SEDE: An ontology for scholarly event description

Article *in* Journal of Information Science · March 2010

DOI: 10.1177/0165551509358487

2 authors:

Senator Jeong
National Institute of Health, Korea

Hong-Gee Kim
Seoul National University

SEDE: The Scholarly Event Description Ontology

Senator Jeong; Hong-Gee Kim[1]

Biomedical Knowledge Engineering Laboratory, Seoul National University, Korea

Abstract

Scholarly events, such as conferences, workshops, and seminars, are important scientific communication channels. Our research goal is to satisfy scientists' basic information needs by collecting, archiving, and providing access to scholarly event information. Furthermore, we aim to satisfy users' in-depth information needs by excavating scholarly meaningful information through reasoning about knowledge. A prerequisite to accomplishing this end is to define a description base for scholarly events to enable software agents to crawl and extract scholarly event data, and to facilitate the unified access to them and reason about the collected data. We present the design and implementation of the Scholarly Event Description Ontology (SEDE) to achieve the research goal, and its application use case scenarios in scholarly event information space – scholarly event information extraction; semantic search & retrieval on scholarly event; knowledge organization; scholarly event similarity calculation; knowledge structure analysis; and academic performance evaluation.

Keywords: scholarly event; ontology; scholarly information

1. Introduction

Together with scientific journals, scholarly events, such as conferences, workshops, and seminars, are important scientific communication channels. They provide scholars with opportunities to publish up-to-date scientific research results, to get feedback from scientific communities, and to exchange research interest and ideas with each other. They also demonstrate current research trends, and as such, they are valuable information sources from which researchers can obtain knowledge on research directions and trends.

We postulate scientists' information needs with regard to scholarly events. First, *information need of a simple magnitude*: which events will be (or have been) held on which topics and when, where and who will organize them and what the due dates are for calls for submissions. Though the information of this magnitude is quite rudimentary, it is not often the case that a scientist gets a full and exhaustive picture of scholarly events held in the world. The reason for this may be due to the sheer volume of scholarly events held by various academic societies and organizations – no single information channel has been successful at keeping track of ever-growing

Correspondence to: Hong-Gee Kim, Biomedical Knowledge Engineering Laboratory, #102 Bio Material Research building., Seoul National University, 28-22 Yeon-GeonDong JongnuGu, Seoul 110-749, Korea , hgkim@snu.ac.kr.

JIS-0976-v2 Received: 1st July 2009 Revised: 2nd July 2009

conferences and providing their information to scientists. Information is also needed for *scientifically meaningful inference*: who is regarded as prominent in a specific topic; which event is the most prominent in one's field of study; who are the best scientists suited for questions, consultations, and collaboration. Researchers, especially newcomers or beginners in a given field, may have special interests. These information needs might be met partially at a minimal level since almost all event websites list leadership members such as general chairs, committee members, invited speakers and/or award winners. Users, however, are not yet satisfied, since they are not able to get a whole picture and existing library services, with regard to scientific events, do hardly provide this kind of meaningful information in an integrated and collective manner, hence, making the users' search for relevant information a cumbersome endeavour.

Our research goal is to satisfy scientists' basic information needs by collecting, archiving, and providing access to scholarly event information. Furthermore, we aim to satisfy users' in-depth information needs by excavating scholarly meaningful information through reasoning about knowledge. A prerequisite to accomplishing this end is to define a description base for scholarly events to enable software agents to crawl and extract scholarly event data, and to facilitate the unified access to and reason about the collected data. We present design and implementation of the Scholarly Event Description Ontology (SEDE) to achieve the research goal, and its implication for scholarly-events information space.

This paper begins with the description of the related work on scholarly event descriptions and systems. Then we present the methodological approach including the generic event description model, ontology structure and key concepts. Next, we cover the use scenarios to demonstrate the benefits of the ontology in real life applications. At the end, we provide the discussion and conclusions.

2. Background

There are several services in place that offer basic information on scholarly events: EventSeer (http://eventseer.net), PapersInvited (http://papersinvited.com/), and conference Alerts (http://conferencealerts.com/), to name a few. They mainly focus on call-for-paper where event titles, locations, deadlines are described. Useful as they may be, their functionalities are limited in that they provide just the simple metadata about forthcoming events and they do not "talk" to each other, due to the proprietary description formats used that hinder unified access to, sharing and exchanging of, event information. On the other hand, our research is not only to provide users with access to basic event information, but also to aim to crawl, extract, archive event data and infer scholarly meaningful information with greater detailed description base. We took an ontological approach to achieve these goals.

The ontology-based description approach is not a new one. The ESWC Conference ontology [1] was designed to describe the Semantic Web Conference (it started from ESWC2006). It uses FOAF [2], Geo [3], and SWRC [4] as core namespaces. Though it provides a suitable level of description power, it works best only for ESWC conference, because other events may not have the exactly same structure and terminology. For example, in some event, a session is grouped around a specific time interval, while, in others, the same term 'session' can mean a track grouped around a specific research topic or theme. On the other hand, our research attempts to build a more axiomatic scholarly event description basis. The Iugo project (http://iugo.ilrt.bris.ac.uk/) extends the ESWC Conference Ontology [5]. And it is designed to present data for discovery and retrieval of materials that relate to particular events at user interface level. To this end, it subclasses the ESWC Conference Ontology to model its proprietary major classes and properties. It inherits the ESWC ontology and so does its problems. Hence, the model is not enough to comprehend entities to harvest scholarly event data, and to infer scholarly meaningful information from them. Furthermore, its prototype implementation (http://iugo.ilrt.bris.ac.uk/IugoPortal/) has no difference with other services in place except that it provides a little bit more details on events and enables registered users to make annotations in the website interface.

To represent sound and comprehensive scholarly event information, we need a generic common ground for event description. There are several initiatives to model event-focused concepts in knowledge representation[6],

JIS-0976-v2 Received: 1st July 2009 Revised: 2nd July 2009

medical informatics[7], culture and heritage[8] and multimedia community[9-11] –the last three describe resources to facilitate the interoperability across systems and domains.

