processing ontology design pattern is proposed in [18], in which *EventObject* is a sub-concept of *InformationObject* from the DOLCE ontology [19]. Event objects can be simple and complex event objects. There is a relationship, *InformationAbout* between *EventObject* and *Event*. In representing temporal extent, the OWL-Time ontology[8] offers models for relevant temporal concepts and their relationships. For spatial information representation, GeoSPARQL [20], established as a standard query language for geospatial data in the Semantic Web by the Open Geospatial Consortium (OGC)[9], provides an ontological model. Another initiative, GeoLink [17] aiming at interlinking geospatial data across diverse datasets and domains, also provides an ontology for representing geospatial data related to events.

3. Ontology Development

We employed the NeOn ontology engineering methodology [21] to develop BadmintONTO. Specifically, our focus was on requirements analysis and the reuse and re-engineering of ontological resources according to the NeOn methodology. While other methodologies for ontology development such as the On-To-Knowledge (OTKM) methodology [22] and eXtreme Design (XD) [23] exist, we chose to follow the NeOn methodology as it permits us to also take future possible scenarios (e.g., the scenario about ontology alignment) into account.

3.1. Requirements Analysis

To develop an ontology for the badminton domain, we conducted a study on: **(1)** existing ontologies related to sports and general domains as mentioned in Section 2, **(2)** state-of-the-art data analysis applications in badminton and then discussed with domain experts in terms of domain interests of information retrieval and data analysis. The domain experts include the first author and an expert with coaching experience. Basically, the requirements analysis includes use case identification and competency question identification.

Use Case. BadmintONTO aims to capture semantics that model the details of badminton matches, enabling the annotation of play-by-play (i.e., shot-by-shot) data, not only of match statistics but also techniques and tactics performed by players. Additionally, it should facilitate semantics-aware searching over play-by-play data.

[7]The Event Ontology: https://motools.sourceforge.net/event/event.html
[8]OWL-Time ontology: https://www.w3.org/TR/owl-time/
[9]Open Geospatial Consortium: https://www.ogc.org

Competency Questions. After discussions between domain experts and ontology engineers, we formulated competency questions that BadmintONTO should be able to address.

- **CQ1**: What are the contextual details of a competition or match, such as time, location, and match type?
- **CQ2**: How many matches reach to a third set in a tournament competition?
- **CQ3**: Which techniques (shots) does a player/pair primarily use in a match?
- **CQ4**: (On average), how many shots does a player/pair play to win a point in a match?
- **CQ5**: What are the reasons for a player/pair getting or losing a point?
- **CQ6**: What is the number of shots for the longest rally in a match?
- **CQ7**: How many rallies end within the first three shots in a match?
- **CQ8**: Are there any patterns of shots leading to a player/pair getting points in matches?
- **CQ9**: Who are the winners of competitions (e.g., in different match types)?
- **CQ10**: In doubles matches, what are the stance positions of the winning pair of players (parallel stance or front-back stance) for all winning points?

3.2. Development and Implementation Process

The NeOn methodology outlines two essential activities: conceptualization and formalization, which are based on specifications derived from requirement analysis. These two activities also involve considering existing ontological resources, such as ontologies or patterns. Therefore, we extract key terms from the requirements analysis results, which could represent concepts and relationships. Based on these key terms, we explore existing resources for potential reuse or re-engineering, as discussed in Section 2.

Our ontology implementation process involves several steps. We implement the ontology using Protégé,[10] hosting it on a GitHub repository,[1] and creating documentation using tools, pyLODE[11] and WebVOWL.[12] To evaluate the ontology, we follow a multi-step approach. Firstly, we write RML (RDF Mapping Language) mappings [24] to construct an RDF dataset based on BadmintONTO and an example match from ShuttleSet. Secondly, we utilize an Apache Jena Fuseki-based SPARQL server[13] to execute and test SPARQL queries over this RDF dataset. Additionally, we employ ontology pitfall scanners and validators such as OntOlogy Pitfall Scanner (OOPS!) [25] and OOPS! for FAIR (FOOPS!) [26] to identify and address potential issues. Further details are elaborated in Section 4. It is worth noting that our development and implementation process follows an iterative-incremental model. We prioritize writing RML mappings to annotate play-by-play data and conduct OOPS! and FOOPS! checks once we achieve a relatively stable ontology implementation. This approach enables us to identify and address potential issues promptly. In the subsequent sections, we delve into more detailed discussions about conceptualization and formalization.

