

Semantic Modeling of Events Using Linked Open Data

Sehrish Jamil¹, Salma Noor^{1,*}, Iftikhar Ahmed², Neelam Gohar¹ and Fouzia¹

¹Shaheed Benazir Bhutto Women University, Peshawar, 25000, Pakistan

²University of Engineering and Technology, Peshawar, 25000, Pakistan

*Corresponding Author: Salma Noor. Email: dr.salmanoor@sbbwu.edu.pk

Received: 10 February 2021; Accepted: 29 March 2021

Abstract: Significant happenings in terms of spatio-temporal factors are called events. In the digital age, these events and their associated features are scattered in various databases on the Internet. The event data are in heterogeneous formats, which are often not machine-readable. This leads to a lack of unification of event-related knowledge across different domains and results in a research gap in terms of event modeling and representation. Specialized event models are needed to overcome this gap and integrate relevant information of different similar events occurring worldwide. Our research explores the problem of heterogeneity in specialized event modeling and takes modeling for refugee registration and repatriation events as a case study. Our research explores the problem of heterogeneity at the data level and proposes a solution to this problem in the field of refugee registration and repatriation events. Considering refugee migration is one of the biggest crises in the world. The proposed model is designed according to Semantic Web standards to ensure reusability and machine readability. The project uses Protégé to model classes and ontology. Our ontology is called OntoEvent ontology and Karma is used for data mapping over ontology. Heterogeneity for the same concepts collected through the internet and through UNHCR (reports, excel sheets) is analyzed and resolved during the data modeling phase. As a result of this research, a timeline is designed to visualize events over time, along with a semantic data model and Linked Open Data representation of refugee data that we believe is of global significance. The W3C Ontology Validation Service has successfully validated the proposed OntoEvent ontology.

Keywords: Semantic Web; event modeling; ontology; linked open data; semantic representation; data science; machine readability

1 Introduction

The first generation of the web started with the read-only web and the system of cognition. The second generation of the web took a step forward and introduced machine generated web pages and active HTML pages. The common feature between the first two generations of the web is that they are both human oriented/readable. The third generation of the Web is what is commonly known as the “Semantic Web”, and its main difference from the previous two generations is that it targets machine-readable information [1].



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This research contributes a semantic data model for registration and repatriation events that occur during refugee crises. Semantic web technologies are used for their inherent advantages of machine readability, standardization, and discovery of new knowledge. The aim of this research is to propose a specialized event model for refugee *registration* and *repatriation* events. The need for such a model stems from the fact that refugee displacement is a global humanitarian crisis affecting many countries in the contemporary era. The events modeled in this research are refugee registration and repatriation.

Events have become central elements in representing information for various Semantic Web applications [2]. Many Semantic Web vocabularies aim to describe static things and relationships between them, but events can be continuous or non-static (dynamic) and their representation requires time as a factor. Event information can be used in different domains and in different context. It is important that this information is modular. The idea is to create a model that is interoperable, reusable, and accessible [3].

Explicit modeling of events and event-based systems is receiving a lot of attention in research and industry. A growing number of systems handle events, e.g., in media delivery, surveillance video, or emergency management. Therefore, building an event model is important for processing information in a variety of domains such as emergency response, sports, news, and law, to name a few [4]. A study on the conceptual modeling of events and the extensive use of Web Ontology Language (OWL) for events is observed in Hanzal et al. [5]. Typically, types/classes from taxonomies or ontologies are semantically assigned to events and the entities of their participants [6].

There are many ontologies that model events, such as Simple Event Model (SEM) and The Linking Open Description of Events (LODE), The EventsML-G2, The Event Ontology and E-multimedia Event Model. However, currently there are no specific ontologies for representing refugee data. The reasons for displacement of refugees are many, such as instability in their countries or terrorism, hence the desire for a better quality of life and the resolution of ethnic problems. This work is a pioneer in developing an ontology and data representation model for refugee data.

The event model presented here is applicable to refugee data representation worldwide and can be extended as needed in different situations.

From refugee displacement data, the events of registration and repatriation of refugees are extracted. Registration is the collection, verification and updating of information on affected persons with the aim of protecting and documenting them and implementing durable solutions. The events include registration of refugees with UNHCR organization, providing them with facilities such as education, health, and others. Registration events include the presentation of information such as location, date, number of people, gender, and families. Repatriation is the process of returning a person—voluntarily or forcibly—to their caretaker or place of origin or nationality. Repatriation event includes information on areas, date and year, number of individuals and families of refugees returning to their countries. The following are examples of events:

a. Example from Registration Event:

There are 19,560 registered births of Afghan refugees in September 2018.

b. Example from Repatriation Event:

In March 2018, 365 Afghan families and 1,152 individuals were repatriated from Peshawar.