The ABC ontology is based on an event-aware model to address semantic interoperability of individual metadata throughout the lifecycle of multimedia content and events on it, such as transition, transformation, and so on[9, 11]. It was designed in line with several metadata models such as INDECS[12], FRBR[13], CIDOC-CRM[8], MPEG-7[14], and Dublin Core. They identified common entities and their relationships to each other. As such, the top classes of the ABC model are “Resource”, “Event”, “Event Relation”, “Agent”, “Act”, “Role”, and “Context” (Time, Place). The ABC model is a good reference for modeling our generic event structure. The INDECS framework and OntologyX meta-model provide a meaningful starting point to define what comprises an event. INDECS explains the relationships among entities in light of events in the intellectual rights management domain. The events in INDECS are the glue in the model: all metadata relationships are either events in themselves, or rely on events to establish them. The INDECS framework was partially reflected by the CIDOC Conceptual Reference Model which is a formal ontology intended to facilitate the integration, mediation and interchange of heterogeneous cultural heritage information. The OntologyX, the INDECS framework’s extension, was designed to provide a hub for the interoperability of meaning to support sophisticated machine-to-machine communication on the network [15]. It aims to be a generic master ontology for mapping terms between different ontologies. The schema includes a context term of the basic term set that is assigned a meaning signifying a circumstance in which at least the basic action occurs. As such, OntologyX provides a powerful reference model to determine the primary abstract classes for the generic event model. The Enterprise Architecture[16] exhibits many similarities with OntologyX, ABC model, and the CIDOC-CRM at the upper level of abstraction. Its classification—What (Data), How (Function), Where (Location), Who (People), When (Time), and Why (Motivation)—seems to comprise ingredients of event model. Our generic event model takes on this abstraction. The Event Ontology [17] defines an “Event” concept. It is the base ontology of the Music ontology [10] which represents music–related information such as performances, composition events, recordings, arrangements, and so on. In the ontology, an event consists of Agent, Thing, Spatial Thing, and Temporal Entity. This ontology model also gives an insight into how to model our generic event structure. We synthesize these previous works to design the generic event model.

3. Generic Event Model

The SEDE ontology was developed in ways that could provide enough description power with fine granularity to span over multiple scientific disciplines and capture as many varied event types as possible. This development approach comes from the realization that there exist discrepancies among the terms used by scholarly events, such as conferences, workshops and seminars. And a scholarly event description ontology, to be of any use, must reflect the variations and, at the same time, provide a framework that enables consistency, extensibility, and a high degree of interoperability between event descriptions. For this reason, we started from designing a generic event model that provided building blocks for scholarly event representation. We adopted an event definition in English dictionaries: *event is something that happens at a given place and time*. This definition is broad enough to take a snapshot of events that we are interested in. We might paraphrase this definition: a collateral combination of five English interrogatives: Who, How, What, When, and Where. These interrogatives are then translated into event primitives to formalize an event description as following:

Event≡ (□*Agent*) □ (□*Action*) □ (□*Patient*) □ (□*Place*) □ (□*Time*)

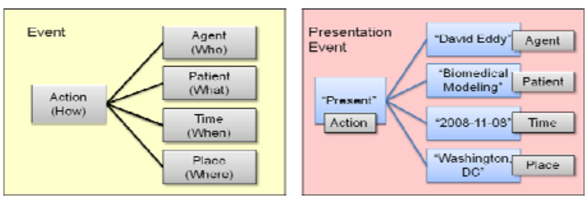


Figure 1. Primitives in the generic event model (left) and its exemplary event description (right).

JIS-0976-v2 Received: 1st July 2009 Revised: 2nd July 2009

This formalism will be an axiom of the generic event representation as illustrated in Figure 1 and Table 1. The exemplary event description tells us that "David Eddy" present on "biomedical modeling" November 8, 2008 in Washington, DC, which could be stated as:

$(\exists Agent(\text{D.Eddy})) \wedge (\exists Action(\text{present})) \wedge (\exists Patient(\text{Biomedical Modeling})) \wedge (\exists Place(\text{Washington})) \wedge (\exists Time(\text{2008-11-08}))$.

Table 1. The classes of the Generic Event Model.

Class	Definition	Example
Event	A collateral combination of five English interrogatives.	conference, workshop, symposium, talk, meeting, and so on.
Agent	The entity playing an initiative role in an event.	person, organization, group, committee, company, and so on.
Action	The characteristic activity in an event.	organize, present, demonstrate, disseminate, discuss, and so on.
Patient	Any physical or abstract thing that is acted upon or undergoes an action in an event.	topic, theme, and so on.
Time	A point or portion of time which forms a temporal parameter of an event.	year, month, date/time, and so on.
Place	A point or portion of three-dimensional or virtual space which forms a spatial parameter of an event.	country, city, university campus, URL, and so on.

4. The SEDE Model

4.1. Ontology Modelling Principle

In order to ensure its generality, granularity and resolvability, SEDE is modeled according to a number of principles. The SEDE ontology extends the generic event model to induce its own classes. We also tried to find a common ground with existing metadata and ontologies. We drew on some of the ontological structures found in previous ontology projects—the agent from FOAF, the topic from SKOS[18], the place from Geo[3]— to make SEDE a more comprehensive and interoperable ontology. From the survey and analysis of various scholarly event types, we identified scholarly event-specific properties and attributes such as "committee", "track", "session", and "atom event".

JIS-0976-v2 Received: 1st July 2009 Revised: 2nd July 2009

4.2. Scholarly Event Description Structure

A scholarly event may be made of more than one "track", grouped around a specific domain area, audience, format (e.g., poster, paper etc.), or purpose (e.g., industry, application). A track may have more than one "session", in a specific duration, in which "atom events" having similar topics are grouped in sequentially ordered manners. An "Atom event" is the minimum level of description in our ontology, which is made of presenter, topic, duration, and place as its core elements. Some domains require different levels of complexity depending on the size of the event, which are determined by the number of topics and participants. A scholarly event may be organized with several tracks. Or only a session may become a scholarly event. Furthermore, a scholarly event may contain other scholarly events as child events. Figure 2 illustrates the components of a scholarly event.