[10] Protégé: http://protege.stanford.edu
[11] pyLODE: https://github.com/RDFLib/pyLODE
[12] WebVOWL: https://github.com/VisualDataWeb/WebVOWL
[13] Apache Jena Fuseki: https://jena.apache.org/documentation/fuseki2/

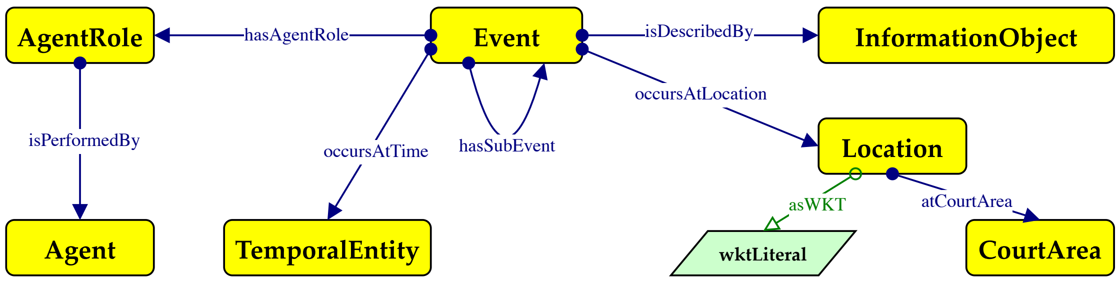


Figure 1: The core concepts and relationships of BadmintONTO.

3.3. General Conceptualization and Formalization

Based on the results of the requirements analysis, we identify some general key concepts and relationships to be modeled in BadmintONTO as shown in Figure 1. They are relevant to general events. We introduce formulation details as follows.

Event and Agent. An event typically involves participating agents who play specific roles (e.g., players or umpires in a sport event). At this stage, we focus on modeling the roles of participants in events, such as participating players and teams. We specify in the ontology that each event has some participating agents through the roles performed by agents (Axioms 1 and 2). This approach to modeling roles and participating agents is well discussed and used as shown in [16] and other design patterns.[14] In addition, we follow the approach in [16] for modeling a sub-event relationship, which is a transitive relationship with event as both the domain and range (Axioms 3, 4 and 5).

$$Event \sqsubseteq \exists hasAgentRole.AgentRole \sqcap \forall hasAgentRole.AgentRole \quad (1)$$

$$AgentRole \sqsubseteq \exists isPerformedBy.Agent \sqcap \forall isPerformedBy.Agent \quad (2)$$

$$hasSubEvent \circ hasSubEvent \sqsubseteq hasSubEvent \quad (3)$$

$$\top \sqsubseteq \forall hasSubEvent.Event \quad (4)$$

$$\exists hasSubEvent.\top \sqsubseteq Event \quad (5)$$

Event and SpatioTemporalExtent. Another modeling concern of BadmintONTO is representing events and their temporal and spatial information. We utilized the OWL-Time ontology[8] for this purpose. In detail, a **TemporalEntity** in OWL-Time is specified by **Instant** and **Interval**, while the latter can have time instants inside. In addition, OWL-Time provides a vocabulary to describe time, temporal duration, etc. In our ontology, we connect **Event** and **TemporalEntity** through the **occursAtTime** relationship (Axiom 6). To represent spatial information, we draw inspiration from GeoSPARQL [20] and GeoLink [17]. For instance, we specify that each event takes place at certain locations (Axiom 7), with each location being a sub-concept of **Geometry** defined in GeoSPARQL (Axiom 8). Moreover, considering the need to express

[14] E.g., http://ontologydesignpatterns.org/wiki/Submissions:Objectrole, http://ontologydesignpatterns.org/wiki/Submissions:ParticipantRole

positions of players and shuttles in terms of their corresponding court areas (e.g., each court can be divided into 16 areas [7]), we specify that a location can also be expressed in terms of **CourtArea** (Axiom 9). Events can take place at multiple locations and times. This arises from the possibility of an event being suspended and conclude at different times and locations.

$$Event \sqsubseteq \exists occursAtTime.TemporalEntity \sqcap \forall occursAtTime.TemporalEntity \quad (6)$$

$$Event \sqsubseteq \exists occursAtLocation.Location \sqcap \forall occursAtLocation.Location \quad (7)$$

$$Location \sqsubseteq Geometry \quad (8)$$

$$Location \sqsubseteq \exists atCourtArea.CourtArea \sqcap \forall atCourtArea.CourtArea \quad (9)$$

Contextual Information of Events. To represent contextual information of events, such as scores of sets or matches and reasons for winning or losing rallies, we adopt established practices as presented in [16, 17], utilizing the **InformationObject** based on DOLCE [19]. Additionally, the **InformationObject** captures information about event entities and other related entities, such as the winning player of a match event or a specific shot type of a shot.

$$Event \sqsubseteq \exists isDescribedBy.InformationObject \sqcap \forall isDescribedBy.InformationObject \quad (10)$$

$$\top \sqsubseteq \forall isDescribedBy.InformationObject \quad (11)$$

$$\exists isDescribedBy.\top \sqsubseteq Event \quad (12)$$

3.4. Conceptualization for Other Badminton Domain Knowledge

Figure 2 illustrates additional concepts essential for annotating badminton events. Specifically, we define a taxonomy of **Event**, which encompasses sub-concepts such as **CompetitionEvent**, **MatchEvent**, **SetEvent**, **RallyEvent**, and **ShotEvent**. Accordingly, we establish a taxonomy of **InformationObject** with sub-concepts **CompetitionInformation**, **MatchInformation**, **SetInformation**, **RallyInformation**, and **ShotInformation**. These sub-concepts intend to capture information such as scores, locations, and participating players as shown in Figure 3, all of which are covered in the play-by-play data. Specific events occurring during a match, such as challenges to the shuttle's landing position and penalty judgments, are modeled within BadmintONTO for potential future utilization despite not being currently represented in the badminton play-by-play data (i.e., ShuttleSet).