1.1 Problem Statement

“The problem of heterogeneity arises in event modeling when the same event occurs in two different domains, or when the event occurs in the same domain, but different language constructs are used.” Both conditions pose challenges for modeling event information. As a result, heterogeneity is observed at three levels. (1) Ontological heterogeneity (2) Temporal heterogeneity and (3) Data-level heterogeneity.

- Ontological heterogeneity occurs when ontologies exhibit structural differences and are expressed using different ontology languages.

- Temporal heterogeneity arises when entity values or definitions belong to different times, or time intervals.
- Data-level heterogeneity is observed when databases reporting on the same entity adopt different representation choices.

For example, in our research, the same data collected from different databases, UNHCR demographic reports, Excel spreadsheets and websites have different names for the common attributes of a similar event. So, when the events are collected from these different representations, it is difficult to process the data. Here semantic modeling is used to build a generalized/unified model.

Our research explores the problem of data-level heterogeneity in Semantic Representation and Event Modeling of similar events over time using linked open data and proposes a solution to this problem in the field of refugee registration and repatriation events. The proposed model is designed according to Semantic Web standards to ensure reusability and machine readability.

2 Related Work

Event ontology is a new paradigm for representing and inferring events. It emphasizes the representation of event classes and relations. Formalisms for knowledge representation provide structures for organizing this knowledge, but do not provide mechanisms for sharing it. Ontologies provide a common vocabulary to support knowledge sharing and reuse. The existing event ontologies are discussed below.

The Event ontology [7], The Simple Event Model [8], The Linking Open Descriptions of Event [9], The Event-Model-F [4], The EventsML-G2 [10], The E multimedia Event model [11] these all ontologies are specifically event centric ontologies and Agora Project [12] is event centric project.

2.1 State of the Art in Existing Event Ontologies

Event ontology is a new paradigm for representing and reasoning about events. It emphasizes the representation of event classes and relations. Formalisms for knowledge representation provide structures for organizing this knowledge, but do not provide mechanisms for sharing it. Ontologies provide a common vocabulary to support knowledge sharing and reuse. Existing event ontologies are discussed in [Tab. 1](#).

The above table shows some prominent event ontologies with their domains, advantages, and disadvantages. In the review we conducted for our research we concluded that research on refugee event ontologies has not been attempted before. The ABC model [13], The CIDOC/CRM model [14] are promising models but they only address events for culture domain. Our work and contribution of the current research are the first of its kind. The existing ontologies can be reused or extended to form new ontologies. In the current research work, the ontology SEM has been extended to OntoEvent ontology by inheriting its classes and relationships.

3 Data Acquisition

In obtaining data sets for Afghan refugees, the three authentic data sets from UNHCR, Humanitarian Data Exchange and Operation Portal Refugee Situations were selected. The collection and use of refugee data is mandated by the 1951 Refugee Convention and by the Statute of the Office of the High Commissioner for Refugees. UNHCR teams conduct the surveys in refugee camps to collect the data and publish the reports on the website. We have collected data from the UNHCR office, website, and other sources. These data sets from different sources use different terminologies for the same concept. Introduction of Linked Open Data Cloud would help synchronize, update, and manage the data in a better way and would facilitate new data discovery. To overcome the shortcomings during the preprocessing phase, we compared, analyzed, and collected concepts from these datasets to model ontology classes.

Table 1: Represents advantages and disadvantages of event models in comparison with other event models