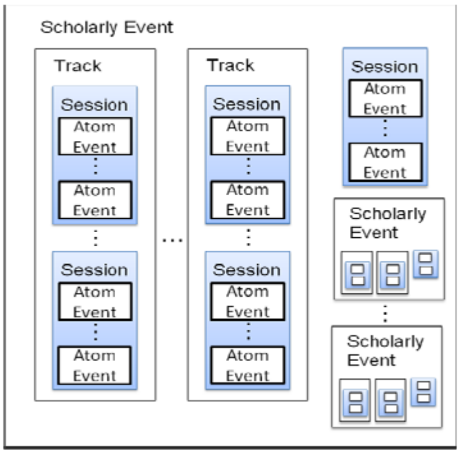


Figure 2. Scholarly Event Structure. A scholarly event may consist of session(s) which contains atom events, or track(s) containing more than one session.

4.3. Key concepts in the SEDE Ontology

All classes in the SEDE ontology, following the generic event model, is categorized into Agent (e.g., *Person*, *Committee*), Patient (e.g., *Concept*, *Artifact*), Place (e.g., *City*, *Venue*), and Time (e.g., *startDate*, *endDate*). Action is reflected by properties (e.g., *heldAt*). A scholarly event may have `Committee` for its organization and management and is a subclass of the `foaf:Agent`. A Patient group has `Artifact` and `skos:Concept`. The `Artifact` class subsumes `foaf:Document` which in turn subsumes `Call`, `Program`, `Proceedings`, and so on. The `geo:SpatialThing` class is a super class of place: `Country`, `City`, and `Venue`. `Time` is a super-property of `startDate`, `endDate`, and so on. Figure 3 shows the top level classes and their groupings in SEDE.

Journal of Information Science, XX (X) 2009, pp. 1–15 DOI: 10.1177/0165551506nnnnnn

JIS-0976-v2 Received: 1st July 2009 Revised: 2nd July 2009

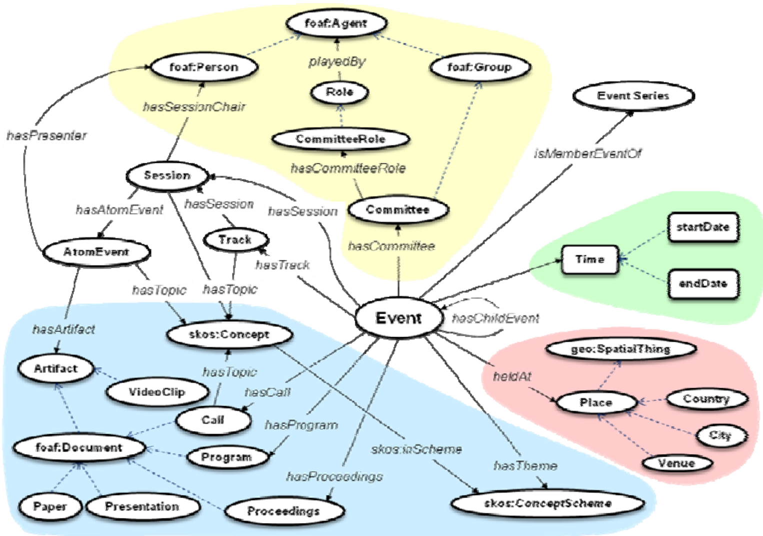


Figure 3. Some of the top level classes in the Scholarly Event Description Ontology. The ovals are classes and rounded rectangles are attributes. The dotted arrows indicate subclass (sub property) relationships, while the labelled line arrows represent binary relationships between two classes. Light yellow, green, red, and blue groups Agents, Time, Place, and Patient, respectively.

5. Implementation

In our current implementation, SEDE is coded in a set of RDFS classes and RDF/OWL properties used to express the content and structure of a scholarly event as an RDF graph. SEDE contains 30 classes with its own namespace (sede, http://eventography.org/sede/0.1/), two FOAF classes (`foaf:Agent`, `foaf:Document`), one GEO class (`geo:SpatialThing`), and two SKOS classes (`skos:Concept`, `skos:ConceptScheme`). The RDFS version of the ontology is available at http://www.eventography.org/sede/0.1/sede.rdf. Figure 4 shows a UML diagram of scholarly event and Table 1 describes the classes and properties.

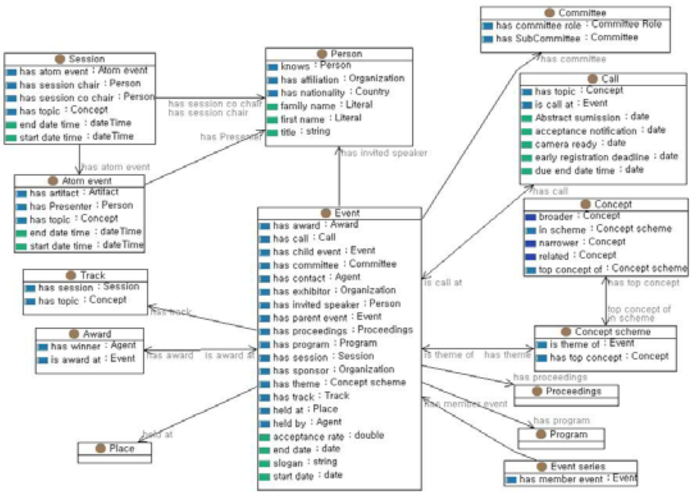


Figure 4. UML representation of Scholarly Event.

JIS-0976-v2 Received: 1st July 2009 Revised: 2nd July 2009

Table 2. A list of the major classes and properties in the SEDE ontology version 0.1.