We model different instances of roles as outlined in Figure 2 (e.g., **SingleRoleA** and **SingleRoleB**) to distinguish the two players/pairs/teams in a game. While in some contexts, these roles may be interpreted as *home role* or *opponent role*, we opted for a more generic approach, modeling them as *Role A* and *Role B*. These roles persist throughout an entire match. Additionally, two specific design decisions regarding role modeling have been made in BadmintONTO, which may be subject to future updates. The first decision involves only modeling players' role information at the **SetEvent** level for ease of annotating the play-by-play data. Future work could explore extending these roles to all event levels or limiting them to the **MatchEvent**. The second decision pertains to modeling score information of two agents. Currently, we define the scores' sources in terms of agents playing *Role A* and *Role B* through relationships (**matchScoreA**, **matchScoreB**, **setScoreA**, **setScoreB**, **roundScoreA**, and **roundScoreB**).

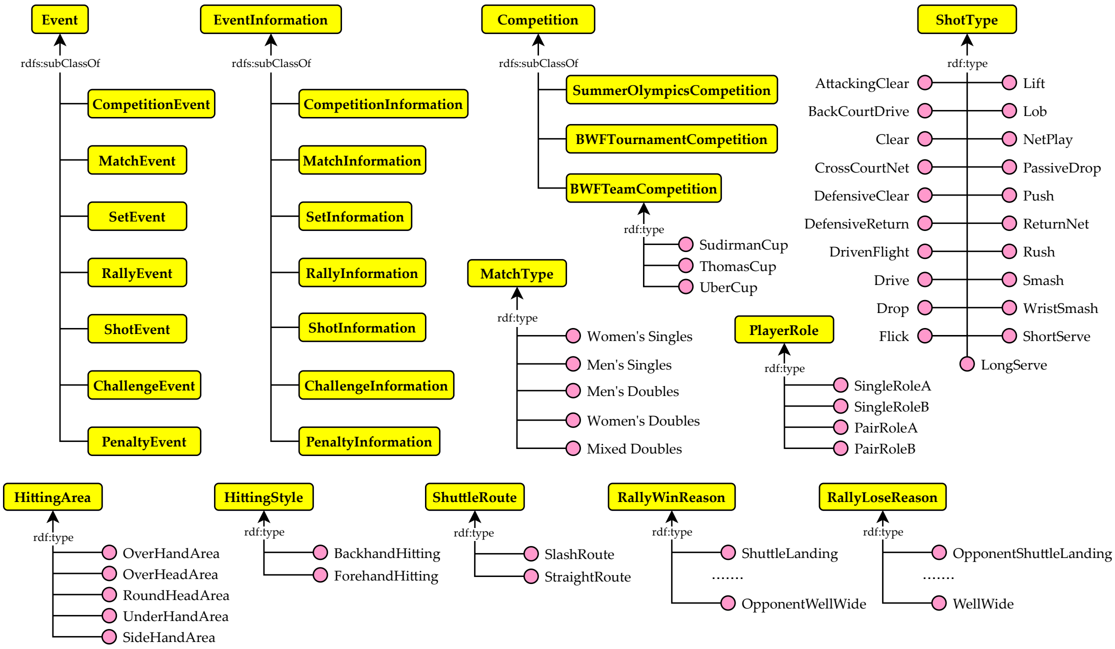


Figure 2: Specific badminton domain-related concept hierarchy.

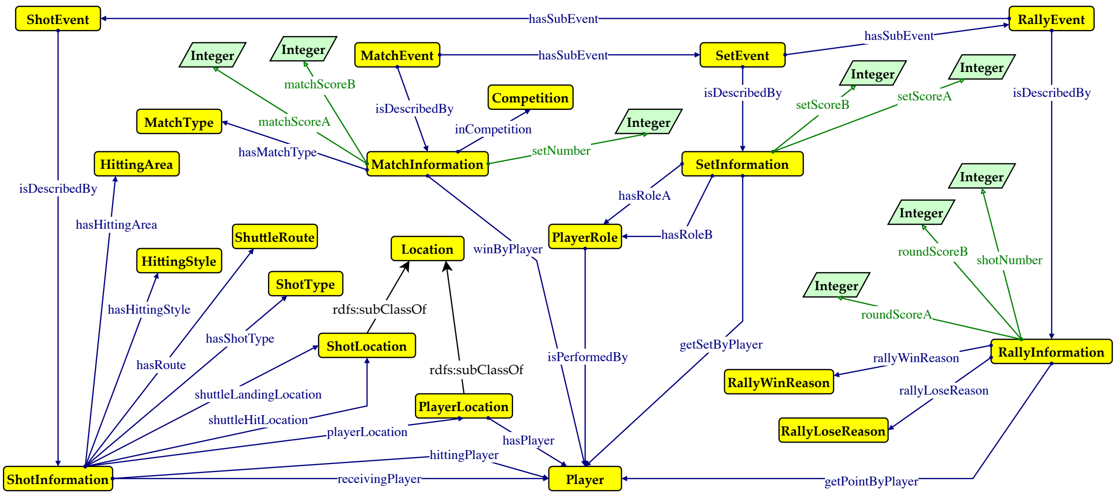


Figure 3: Event and Information Details.