Event Model/ Ontology	Domain	Advantages	Disadvantages
<i>The Event Ontology</i>	Conferences, festivals, concerts, or music chords	It relies on vocabularies defined externally to refine the knowledge expressed. For example, no Agent class is defined <i>per se</i> , but their co-agent property has foaf: Agent as a range: EO benefits therefore from the richness of the FOAF vocabulary. The explicit linking to vocabularies brings EO its richness.	<ol style="list-style-type: none"> 1. EO defines a minimal event, Roles, Types, Views and Temporary are not defined in EO, 2. Place, Time and Agent are defined via range restrictions on EO's properties. 3. This model constrains the possible values for mentioned properties above.
<i>The Simple Event Model (SEM)</i>	<ol style="list-style-type: none"> a) historic events (geography, history, cultural heritage and multimedia) b) maritime safety and security 	<ol style="list-style-type: none"> 1. SEM is compatible with more Place, Time and Actor representations as compared to EO. 2. The main common point between SEM and EO is the modularity in the design: most classes are optional in EO; In SEM, the sem: Event class is optional. It was important to model a (potential) actor or a place independently of the events they might participate in. 3. This model does not make formal mappings and functional property restrictions, and do not conform to one single vocabulary for properties as compared to LODE. 4. The SEM model can be connected to other models for vocabularies not directly but reserves placeholder for Role and Type. This will keep external hierarchies out from core classes. 5. The CIDOC-CRM's model is generalized with SEM, and representation of View is added to fulfill requirements of SEM. 6. SEM gathers light-weight description of events without strong semantic definitions leads to inconsistency. 7. SEM specifies the necessary additions for dealing with heterogeneous and messy data from the web. 	<ol style="list-style-type: none"> 1. SEM do not define an explicit vocabulary of types (e.g., of events, roles etc.), but provide placeholder classes for using other vocabularies instead. 2. SEM do not deal with relations between events other than sem:hasSubEvent. 3. SEM do not have strong mappings to other models, as a result SEM cannot use expressivity for direct inferences.
<i>The Linking Open Descriptions of Event (LODE) ontology</i>	Historical events	<ol style="list-style-type: none"> 1. Both the class and the properties are formally mapped to other event models like the CIDOC-CRM, EO and DUL using owl: sameAs and rdfs:subPropertyOf, by this interoperability is enabled and a user can benefit from existing more complex vocabularies. 2. The properties of Role, Type and View can be expressed via their mapping to DUL, by using the Description and Situation patterns of F-model. F specializes D&S patterns from DUL. 	<ol style="list-style-type: none"> 1. This model aims at a minimal modeling of events. 2. LODE itself keeps its own classes and properties to the lowest possible number. This model get benefit from the other models or vocabularies directly.
<i>The Conceptual Reference Model (CIDOC/CRM)</i>	Cultural heritage and museum documentation	<ol style="list-style-type: none"> 1. In this model Roles are represented as constraint on property. 2. The property of Types can be applied to all entities of CIDOC-CRM. 	<ol style="list-style-type: none"> 1. In this model Roles are assigned to Actors only. 2. The Timestamps can only be applied to Temporal Entities: Roles, Types only. Other constituents of event remain without Timestamped.

4 Methodology

The methodology in current research is related to events, their modeling, mapping, and publishing on LOD and query using SPARQL. The general methodology for event modeling is given in Fig. 1.

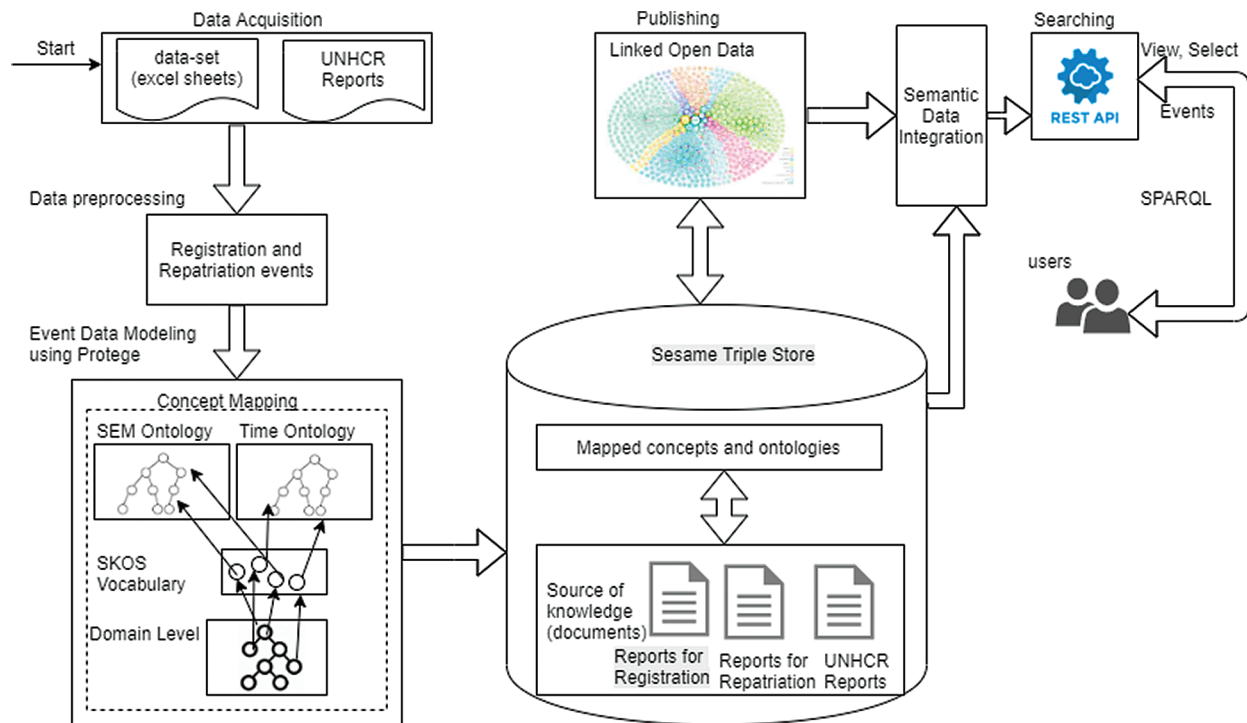


Figure 1: Flow chart representing methodology

Fig. 1 shows flowchart representing methodology. The datasets collected are in different resources and formats. These datasets are preprocessed by transforming and cleansing the data and as a result retrieved events of *Registration and Repatriation*. The Event datasets are mapped with concepts in proposed ontology called OntoEvent so that users query the RDF format stored in sesame triple store using API.