Class	*Description*	*Property*
Event	Event includes conference, workshop, symposium, tutorial, workshop, seminar, and so on.	`hasAward, hasCall, hasChildEvent, hasCommittee, hasContact, hasExhibitor, hasInvitedSpeaker, hasParentEvent, hasProceedings, hasProgram, hasSession, hasSponsor, hasTheme, hasTrack, heldAt, heldBy, isMemberEventOf, acceptanceRate, startDate, end Date, slogan`
Track	Track represents a group of Sessions which are grouped around a specific domain area, audience, format (e.g., poster, paper etc.), or purpose (e.g., industry, application track).	`hasSession, hasTopic`
Session	Session represents a gathering of Atom Event's in which atom events having similar topics are, in a specific duration, grouped in sequentially ordered manners.	`hasAtomEvent, hasSessionChair, hasSessionCo-chair, hasTopic, startDateTime, endeDateTime`
AtomEvent	Primitive constituent of a scholarly event. An Atom Event has Agent, Action, Patient, Time, and Place as mandatory elements.	`hasArtifact, hasPresenter, hasTopic, startDateTime, endDateTime`
Call	Call for submission in an event, such as Call for Paper, Call for Workshop, Call for Proposal, and so on.	`abstractionSubmission, acceptanceNotification, cameraReady, earlyRegistration, hasTopic, isCallAt, submissionDeadline`
Person	Person is one who participates in an Event. Tutor, Invited Speaker, Author, Committee Member, and so on.	`foaf:knows, hasAffilliation, hasNationality, foaf:familyName, foaf:firstName, foaf:familyName, foaf:title`
Committee	Committee is a group in an event, such as programme committee, organizing committee, and so on.	`hasCommitteeRole, hasSubCommittee, isCommitteeOf`
Role	A function or position.	`playedBy, roleTitle`
Place	A place where an event takes place, such as venue, city, and country.	`Geo:based_near, geo:altitude, goe:latitude, geo:longitude`

JIS-0976-v2 Received: 1st July 2009 Revised: 2nd July 2009

5.1. n-ary relations and reification heuristics

Some entity-to-entity relationships need reification, for instance, the committee-person relation, because RDF Schema and OWL do not allow n-ary relations. Someone, for example, is a member of a committee: she has a role in the committee. We may describe this committee-to-person relationship with properties: `hasGeneralChair`, `hasTrackChair`, `hasPublicationChair`, `hasPublicityChair`, and so on. In many case, however, a committee may have various types of member role. Thus, creating many properties, with their range being the `Person` class, to enumerate endless lists of member roles is not desirable. One way to handle this problem is to introduce a new binary relation. Then, an n-ary relation can be represented by a new entity (i.e., `Role`) linked by binary relations to all of its related entities. Ideally the `Role` class should be a superclass of all reified classes for relations between `Agent` and `Event`. However, there is still a problem in the `Role` class because every role type should be enumerated and not all entity-to-entity relationship necessarily be reified for practical reason. In our heuristics, as Figure 5 shows, only the `Agent-Committee` relations are reified. In our current version, we predefined only one subclass of the `Role` class, so called `Committee Role`. Then, each role title is described via an OWL data property: `roleTitle`.

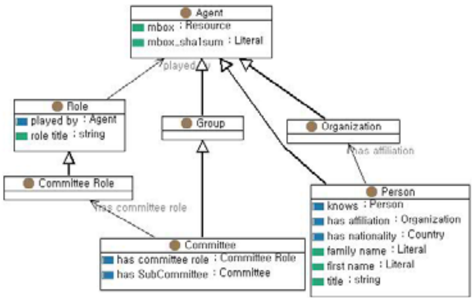


Figure 5. The reified relationship between `Committee` *and* `Agent` *via* `CommitteeRole` *which is subsumed by the* `Role` *class.*

5.2. Ontology improvement

We tested the structural soundness and conceptual clarity of the SEDE ontology by describing several real-world examples—applying to a number of different event types (conference, congress, workshop, and seminar) and various subject domains. Then we refined some of classes and adjusted the domains and ranges of properties to have confidence that the classes and properties in SEDE cover enough of the essential data of scholarly events. Figure 7 exemplifies the RDF/XML representation of a scholarly event with the SEDE ontology.

```
<?xml version="1.0"?>
<rdf:RDF xmlns:sede="http://eventography.org/sede/0.1/" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:skos="http://www.w3.org/2004/02/skos/core#" xmlns:foaf="http://xmlns.com/foaf/0.1/" xmlns="http://eventography.org/sede/0.1/#"
xmlns:vs="http://www.w3.org/2003/06/sw-vocab-status/ns#" xmlns:owl="http://www.w3.org/2002/07/owl#" xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:dct="http://purl.org/dc/terms/" xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" xmlns:geo="http://www.w3.org/2003/01/geo/wgs84_pos#"
xml:base="http://eventography.org/sede/0.1/">
<sede:Event rdf:ID="AMIA2009">
  <skos:altLabel>AMIA2009</skos:altLabel>
  <skos:prefLabel>AMIA 2009 Annual Symposium</skos:prefLabel>
  <foaf:homepage rdf:resource="http://symposium2009.amia.org"/>
  <sede:startDate rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2009-11-14</sede:startDate>
  <sede:endDate rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2009-11-18</sede:endDate>
  <sede:heldBy>
    <sede:Association rdf:ID="AmericanMedicalInformaticsAssociation">
      <foaf:name>American Medical Informatics Association</foaf:name>
    </sede:Association>
  </sede:heldBy>
  <sede:hasTheme rdf:resource="#AMIA2009TopicTrackCategory"/>
  <sede:hasCall>
    <sede:CallForParticipation rdf:ID="AMIA2009CallForParticipation">
```