Furthermore, the ontology incorporates a concept of competition types, **Competition**, with three sub-concepts, each representing specific competition types. For instance, the *All England Open* is defined as a **BWFTournamentCompetition**, occurring annually, with a specific edition linked to an instance of **CompetitionEvent**. Additionally, considering BadmintONTO's objective of annotating play-by-play data with a focus on observed techniques and tactics in matches, the representation of shot types is crucial. Therefore, we introduce the **ShotType** concept with instances to represent specific shot types (e.g., serve shot and smash shot). As interests of the badminton domain, we also introduce the **HittingArea**, **HittingStyle**, **ShuttleRoute**, **RallyWinReason**, **RallyLoseReason** concepts with instances.

4. Usage and Evaluation

This section introduces the usage and evaluation of BadmintONTO.

4.1. Annotating Shot-by-Shot Data

Between ShuttleSet [7] and BadmintONTO [6], we utilized the former to construct an RDF dataset based on BadmintONTO, employing RML [24]. ShuttleSet, a human-annotated dataset, encompasses 44 BWF (Badminton World Federation) tournament matches between 2018 and 2021, involving 27 top-ranking single male and female players. The dataset comprises 104 sets, 3,685 rallies and 36,492 shots across the 44 matches. As mentioned in Section 1, ShuttleSet captures more details from games than BadmintonDB such as locations in terms of court areas.

RML mappings were created across four levels corresponding to four event types: **MatchEvent**, **SetEvent**, **RallyEvent**, and **ShotEvent**. These mappings encompassed participating players, spatio-temporal extent, and contextual information (e.g., shot type, hitting location, point acquisition reasons, scores) for the mentioned events, resulting in the generation of an RDF dataset. Additionally, we publish a SPARQL server[2] to answer and test SPARQL queries over this RDF dataset (further discussed in the next Section).

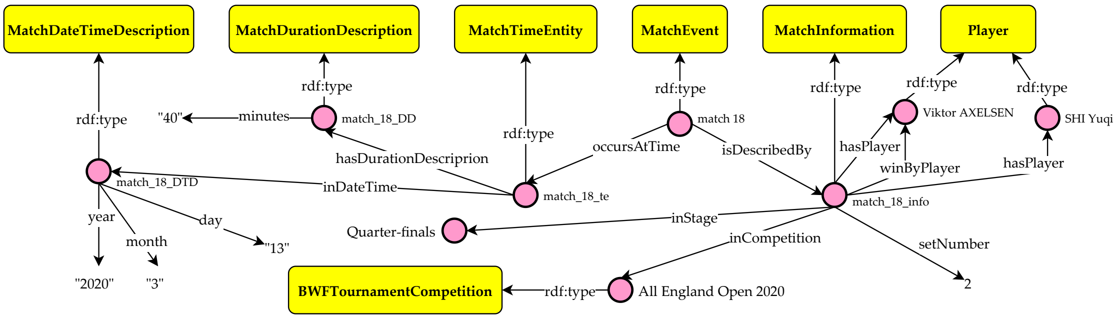
An instantiation example related to a match event is illustrated in Figure 4a, showcasing temporal annotations (e.g., **MatchDateTimeDescription**, **MatchDurationDescription**, and **MatchTimeEntity**). Notably, our focus on representing temporal information primarily pertains to the match event level due to the domain's emphasis on temporal aspects at this level and the current version of the ShuttleSet dataset containing temporal information exclusively at the match event level. Figure 4b demonstrates an example of annotating shot-by-shot data at the set event level, where a set event serves as a sub-event of a match event and involves participating players performing two different roles.

In Figure 4c, shot-by-shot data is annotated at the rally event and shot event levels. A rally event is characterized by information such as shot count, round scores, win reason and the player getting the point. Additionally, a rally event includes at least one shot event (i.e., a serving shot) as its sub-event. Incorporating location information is important for describing a shot, enabling the analysis of shot trajectories. Thus, shot information encompasses details such as the shuttle's hitting and landing locations.