4.1 Data Preprocessing

The data we acquired are in the form of demographic reports, which required preprocessing to extract relevant information about repatriation and registration events. During preprocessing, we performed two steps. First, we manually converted the data format from a .pdf to an .xls file; this format is lightweight and therefore quick to use and access. Second, we prepared the data cleaning in Open Refine software. The empty rows and redundant data are removed from our dataset.

4.2 Data Modeling

The dataset is collected, transformed, and cleaned. Next, we need to create a model in an automated way that shares and reuses event information. To this end, we have developed an ontology called OntoEvent that facilitates users to represent event-related information in a standardized and machine-readable form. Concepts are selected from demographic reports for data modeling, using the Protégé tool, these concepts are modeled in a hierarchy as classes and subclasses.

The OntoEvent ontology contains concepts (classes) collected from reports and relevant ontologies, e.g., FOAF, SEM, Geospatial Ontology.

4.2.1 *Ontology Design*

For the design of the ontology, the tool Protégé version 5.5.0 build beta 7 was chosen. In Protégé, we have given the ontology IRI (Internationalized Resource Identifier) as <http://www.semanticweb.org/refugees/OntoEvent> and our ontology version IRI is <http://www.semanticweb.org/refugees/OntoEvent/1.0>. The ontology SEM is imported as Foundational Ontology. The classes Core, Event, Time and Place in SEM are reused in OntoEvent ontology. In OntoEvent, the Event class has three subclasses, namely *the EventID*, *Registration* and *Repatriation* classes. The Registration and Repatriation classes are extended with additional subclasses. Further subclasses are the classes *Population* and *Organization*.

Then the hierarchy of the classes is designed, the data properties and object properties are defined for the classes. We have specified domain, range and cardinalities in the object properties and the data properties are specified with domain and data type of the data. In data modeling, some classes are disjoint with other classes, for example, the classes “Registration” and “Return” are disjoint because both are sub-events of the class “Event”.

4.2.2 *Event Data Mapping*

After completing data modeling, the next phase is event data mapping. In event data mapping, data is mapped to classes (concepts) using integration and mapping tools. We used Karma, a data integration tool, for data mapping. It learns to recognize the mapping of data to ontology classes and then uses the ontology to propose a model that connects these classes.

- Firstly, we import OntoEvent ontology and data (.xls file) to integrate data according to OntoEvent using a graphical user interface that automates much of the process.
- Secondly, we set the semantic type for each mapped class, in which a class and property with ontology name are selected from given list. Karma gives suggestions for data mapping if the suggestion is appropriate according to our data mapping need, we often just selected the suggestion, otherwise we deliberately set class and property for data mapping.
- Lastly, We saved the data mapping model to be used later for next phase.

4.2.3 *Data Transformation in RDF*

- After data mapping is completed, we have transformed our mapped data to Resource Description Framework (RDF).
- In Karma an option for converting mapped data to RDF format is OpenRDF. After mapping we clicked OpenRDF our mapped data transformed to RDF language as triples (subject, predicate, and object).
- We have saved RDF file.

4.2.4 *Triple Store*

- The Karma generates RDF triples in triple store named as sesame triple store that provides triple store services, SPARQL endpoints and the REST web service. The mapped concepts and ontologies along with documents are stored in Sesame triple store to be able to query data using SPARQL query language.
- We created our repository named OntoEvent in sesame triple store by giving ID, description and a location is provided as URL.
- For example, <http://localhost:8080/openrdf-sesame/repositories/OntoEvent>
- The RDF triples stored in our repository.

4.2.5 Query Interface

For query purpose Karma provides Query repository for writing and executing SPARQL queries. For querying, Sesame REST API is used to query Linked Data. The SPARQL is a query language used to query the semantic data.

- We open OntoEvent repository a Query Repository page opened.
- According to experiment designed we write different queries.
- Results are displayed after executing the queries.

4.3 Publishing an Enrich Event Model

The proposed OntoEvent ontology is enriched by linking events to Linked Open Data (LOD) and publishing it as part of the LOD open access data graph. Gephi software is chosen to publish our data as a network/graph. It contributes to the Linked Open Data Cloud a sub-graph of information regarding the refugee registration and repatriation event. This open linked cloud could now serve as a resource for other related collections and promote interoperability between them. Publishing the enriched event data on LOD, makes it available for public access for future reuse and research, see Fig. 2.