JIS-0976-v2 Received: 1st July 2009 Revised: 2nd July 2009

```
        <rdfs:label>AMIA2009 Call for participation</rdfs:label>
        <foaf:homepage rdf:resource="http://symposium2009.amia.org/files/symposium2008/AMIA_Call_for_Participation_2009_0.pdf"/>
        <sede:submissionDeadline rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2009-03-13</sede:submissionDeadline>
        <sede:isCallAt rdf:resource="#AMIA2009"/>
        <sede:hasTopic rdf:resource="#TerminologyAndStandards"/>
        <sede:hasTopic rdf:resource="#Developing-RefiningEhrDataStandards"/>
        ...............
        <sede:hasTopic rdf:resource="#ConsumerInformaticsAndPHRs"/>
      </sede:CallForParticipation>
    </sede:hasCall>
    <sede:hasCommittee>
      <sede:Committee rdf:ID="AMIA2009SPC">
        <rdfs:label>AMIA2009 Scientific Program Committee</rdfs:label>
        <foaf:homepage rdf:resource="http://symposium2009.amia.org/SPC"/>
        <sede:hasCommitteeRole>
          <sede:CommitteeRole rdf:ID="AMIA2009SPCMember">
            <sede:roleTitle>Scientific Program Committee Member</sede:roleTitle>
            <rdfs:label>AMIA2009 SPCMember</rdfs:label>
            <sede:playedBy>
              <foaf:Person rdf:ID="RickyTaira">
                <foaf:name>Ricky Taira</foaf:name>
                <sede:hasAffiliation>
                  <sede:University rdf:ID="UniversityOfCaliforniaLosAngeles">
                    <rdfs:label>University of California-Los Angeles</rdfs:label>
                  </sede:University>
                </sede:hasAffiliation>
              </foaf:Person>
            </sede:playedBy>
        ...............
          </sede:CommitteeRole>
        </sede:hasCommitteeRole>
      </sede:Committee>
    </sede:hasCommittee>
    <sede:isMemberEventOf>
      <sede:EventSeries rdf:ID="AMIA">
        <rdfs:label>AMIA</rdfs:label>
        <sede:hasMemberEvent rdf:resource="#AMIA2009"/>
      </sede:EventSeries>
    </sede:isMemberEventOf>
  </sede:Event>
  </rdf:RDF>
```

Figure 6 . A fragment of the “AMIA 2009 Annual Symposium” with the SEDE ontology.

6. Application Use Cases

In this section, we present application use cases within the SEDE Application Framework. In the framework, the Event Data Crawler automatically identifies scholarly events’ web sites and fetches their URLs, and then the Event Data Extractor extracts and adds all found and relevant information to the SEDE ontology knowledge base (RDF repository). Various applications can use this knowledge base, as shown in Figure 7.

JIS-0976-v2 Received: 1st July 2009 Revised: 2nd July 2009

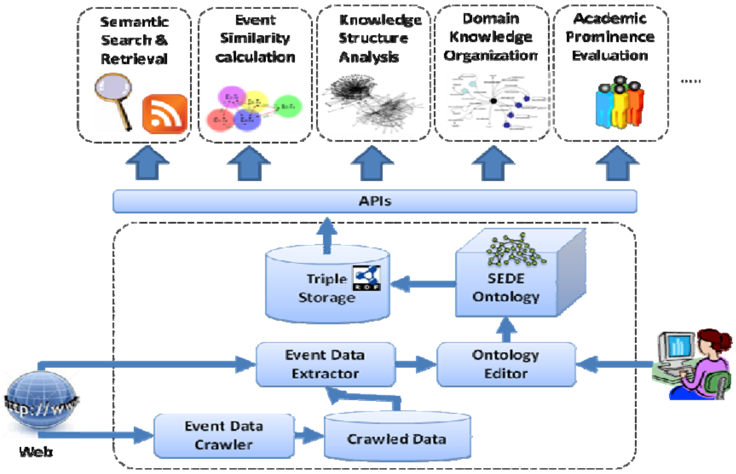


Figure 7. The SEDE Application Framework.

6.1. *Ontology-based Information Extraction*

The SEDE ontology provides a unified framework to extract scholarly event data. Challenges in information extraction are that information on events on the web comes in different presentation forms and styles. There is no standard for the posting of this heterogeneous information. Also, the sheer amount of data to be extracted from a single scholarly event is daunting. Manual instance population is very cost intensive. Thus, automatic information extraction techniques are preferred methods than pure manual efforts. The value of such techniques depends crucially on the accuracy of the extraction. However, the current fully automatic information extraction techniques [19-22]—especially for the types of data we are interested in —are not satisfactory in that human intervention to ensure the accuracy of the extracted data is still a required step. Thus, we developed a semi-automatic data extraction tool, the SEDE Data Extractor, to facilitate data extraction. The extractor parses event pages and extracts data. And then the user cleanses and save them as SEDE instances. The ontology model allows the extractor to recognize and instantiate data in a unified way.

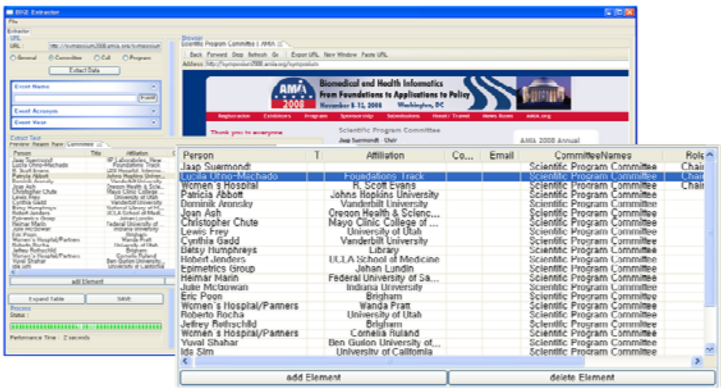


Figure 8. Instance Population with the SEDE Data Extractor.