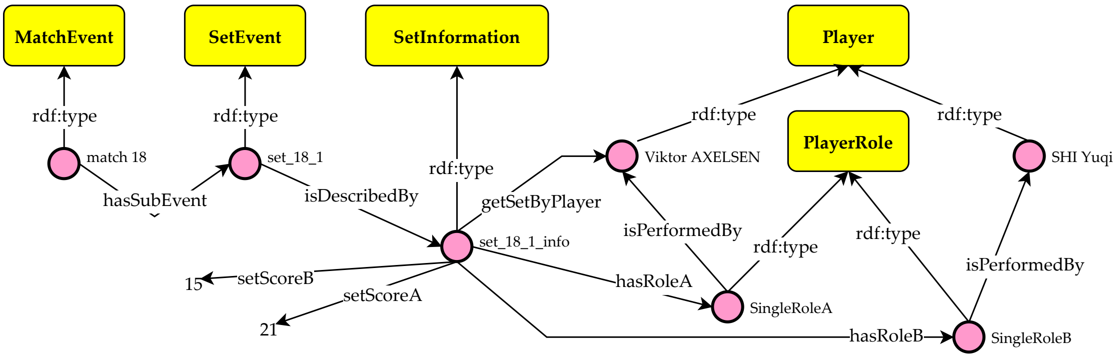
4.2. SPARQL Query Examples

An intended usage of BadmintONTO is to facilitate semantics-aware searching over play-by-play data, aligning with the outlined use case. Following the generation of an RDF dataset, we built a SPARQL server[2] and tested several SPARQL queries.[15]

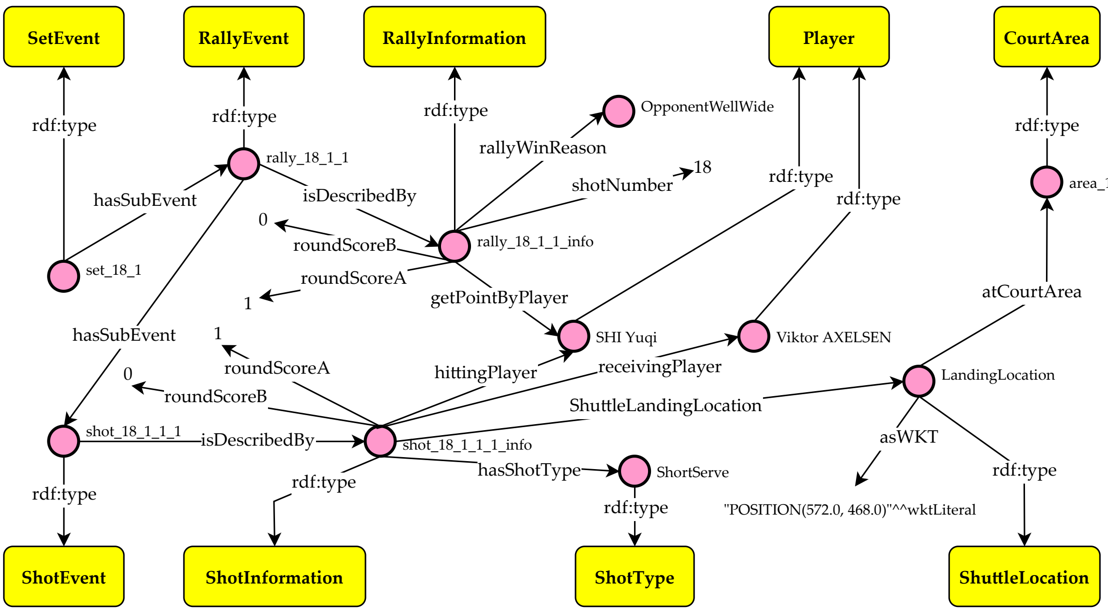
[15] All SPARQL queries are available at https://github.com/huanyu-li/BadmintONTO/blob/main/sparql_query/



(a) **MatchEvent** Instantiation.



(b) **SetEvent** Instantiation.



(c) **RallyEvent** and **ShotEvent** Instantiation.

Figure 4: Events Instantiation.

The SPARQL queries are written in accordance with competency questions outlined in Section 3. Table 1 shows the concepts and relationships from BadmintONTO necessary for writing the example SPARQL queries. Notice that currently, these queries only utilize a subset of BadmintONTO (10 concepts and 16 relationships). Future work will explore additional domain-specific queries and assess their compatibility with BadmintONTO. In Listings 1 and 2, we provide example SPARQL queries corresponding to CQ3 and CQ7, respectively. For each competency question (CQ1-CQ7), we formulated one or more SPARQL queries. However, we currently cannot write SPARQL queries covering CQ8, CQ9, and CQ10. The unavailability of queries for CQ8 is attributed to the absence of direct event sequence modeling in the current version of BadmintONTO, which was left for future work. For CQ9, the limited size of the ShuttleSet data, containing only 44 matches across various tournament competitions, does not have examples featuring all matches from a specific competition. The test of such competition-level contextual information is expected to be a future work. Regarding CQ10, the current ShuttleSet data lacks double and mixed double matches.