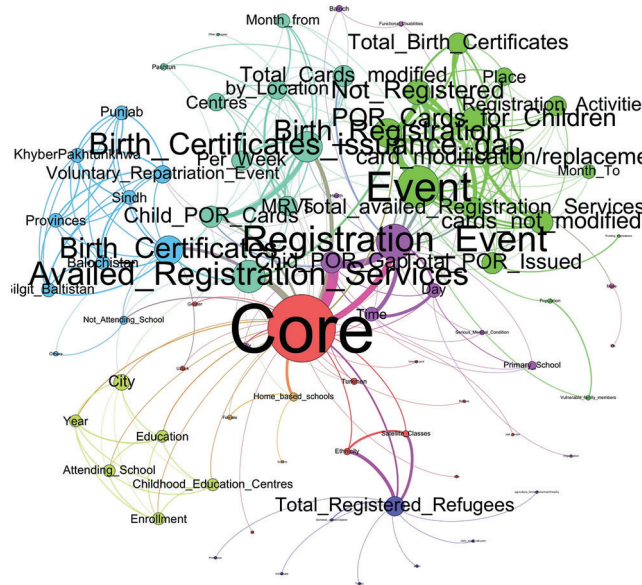


Figure 2: Graph visualization of ONTOEVENT ontology as LOD cloud

5 Implementation

5.1 Ontology Description

The ontology design process takes into consideration various issues to achieve a “data-centric” model such as avoiding inconsistencies, self-explanatory concept, and giving examples of usage. The ontology design details are discussed in Noor et al. [15]. This section describes the main entities in the OntoEvent ontology. We focus on core classes and properties provided by the ontology.

5.1.1 Core Classes

The OntoEvent ontology borrows some of the main classes from the foundational ontologies. For the ones not explicitly matching with the concepts addressed by OntoEvent, new concept definitions have been developed. The core entities of the refugee events in OntoEvent are:

- **Event:** as the entity of main interest, including metadata such as event type (registration, repatriation).
- **Agents:** describes the Organizations hosting or funding the Registration and Repatriation events and other information related to Persons (refugees) e.g., health, education, gender etc.
- **Location:** the city and country in which the Registration and Repatriation events occurred.
- **Time:** to describe the duration of Registration and Repatriation events with year and date.

5.1.2 Specialized Classes

During the ontology requirement analysis and design process it was observed that some of the defined or reused classes require further specialization to address the complexity and diversity of the concepts, so appropriate subclasses were introduced in OntoEvent.

For example, *availed_registration_services*, *Birth_Registration*, *Birth_certificates*, *not_registered*, *Birth_certificate_issuance_gap*, *POR_cards_for_children* are the specialized classes that were added as a subclass of the parent class *RegistrationEvent*.

In addition, a number of classes were added that are missing in existing ontologies, e.g., *Refugee type*, *Occupation*, *EventID* and *Voluntary_Repatriation_Event* classes are some of the specialized classes added to the parent class agent and person. The other specialized subclasses are *Place*, *Centers*, *City*, *Provinces*, *Time*, *Day*, *Month_from*, *Month_to* and *Year*.

5.1.3 Class Disjointness

We ensure pairwise disjointness between classes in the ontology where appropriate. For example, the class *RegistrationEvent* is disjoint with *RepatriationEvent* and *Month_from* is disjoint with *Month_to*.

The conceptual model of OntoEvent classes and features with data types are shown in Fig. 4. The model shows the classes as SEM, since the underlying ontology reuses concepts from other ontologies, e.g., *Dublin Core: identifier*. We have linked the OntoEvent ontology with other classes, such as *Place* with the vocabulary *SpatialThing*, *Time* with the vocabulary *TemporalEntity*, the class *Population* from the FOAF ontology and Dublin Core with the vocabulary Identifier. All these vocabularies and ontologies are used in the OntoEvent ontology, as shown in Fig. 3.

Fig. 3 shows the relationship between main classes with their respective subclasses. The main class is SEM: *Event*, and its subclasses are *Registration*, *Repatriation*, *Time*, and *Place*. These classes have many subclasses with corresponding relationships. For example, the class *Time* has a *time:hasTime* relationship with the class SEM: *Event*, the class *Place* has a *geo:hasPlace* relationship with the class SEM: *Event*. These classes and relationships are shown in Fig. 4. Arrows with open arrowheads denote *rdfs:subClassOf* properties between the classes.

5.2 Defining Classes

The main classes in the OntoEvent ontology are five. Each class has subclasses. The total number of classes in the OntoEvent ontology is 105.