6.2. *Semantic Search & Retrieval on Scholarly Events*

The ontology allows users to semantically search and retrieve scholarly event related information, Here we describe some exemplary real-life semantic search scenarios that use the SEDE ontology

JIS-0976-v2 Received: 1st July 2009 Revised: 2nd July 2009

- Find events with a specific call-for-paper topic, a submission deadline, and a start date.

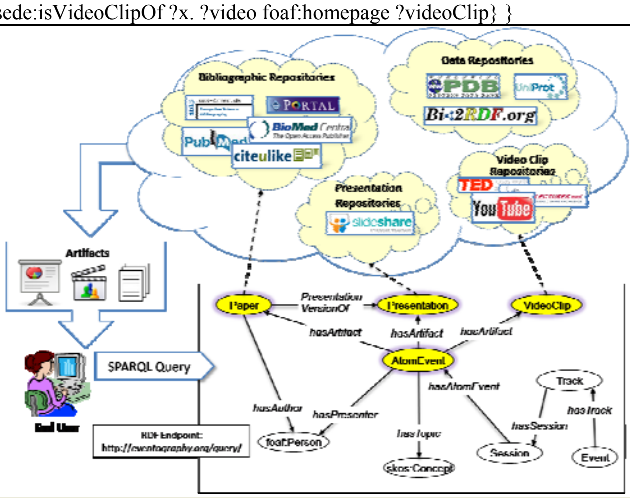
```
SELECT ?Event ?Topic ?Deadline ?When
WHERE {
?x a sede:Event; rdfs:label ?Event. ?x sede:hasCall ?y.?y rdfs:label ?Call.
?y sede:hasTopic ?z. ?z skos:prefLabel ?Topic.
?y sede:submissionDeadline ?Deadline. ?x sede:startDate ?When.
FILTER ( (regex(?Topic, "visual*")||regex(?Topic, "Visual*") )&& (regex(?Topic, "Bio*")||regex(?Topic, "bio*") ) )
}ORDER BY ?Topic ?Event
```

Event	Topic	Deadline	When
BIBE2007	Bio-Data visualization	2007-06-07	2007-10-14

Figure 9. A SPARQL query to find event which has specific call for paper topic, submission deadline, and event start date.

- Get artifacts presented in an atom event.

```
SELECT ?Paper ?videoClip
WHERE {?x a sede:AtomEvent. ?x sede:hasArtifact ?y. ?y a sede:Paper.
?y foaf:homepage ?Paper.
OPTIONAL{?video sede:isVideoClipOf ?x. ?video foaf:homepage ?videoClip} }
```



[Paper]	videoClip
<http://portal.acm.org/citation.cfm?doid=1...	<http://videolectures.net/foe0...

Figure 10. A query to search specific atom events' artifacts and their resolutions to the real data.

- Find domain experts.

```
SELECT DISTINCT ?Domain ?Expert ?Affiliation
WHERE{
        ?x a sede:Session. ?x sede:hasTopic ?topic. ?topic skos:prefLabel ?Domain.
        ?x sede:hasSessionChair ?chair. ?chair foaf:name ?Expert.
        FILTER ( regex(?Domain, "Decision")|| regex(?Domain, "decision"))
         OPTIONAL{?chair sede:hasAffiliation ?y. ?y rdfs:label ?Affiliation.}
}ORDER BY ?Domain
```

Domain	Expert	Affiliation
Clinical Decision Support Outcomes and Patient Safety	Charles Barr	
Clinical Decision Support Outcomes and Patient Safety	R. Scott Evans	University of Utah
Clinical Decision Support Outcomes and Patient Safety	Pierre Zweigenbaum	
Clinical Decision Support Outcomes and Patient Safety	Robert A. Jenders	UCLA School of Medicine
Clinical Decision Support Outcomes and Patient Safety	Patricia Abbott	Johns Hopkins University
Clinical Decision Support Outcomes and Patient Safety	Dominik Aronsky	Vanderbilt University
Clinical Decision Support Outcomes and Patient Safety	Karl Gumpper	
Clinical Decision Support Outcomes and Patient Safety	Adam Wright	

Figure 11. The result to the query" find expert on "decision" or "bio".

JIS-0976-v2 Received: 1st July 2009 Revised: 2nd July 2009

6.3. *Domain Knowledge Organization*

Scholarly events have c*all-for-paper* topics which are encoded in SKOS (Simple Knowledge Organization System)[18]. In SKOS, knowledge organization systems are comprised of conceptual resources. Thus, a SKOS format-encoded event topics act as semantic binding nodes to construct a new knowledge organization sysetm(KOS) and/or extend an existing one, such as classification systems, taxonomies, thesauri or ontologies. For example, one event (BIBE2007) has coarse-grained call-for-paper topics (`Bio-Ontology` is sub concept of `Biomedical Informatics and Computation`), while another event (Bio-Ontology and BioLink Workshop2006) has finer grained topics (`Bio-ontologies` has a number of narrower concepts). With these conceptual resources, we can generate a new compound domain KOS. Furthermore, each concept may be mapped to a semantic web knowledge base, such as DBpedia[23], which aims to connect related data which are not previously linked.

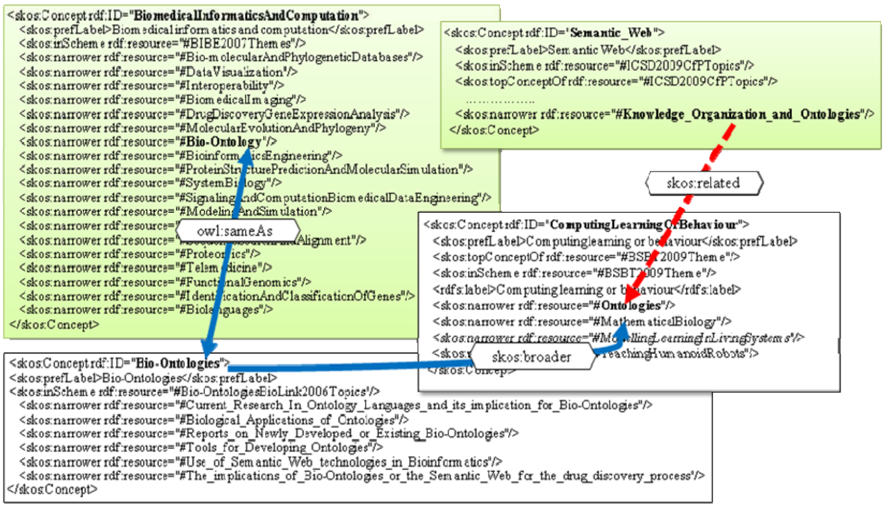


Figure 12. An example of domain knowledge organization using the SKOS semantic relation.