CQ	relevant concepts (10)	relevant relationships (16)
CQ1	MatchEvent, MatchInformation, MatchTimeEntity, MatchDateTimeDescription, MatchDurationDescription	isDescribedBy, inCompetition, hasMatchType, occursAtTime, inDateTime, year, month, day, hasDurationDescriprion, minutes
CQ2	MatchEvent, MatchInformation	isDescribedBy, setNumber
CQ3	ShotEvent, ShotInformation, Player	isDescribedBy, hasShotType, hittingPlayer, personFullName
CQ4	RallyEvent, RallyInformation, Player	isDescribedBy, getPointByPlayer, personFullName, shotNumber
CQ5	RallyEvent, ShotEvent, RallyInformation, ShotInformation, Player	isDescribedBy, getPointByPlayer, hasShotType, hittingPlayer
CQ6	RallyEvent, RallyInformation, Player	isDescribedBy, getPointByPlayer, personFullName, shotNumber
CQ7	RallyEvent, RallyInformation, Player	isDescribedBy, getPointByPlayer, personFullName, shotNumber

Table 1
Coverage of concepts and relationships in CQs.

Listing 1: An example SPARQL query of CQ3 (What are the quantities of different types of shots that SHI Yuqi hits during matches?).

```
PREFIX badmintonto: <http://w3id.org/BadmintONTO/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

SELECT ?shottype (COUNT(?shottype) as ?num)
WHERE {
    ?shotevent  rdf:type   badmintonto:ShotEvent .
    ?shotevent  badmintonto:isDescribedBy   ?shotinfo .
    ?shotinfo   badmintonto:hasShotType   ?shottype .
    ?shotinfo   badmintonto:hittingPlayer   ?player .
    ?player   badmintonto:personFullName   "SHI Yuqi" .
}
GROUP BY ?shottype
```

Listing 2: An example SPARQL query of CQ7 (How many rallies end in the serving shots?).

```
PREFIX badmintonto: <http://w3id.org/BadmintONTO/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

SELECT ?player (COUNT(?rallyevent) as ?num)
WHERE {
    ?rallyevent  rdf:type  badmintonto:RallyEvent .
    ?rallyevent  badmintonto:isDescribedBy  ?rallyinfo .
    ?rallyinfo  badmintonto:getPointByPlayer  ?player .
    ?rallyinfo  badmintonto:shotNumber  ?shotnum .
    FILTER (?shotnum = 1)
}
```

4.3. Evaluation based on OOPS! and FOOPS!

We have demonstrated BadmintONTO's ability in annotating badminton shot-by-shot data and enabling complex query answering over an RDF dataset generated based on BadmintONTO. Furthermore, we employed the OntOlogy Pitfall Scanner (OOPS!) [25] and OOPS! for FAIR (FOOPS!) [26] as ontology pitfall scanners and validators to identify potential issues requiring attention. OOPS! classifies pitfalls in three importance levels that are critical, important and minor [25]. For instance, missing annotations are categorized as minor pitfalls while missing domain/range definitions of properties are important pitfalls. A cyclic definition in the concept hierarchy is a critical pitfall. FOOPS! basically evaluates an ontology adhering to the FAIR principles [9] such as the usage of a persistent, resolvable URI (Findable); availability in different serialization formats (Accessible); metadata annotations using existing vocabularies (Interoperable) and human-readable documentation (Reusable).

The results generated by these tools offer insights into various design aspects (e.g., as shown in Table 2). However, a deliberate modeling decision may trigger a reported pitfall by OOPS! or FOOPS!. For example, deliberately leaving the domain or range of a property undefined to enhance flexibility may be flagged by OOPS! as an important issue. Similarly, FOOPS! checks if an ontology's metadata includes bibliographic citation information, which may not be deemed necessary at a certain stage of development. As a result, the results provided by

OOPS! Example Result and Decision			
Pitfall description	importance	action	reason
Creating unconnected ontology elements	Minor	future consideration	may be for future work
Missing domain or range in properties	Important	Update for some	for flexibility

FOOPS! Example Result and Decision		
Pitfall description	action	reason
Detailed metadata (The following metadata was not found: doi, publisher, logo, status, source)	Disregard	not relevant at this stage

Table 2
Example OOPS! and FOOPS! results and our actions throughout the ontology development process.

these validators primarily serve as guidance rather than offering an exhaustive and absolute list of issues requiring correction. We address major issues identified by these tools and track the decisions (i.e., correction, disregard, defer to future consideration) made for them using the GitHub repository's issue tracker.[16]

5. Discussion and Future Work

BadmintONTO captures semantic information within badminton play-by-play data, offering potential advantages for data analytics applications necessitating semantic comprehension and data integration purposes. Currently, our implementation of the ontology primarily focuses on annotating single matches sourced from ShuttleSet given that the ShuttleSet dataset exclusively comprises play-by-play data for single matches. Further investigation is required to explore how to annotate data from (mixed) double matches. For example, while in a single match, we can infer that a hitting player's location corresponds to that of the shuttle, as provided by ShuttleSet. In a double match, it is essential to also document the location of the other player within the pair (e.g., as indicated in CQ10 in Section 3). Additionally, we currently do not test BadmintONTO for representing team competitions, as the organizational structure of a team competition differs slightly from that of a regular tournament. Therefore, additional information may need to be recorded beyond the scope of ShuttleSet's coverage to accurately capture the play-by-play dynamics of team competitions. In the future, we will investigate how to model such team matches. For instance, a **TeamCompetitionEvent** may be introduced as a sub-concept of **Event**.