5.3 Properties

Two types of properties are defined in Protégé for the OntoEvent Ontology, namely object and data. The object properties show the relationship between classes. While data properties show the relationship between

individual and its data type. For object properties, a domain and range are defined, e.g., *Date* is an object property whose domain is *Day* and range is *Time*. The domain is the area of the property and the range is the boundary where it can access its superclass. Similarly, in the OntoEvent ontology, *Year* is a data property, it has domain *Year* (class) and range *DateTime* (data type).

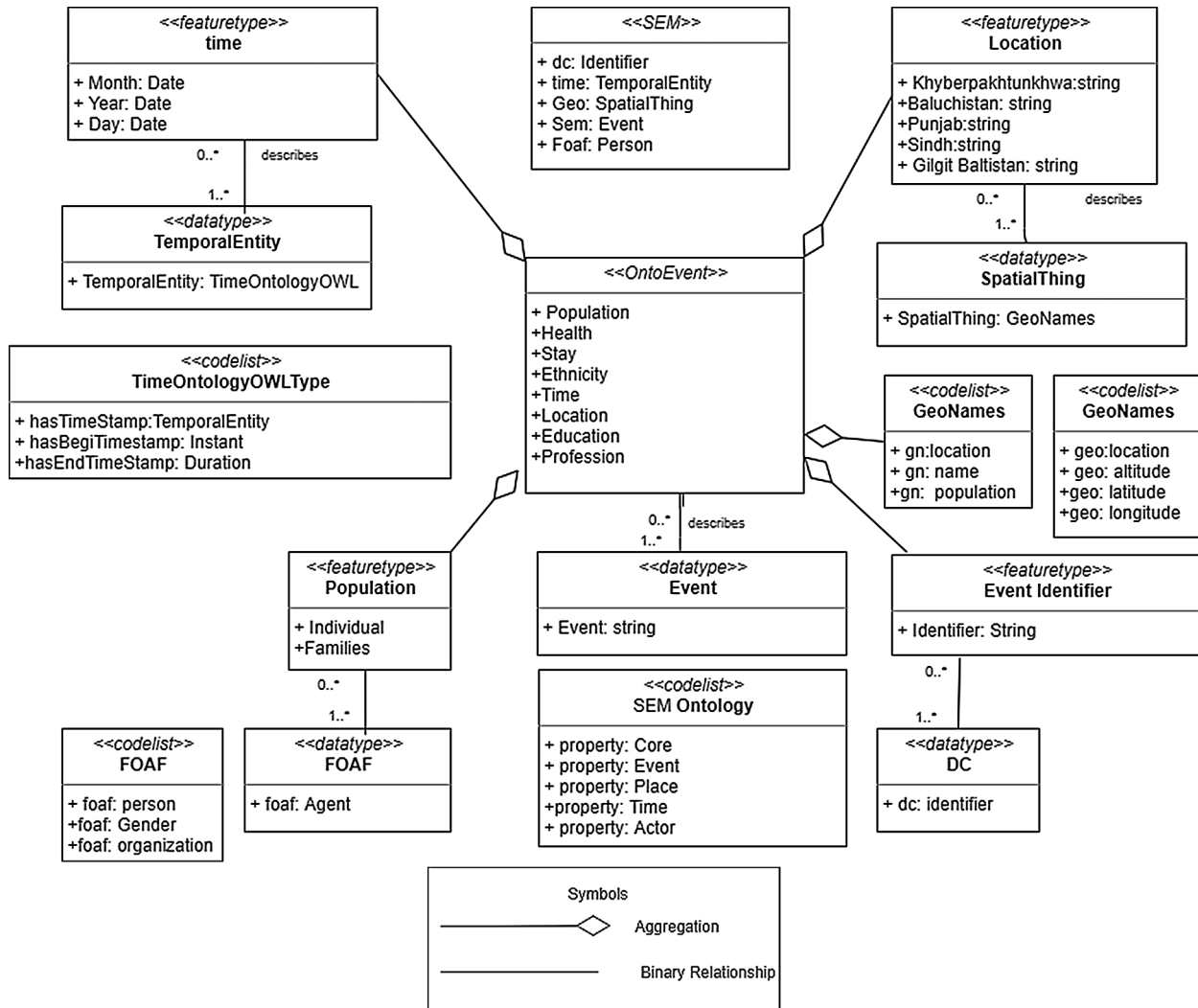


Figure 3: A general outlook of conceptual event data model consist of classes, features, and data types

Some classes in the base ontology, i.e., the Simple Event Model ontology and the proposed OntoEvent ontology, are the same. These classes are marked as equivalent classes with Protégé, e.g., SEM:Event and OE:Event.