6.4. *Event Similarity Calculation*

The SEDE application framework enables a service to measure the topical similarities between scholarly events to couple events based on topics they cover. Similarity calculation may adopt the vector space model which is widely used for analyzing association strength between documents and query terms. The calculation uses topic terms of *call-for-papers* which represent events. Pairs of topics and events use TF/IDF term weights. Then, using the similarity algorithms (e.g., cosine coefficient), the method determines the degree of similarity between each pair of events.

Topic by Event matrix

	Event$_1$	*Event$_2$*	*Event$_3$*	…	*Event$_n$*
Topic$_1$	0.4	0.2	0	…	0
Topic$_2$	0.1	0.56	0	…	0
Topic$_3$	0.25	0	0.12	…	0
Topic$_4$	0.34	0.16	0.24	…	0.04
…	…	…	…	…	…
Topic$_n$	0	0.32	0	…	0.62

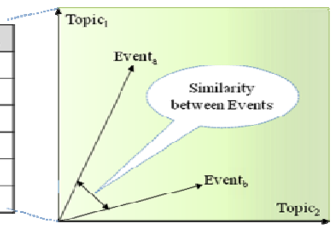


Figure 13. Event Similarity Calculation.

JIS-0976-v2 Received: 1st July 2009 Revised: 2nd July 2009

6.5. Knowledge Structure Analysis

Scholarly events are organized to communicate the latest state of research, and they, in nature, provide opportunities to identify updated scientific research results. Thus the use of scholarly event as a data source for a discipline's intellectual structure analysis is well warranted. Topics of call-for-papers and sessions are clear-cut indicators of a scientific community's interest and orientation. They represent either some knowledge which is not yet mature or specialized knowledge which is not yet generally applicable. They are a potentially valuable data source for knowledge structure analysis, and the social interactions and topologies of research topics. Following questions can be explored: What are the important topics in, for instance, biomedical informatics? What are the newly rising topics in a domain? To explore a domain knowledge structure, we may utilize scholarly events topics. This approach quantitatively explores the intellectual structure of domain knowledge to understand the domain better.

6.6. Academic Performance Evaluation

The ontology has practical implications for evaluating scientists' academic prominence and scholarly events. A scholar's prominence, based on the instances of SEDE, can be deduced from his/her total number of the elite group memberships (committee chairs and members, invited speaker, session chair) in scholarly events[24].

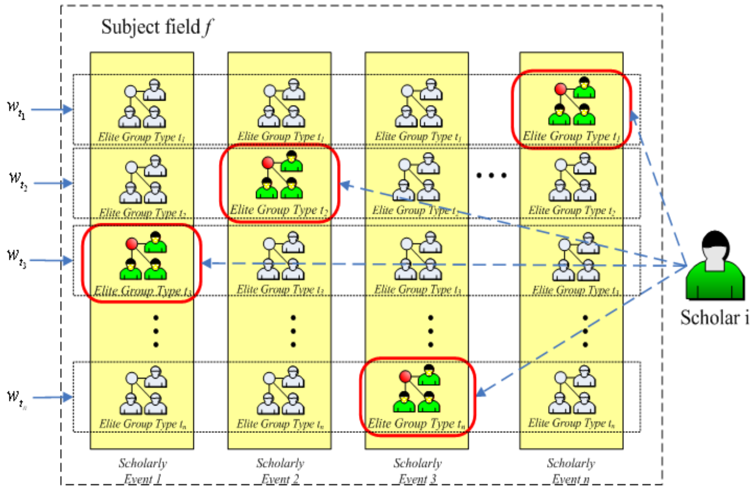


Figure 14. Academic Performance Evaluation.

Let T be the elite group, t be the elite group type (t□T), K_t be, for scholar S, the number of memberships for elite group types in scholarly events, W_t be the weight assigned to elite group type t, n_f be the total number of scholarly events in a given subject field f, and τ be the normalization factor to make $P(S)$ consistent. Then, the Subject Prominence Index of Scholar S is:

$$P(S) = \tau \frac{\sum_{t \in T} {}_{t=1}^{n} (w_t k_t \mid f)}{n_f}$$

where, scholar S_1 is said to be more prominent than scholar S_2 if $P(S_1) > P(S_2)$. Where, $0 \le P(S) \le 1$.

An event can be measured by accumulating the instances of participation of the elite group members in other events. This method enables us to numerically evaluate and determine rankings on scholars and scholarly events.

JIS-0976-v2 Received: 1st July 2009 Revised: 2nd July 2009

7. Discussion and Conclusion

In this paper we presented the development of the Scholarly Event Description Ontology to provide a backbone to represent, collect, extract, and infer from scholarly event information in a logical way. We also described its application use case scenarios—from the queries to satisfy users' basic information needs to the services to meet their scientifically meaningful information needs. Describing scholarly events in a unified model helps to improve information accessibility through greater semantic interoperability of information. It also makes it possible to build a *scholarly semantic web*—the isolated piece of scholarly event data integrated to automatically gain scientific value through its relationships with other scientific data on the web.

We do not claim that the model represented in SEDE completely captures every detailed and varying structure of scholarly events. We may not have included every concept that might be useful. We believe, however, the resulting scholarly event model is flexible enough to describe various event types and data, and is sound enough to be used in excavating scholarly meaningful information.