We also intend to keep track of the evolution of badminton-related datasets, including ShuttleSet, and extend BadmintONTO with new concepts and relationships. This extension may involve representing event sequences, detailed trajectory information of shuttles and players, and validating unaddressed competency questions.

6. Concluding Remarks

This paper introduces BadmintONTO, an ontology designed for the badminton domain. It follows the NeOn methodology and integrates elements from existing ontologies such as DOLCE, the OWL-Time ontology, and various public ontology patterns. Additionally, we showcase usage scenarios in which BadmintONTO is utilized: **(1)** to annotate the recently published shot-by-shot data - ShuttleSet dataset, and **(2)** for asking complex queries over the annotated dataset. Furthermore, we evaluate the ontology using the tools OOPS! and FOOPS!. Our model also permits generality, flexibility, and the potential for future endeavors, such as establishing certain aspects of the ontology as a standard pattern for ontological modeling in racket-based sports (e.g., tennis, table tennis and padel).

BadmintONTO is publicly maintained on a GitHub repository and accessible via a permanent URI provided by the w3id service. We have provided guidelines, within the GitHub repository, of how to contribute to the development of BadmintONTO. We anticipate that more domain experts, application developers, or badminton enthusiasts could contribute to its future development.

[16]Pitfall report from OOPS! and FOOPS!: https://github.com/huanyu-li/BadmintONTO/issues/5

References

[1] T. González Dos Santos, C. Wang, N. Carlsson, P. Lambrix, Predicting Season Outcomes for the NBA, in: Machine Learning and Data Mining for Sports Analytics, 2022, pp. 129–142. doi:10.1007/978-3-031-02044-5_11.

[2] M. Bendtsen, Regimes in baseball players' career data, Data Mining and Knowledge Discovery 31 (2017). doi:10.1007/s10618-017-0510-5.

[3] E. Nsolo, P. Lambrix, N. Carlsson, Player Valuation in European Football, in: Machine Learning and Data Mining for Sports Analytics, Springer International Publishing, 2019, pp. 42–54. doi:10.1007/978-3-030-17274-9_4.

[4] R. Säfvenberg, N. Carlsson, P. Lambrix, Identifying Player Roles in Ice Hockey, in: Machine Learning and Data Mining for Sports Analytics, Springer Nature Switzerland, 2024, pp. 131–143. doi:10.1007/978-3-031-53833-9_11.

[5] K.-S. Chang, W.-Y. Wang, W.-C. Peng, Where Will Players Move Next? Dynamic Graphs and Hierarchical Fusion for Movement Forecasting in Badminton, Proceedings of the AAAI Conference on Artificial Intelligence 37 (2023) 6998–7005. doi:10.1609/aaai.v37i6.25855.

[6] K.-W. Ban, J. See, J. Abdullah, Y. P. Loh, BadmintonDB: A Badminton Dataset for Player-specific Match Analysis and Prediction, in: Proceedings of the 5th International ACM Workshop on Multimedia Content Analysis in Sports, MMSports '22, Association for Computing Machinery, 2022, p. 47–54. doi:10.1145/3552437.3555696.

[7] W.-Y. Wang, W.-W. Du, W.-C. Peng, ShuttleSet22: Benchmarking Stroke Forecasting with Stroke-Level Badminton Dataset, in: Proceedings of the Intelligent Technologies for Precision Sports Science (IT4PSS), 2023. URL: https://arxiv.org/abs/2306.15664v1.

[8] R. Keskisärkkä, H. Li, S. Cheng, N. Carlsson, P. Lambrix, An Ontology for Ice Hockey, in: Proceedings of the ISWC 2019 Satellite Tracks (Posters & Demonstrations, Industry, and Outrageous Ideas) co-located with 18th International Semantic Web Conference (ISWC 2019), volume 2456 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2019, pp. 13–16. URL: http://ceur-ws.org/Vol-2456/paper4.pdf.

[9] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, et.al., The FAIR Guiding Principles for scientific data management and stewardship, Scientific Data 3 (2016) 160018:1–9. doi:10.1038/sdata.2016.18.

[10] S. Kovalchik, M. Reid, A shot taxonomy in the era of tracking data in professional tennis, Journal of Sports Sciences 36 (2018). doi:10.1080/02640414.2018.1438094.

[11] N. Bouayad-Agha, G. Casamayor, L. Wanner, F. Díez, S. López Hernández, FootbOWL: Using a Generic Ontology of Football Competition for Planning Match Summaries, in: The Semantic Web: Research and Applications, 2011. doi:10.1007/978-3-642-21034-1_16.