5.4 Visualization of Event Model

A model for event-based visualization is developed. We designed the interactive visualization using JavaScript and CSS stylesheets with Bracket software. The result is a timeline that depicts the events of refugee registration and repatriation. It is useful for visualizing events over time in an interesting and attractive way. It allows the user to graphically analyze the event data, i.e., when and where an event

6 Results and Evaluation

6.1 Results

The OntoEvent Ontology is evaluated by designing and testing two scenarios. For testing, we have used stest of more than twenty SPARQL queries in both the scenarios. An example query in SPARQL would read as follows. Fig. 6 shows the results obtained from the following query.

```
SELECT Distinct ? Births_Certificate_issued ?Year
WHERE { ?Event1 OntoE:hasissuedbirthcertificates ?Birth_certificates1.
?Birth_certificates1 gn:population Births_Certificate_issued.
?Year1 sem:hasTimeStamp ?Year.
FILTER(?Year = "2017" && ?Births_Certificate_issued > "90%"). }
```

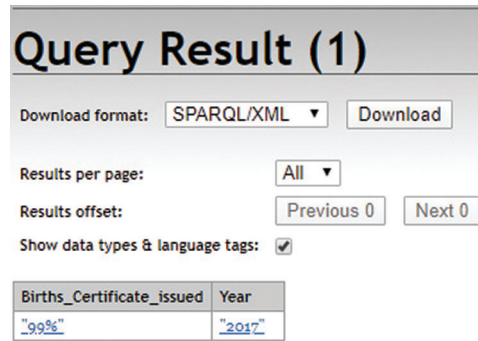


Figure 6: Results for query regarding birth certificate

In the first scenario, a UNHCR administrative officer is tasked with conducting a risk analysis on the influx of refugees in Pakistan. The queries are designed according to the scenarios and the results are shown in Tab. 2.

Table 2: Cumulative percentages of queries of case 1

No. of Queries	Percentage	Cumulative percentage
Query 1	100%	100%
Query 2	100%	
Query 3	100%	
Query 4	100%	

The second scenario was designed and tested for academic purpose in which a student was assigned a project to analyze the situation of refugees in the country. The results showed 100 percent accuracy and precision because the data is represented in machine readable and understandable form through the use of ontology. The result is shown in Tab. 3.

Table 3: Cumulative percentages of queries of case 2

No. of Queries	Percentage	Cumulative percentage
Query 5	100%	100%
Query 6	100%	
Query 7	100%	
Query 8	100%	

6.2 Ontology Validation

The OntoEvent ontology is validated using W3C Validation Service. Validation of ontologies is defined by technical specifications of languages (e.g., RDF, OWL, etc.), which typically include a machine-readable formal grammar (and vocabulary). The act of checking a document against these constraints is called ontology validation. The RDF document of the OntoEvent ontology was successfully validated, as shown in Fig. 7.

Check by Direct Input

```
<?xml version="1.0"?>
<rdf:RDF xmlns="http://www.semanticweb.org/refugees/OntoEvent#"
  xmlns:ns="http://creativecommons.org/ns#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:xml="http://www.w3.org/XML/1998/namespace"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:adms="http://www.w3.org/ns/adms#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
```

Parse RDF Restore the original example Clear the textarea

Display Result Options:
 Triples and/or Graph: Triples Only
 Graph format: PNG - embedded

Paste an RDF/XML document into the following text field to have it checked. More

Figure 7: Uploaded RDF document to W3C validation service

The result is shown as triples with comment that the RDF document is validated successfully is shown in Fig. 8.

Validation Results

Your RDF document validated successfully:

Triples of the Data Model

Number	Subject
1	http://www.semanticweb.org/refugees/OntoEvent
2	http://www.semanticweb.org/refugees/OntoEvent
3	http://www.semanticweb.org/refugees/OntoEvent
4	http://www.semanticweb.org/sshja/ontologies/2018/3/untitled-ontology-15#Event_of_Registration
5	http://www.semanticweb.org/sshja/ontologies/2018/3/untitled-ontology-15#Event_of_Registration
6	http://www.semanticweb.org/sshja/ontologies/2018/3/untitled-ontology-15#RE
7	http://www.semanticweb.org/sshja/ontologies/2018/3/untitled-ontology-15#RE
8	http://www.semanticweb.org/sshja/ontologies/2018/3/untitled-ontology-15#Registration_Event
9	http://www.semanticweb.org/sshja/ontologies/2018/3/untitled-ontology-15#Registration_Event
10	http://www.w3.org/2004/02/skos/core#broadMatch

Figure 8: Validation results of OntoEvent ontology

7 Conclusion and Future Work

We identified the need for an event ontology for refugee data. Events related to refugees are scattered across the Internet, offline databases, and newspapers. To represent the unified resource for management and analytical support, an ontology named OntoEvent Ontology is proposed to provide a platform that can be reused in different domains for knowledge enrichment. Users can retrieve knowledge from the OntoEvent Ontology as various classes and features are identified in the model for querying. Semantic Web technologies, i.e., RDF, are used to represent event data semantically. The model is linked with Linked Open Data and other Event Ontologies to enrich the data model and integrate concepts of domain ontology with basic ontologies to solve heterogeneity problems. In the OntoEvent ontology, new classes, concepts, and relationships were added for the humanitarian domain of refugees. These classes have enriched the event model compared to previous SEM or event models. The contribution of the OntoEvent ontology is in the humanitarian domain, i.e., no such work has been done for refugee events to our knowledge.