One of the difficulties encountered by this study was that it was hard to evaluate the ontology with a golden standard. This comes from the fact that there is no *canonical* way to judge the result of ontology modeling, and that ontology evaluation is to assess a given ontology from the view point of a particular application[25]. Thus, the ontology evaluation needs to be connected to the application use-case scenarios. The scenarios presented here suggest that the SEDE ontology satisfies various scholarly information needs. But the use-case scenarios are in nature ideological. Thorough testing is required to determine plausibility and suitability of the SEDE ontology to backup the proposed use case scenarios and enhance the current design.

References

[1] T. Heath, *ESWC2006* Conference *ontology* (2006). Available at: http://www.eswc2006.org/technologies/ontology-content/2006-09-21.rdf (accessed 10 March 2007).

[2] D. Brickley and L. Miller, *FOAF vocabulary* Specification (2005). Available at: http://xmlns.com/foaf/0.1/ (accessed 10 March 2009).

[3] D. Brickley, *Geo* (2003). Available at: http://www.w3.org/2003/01/geo/wgs84_pos (accessed 12 June 2007).

[4] Y. Sure, S. Bloehdorn, P. Haase, J. Hartmann and D. Oberle, The SWRC Ontology - Semantic Web for Research Communities In: C. Bento, A. CardosoandG. Dias (eds.). *2th Portuguese Conference on Artificial Intelligence - Progress in Artificial Intelligence (EPIA 2005)* (Springer, Covilha, Portugal, 2005) 218 - 231.

[5] N. Rogers, M. Jones, P. Shabajee and J. Tredgold, *Iugo* Deliverable *Report: Project Final Report* (Institute for Learning and Research Technology, 2006).

[6] J.F. Sowa, Knowledge representation: logical, philosophical, and computational foundations (Brooks/Cole, Pacific Grove, 2000).

[7] G. Schadow, C.N. Mead and D.M. Walker, The HL7 reference information model under scrutiny. *Studies in health technology and informatics* 124(1) (2006) 151-156.

[8] M. Doerr, The CIDOC Conceptual Reference Module an Ontological Approach to Semantic Interoperability of Metadata. *AI Magazine* 24(3) (2003) 75-92.

[9] C. Lagoze, J. Hunter and D. Brickey, An Event-Aware Model for Metadata Interoperability *4th European Conference on Research and Advanced Technology for Digital Libraries ECDL 2000, (*Lisbon, 2000).

[10] F. Giasson and Y. Raimond, *Music* Ontology *Specification* (2007). Available at: http://musicontology.com/ (accessed 20 December 2008).

[11] J. Hunter, Reconciling MPEG-7 and MPEG-21 Semantics through a Common Event-Aware Metadata Model (http://arxiv.org/html/cs/0210021, 2002).

[12] G. Rust and M. Bide, *The <indecs> metadata framework: Principles, model and data dictionary* (Indecs Framework Ltd., 2000).

JIS-0976-v2 Received: 1st July 2009 Revised: 2nd July 2009

[13] P. Riva, Introducing the functional requirements for bibliographic records and related IFLA developments. *Bulletin of the American Society for Information Science and Technology* 33(6) (2007) 7-11.

[14] B. Manjunath, P. Salembier and T. Sikora, *Introduction to MPEG-7: multimedia content description interface* (John Wiley & Sons, New York, 2002).

[15] G. Barlas, G. Rust, M. Quinlan and M. Dow, Demonstrator: Ontology Integration and Policy Enactment for Managing Rights Metadata, *2nd International Semantic Web Conference, (*Sundial Resort, Sanibel Island, Florida, USA, 2003).

[16] J.F. Sowa and J.A. Zachman, Extending and Formalizing the Framework for Information Systems Architecture. *IBM Systems Journal* 31(3) (1992) 590-616.

[17] Y. Raimond and S. Abdallah, *The Event Ontology* (2007). Available at: http://motools.sourceforge.net/event/event.html (accessed 17 July 2008).

[18] W3C, *SKOS Simple* Knowledge *Organization System Reference* (2009). World Wide Web Consortium. Available at: http://www.w3.org/TR/2009/CR-skos-reference-20090317/ (accessed 20 March 2009).

[19] A. McCallum, K. Nigam, J. Rennie and K. Seymore, A machine learning approach to building domain-specific search engines, *16th International Joint Conference on Artificial Intelligence (IJCAI-99), (*Pasadena, California, USA, 1999).

[20] F. Peng and A. McCallum, Accurate information extraction from research papers using conditional random fields, *Human Language Technology Conference and North American Chapter of the Association for Computational Linguistics (HLT-NAACL), (*Boston, USA, 2004).

[21] K.-M. Schneider, Using Information Extraction to Build a Directory of Conference Announcements, *Fifth International Conference on Intelligent Text Processing and Computational Linguistics (CICLing-2004), (*Seoul, Korea, 2004).

[22] K.-M. Schneider, Information Extraction from Calls for Papers with Conditional Random Fields and Layout Features, *16th Irish Conference on Artificial Intelligence and Cognitive Science (AICS 05), (*Portstewart, Northern Ireland, 2005).

[23] C. Bizer, T. Heath, K. Idehen and T. Berners-Lee, Linked data on the web (LDOW2008), *17th International Conference on World Wide Web 2008, WWW'08, (*Beijing, China, 2008).

[24] S. Jeong, S. Lee and H.-G. Kim, Are You an Invited Speaker? A bibliometric analysis of Elite Groups for Scholarly Events in Bioinformatics. *Journal of the American Society for Information Science and Technology* 60(6) (2009) 1118-1131.

[25] J. Brank, M. Grobelnik and D. Mladenic, A survey of ontology evaluation techniques, *The Conference on Data Mining and Data Warehouses (SiKDD 2005), (*Ljubljana, Slovenia, 2005).

JIS-0976-v2 Received: 1st July 2009 Revised: 2nd July 2009