[12] B. J. Kim, Y. S. Choi, Automatic baseball commentary generation using deep learning, in: Proceedings of the 35th Annual ACM Symposium on Applied Computing, 2020, p. 1056–1065. doi:10.1145/3341105.3374063.

[13] P.-M. Filippidis, C. Dimoulas, C. Bratsas, A. Veglis, A unified semantic sports concepts classification as a key device for multidimensional sports analysis, in: 13th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP), 2018, pp. 107–112. doi:10.1109/SMAP.2018.8501868.

[14] R. Piryani, N. Aussenac-Gilles, N. Hernandez, Comprehensive Survey on Ontologies about Event, in: Proceedings of the Workshop on Semantic Methods for Events and Stories (SEMMES) co-located with ESWC2023, volume 3443 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2023. URL: https://ceur-ws.org/Vol-3443/ESWC_2023_SEMMES_XPEvent.pdf.
[15] X. Chen, Q. Li, Event modeling and mining: a long journey toward explainable events, The VLDB Journal 29 (2019). doi:10.1007/s00778-019-00545-0.
[16] A. A. Krisnadhi, P. Hitzler, A Core Pattern for Events, in: Proceedings of the 7th Workshop on Ontology Design and Patterns (WOP 2016) co-located with the 15th International Semantic Web Conference (ISWC 2016), 2016. doi:10.3233/978-1-61499-826-6-29.
[17] A. Krisnadhi, Y. Hu, K. Janowicz, P. Hitzler, R. Arko, S. Carbotte, C. Chandler, M. Cheatham, D. Fils, T. Finin, P. Ji, M. Jones, N. Karima, K. Lehnert, A. Mickle, T. Narock, M. O'Brien, L. Raymond, A. Shepherd, M. Schildhauer, P. Wiebe, The GeoLink Modular Oceanography Ontology, in: The Semantic Web - ISWC 2015, Springer International Publishing, Cham, 2015, pp. 301–309. doi:10.1007/978-3-319-25010-6_19.
[18] M. Rinne, E. Blomqvist, R. Keskisärkkä, E. Nuutila, Event processing in RDF, in: Proceedings of the 4th Workshop on Ontology Design and Patterns (WOP 2013) co-located with the 12th International Semantic Web Conference (ISWC 2013), volume 1188 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2013. URL: https://ceur-ws.org/Vol-1188/paper_1.pdf.
[19] S. Borgo, R. Ferrario, A. Gangemi, N. Guarino, C. Masolo, D. Porello, E. M. Sanfilippo, L. Vieu, DOLCE: A descriptive ontology for linguistic and cognitive engineering, Applied Ontology 17 (2022) 45–69. doi:10.3233/AO-210259.
[20] R. Battle, D. Kolas, Enabling the geospatial Semantic Web with Parliament and GeoSPARQL, Semantic Web 3 (2012) 355–370. doi:10.3233/SW-2012-0065.
[21] M. C. Suárez-Figueroa, A. Gómez-Pérez, M. Fernández-López, The NeOn Methodology for Ontology Engineering, in: Ontology Engineering in a Networked World, Springer, Berlin, Heidelberg, 2012, pp. 9–34. doi:10.1007/978-3-642-24794-1_2.
[22] Y. Sure, S. Staab, R. Studer, On-To-Knowledge Methodology (OTKM), in: Handbook on Ontologies, Springer, Berlin, Heidelberg, 2004, pp. 117–132. doi:10.1007/978-3-540-24750-0_6.
[23] V. Presutti, E. Blomqvist, E. Daga, A. Gangemi, Pattern-Based Ontology Design, in: Ontology Engineering in a Networked World, Springer, Berlin, Heidelberg, 2012, pp. 35–64. doi:10.1007/978-3-642-24794-1_3.
[24] A. Dimou, M. Vander Sande, P. Colpaert, R. Verborgh, E. Mannens, R. Van de Walle, RML: A Generic Language for Integrated RDF Mappings of Heterogeneous Data, in: Proceedings of the Workshop on Linked Data on the Web co-located with the 23rd International World Wide Web Conference (WWW 2014), volume 1184 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2014. URL: http://ceur-ws.org/Vol-1184/ldow2014_paper_01.pdf.
[25] M. Poveda-Villalón, A. Gómez-Pérez, M. C. Suárez-Figueroa, OOPS! (OntOlogy Pitfall Scanner!): An On-line Tool for Ontology Evaluation, International Journal on Semantic Web and Information Systems (IJSWIS) 10 (2014) 7–34. doi:10.4018/ijswis.2014040102.
[26] D. Garijo, O. Corcho, M. Poveda-Villalón, FOOPS!: An Ontology Pitfall Scanner for the FAIR Principles, in: International Semantic Web Conference (ISWC) 2021: Posters, Demos, and Industry Tracks, volume 2980 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2021. URL: http://ceur-ws.org/Vol-2980/paper321.pdf.