The goals of the research are to centralize event data in one place while eliminating heterogeneity, this is achieved by creating an ontology in Protégé. The OntoEvent ontology is then mapped to the actual data using the Karma data integration tool. The query results for evaluation are obtained using SPARQL query API. The event data in the OntoEvent ontology is shareable and reusable by linking them together LOD. The main contribution of our research is to propose the first of its kind ontology dealing with refugee registration and repatriation events.

In the future, an API for the OntoEvent ontology is proposed so that useful applications can be developed on this data. The applications for humanitarian work, for funding and for academic purposes can be mapped on the OntoEvent Ontology. The OntoEvent Ontology can be used and extended for multiple purposes. Therefore, the OntoEvent Ontology can be used for scenarios such as humanitarian research, foreign funds, government finance department and historians to name a few.

Acknowledgement: Special thanks to UNHCR Peshawar Pakistan for providing dataset/reports regarding Afghan Refugees.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] T. B. Lee, *Weaving the web: The original design and ultimate destiny of the worldwide web by its inventor*. San Francisco, CA: Harper San Francisco, 1999.
- [2] W. Liu, N. Ding, Y. Tan, Y. Zhang and Z. Liu, *Event representation and reasoning based on SROIQ and event elements projection*. In: Hung, J., Yen, N., Li, K. C. (eds.), *Frontier Computing. Lecture Notes in Electrical Engineering*, vol. 375. Springer, Singapore, 2016.
- [3] F. Soualah-Alila, C. Faucher, F. Bertrand, M. Coustaty and A. Doucet, "Applying semantic web technologies for improving the visibility of tourism data," in *Proc. of the Eighth Workshop on Exploiting Semantic Annotations in Information Retrieval*, Melbourne, Australia, pp. 5–10, 2019.
- [4] A. Scherp, T. Franz, C. Saathoff and S. Staab, "F–A model of events based on the foundational ontology dolce+ DnS ultralight," in *Proc. of the Fifth Int. Conf. on Knowledge Capture*, Redondo Beach, California, USA, pp. 137–144, 2009.
- [5] T. Hanzal, V. Svátek and M. Vacura, "Event categories on the semantic web and their relationship/object distinction," in *Formal Ontology in Information Systems, FOIS*, Annecy, France, pp. 183–196, 2016.
- [6] S. Ananiadou, P. Thompson, R. Nawaz, J. McNaught and D. B. Kell, "Event-based text mining for biology and functional genomics," *Briefings in Functional Genomics*, vol. 14, no. 3, pp. 213–230, 2015.

- [7] Y. Raimond and S. Abdallah, *The event ontology*, 2007 [Online]. Available: <http://motools.sf.net/event/>.
- [8] W. R. Van Hage, V. Malaisé, R. Segers, L. Hollink and G. Schreiber, "Design and use of the simple event model (SEM), Web Semantics: Science," *Journal of Web Semantics*, vol. 9, no. 2, pp. 128–136, 2011.
- [9] R. Shaw, R. Troncy and L. Hardman, "Lode: Linking open descriptions of events," in *Asian Semantic Web Conf.*, Berlin, Heidelberg: Springer, pp. 153–167, 2009.
- [10] EventsML-G2 Events ML-G2, 2012, [Online] Available: <https://iptc.org/standards/eventsml-g2/>.
- [11] U. Westermann and R. Jain, "E-A generic event model for event-centric multimedia data management in eChronicle applications," in *IEEE Proc. 22nd Int. Conf. Data Engineering Workshops*, Atlanta, GA, USA, pp. x106, 2006.
- [12] P. M. Jenlink and B. H. Banathy, "The agora project: The new agoras of the twenty-first century," *Systems Research and Behavioral Science*, vol. 19, no. 5, pp. 469–483, 2002.
- [13] C. Lagoze and J. Hunter, "The ABC ontology and model," *Journal of Digital Information. British Computer Society and Oxford University Press*, vol. 2, pp. 1–18, 2001.
- [14] M. Doerr, "The CIDOC conceptual reference module: An ontological approach to semantic interoperability of metadata," *AI magazine*, vol. 24, no. 3, pp. 75, 2003.
- [15] S. Noor, S. Jamil, N. Gohar and L. Shah, "Knowledge retrieval of historic concepts using semantic web," *Journal of Cluster Computing*, vol. 22, no. S3, pp. 7321–7332, 2019.