SocIoS: A Social Media Application Ontology

Konstantinos Tserpes[1], George Papadakis[2], Magdalini Kardara[2],
Athanasios Papaoikonomou[2], Fotis Aisopos[2],
Emmanuel Sardis[2], and Theodora Varvarigou[2]

[1] Department of Informatics and Telematics, Harokopio University of Athens,
70 Eleftheriou Venizelou Str, 17676 Athens, Greece
`tserpes@hua.gr`
[2] Department of Electrical and Computer Engineering,
National Technical University of Athens,
9, Heroon Polytechniou Str, 15773 Athens, Greece
`{gpapadis,nkardara,tpap,fotais,sardis,doravarv}@mail.ntua.gr`

Abstract. The value that social web is adding is undeniable. However, social web is coming at a cost: the so-called "closed" web. Slowly but steadily, a growing portion of internet activity is confined within the spaces of social networking sites (SNS) or platforms that encompass social networking capabilities by default. Furthermore, due to their competitive stance between one another, conceptually common content and functionality are accessed through different mechanisms in SNSs. This work deals with the issue of semantic equivalence between the prevalent notions used in SNSs in order to device a common object model that will enable the integration of social platform APIs. The result is the proposal for an ontology for the implementation of cross-platform social applications. Two particular applications are considered as a guide for this reference model.

Keywords: Social Media, Semantic Web, Ontology, Application Programming Interface.

1   Introduction

Social media platforms are currently concentrating a great portion of web activity. Internet users are generating content and social graphs that carve the paths through which social information is shared. SNSs are exposing functionality and data not only through their usable GUIs but also through web APIs so as to ensure re-usability and interoperability. This enables them to support third party applications that add value to their platform and bring users in, without significant costs.

However, the competitive environment to which the SNSs are operating leads to the adoption of a multitude of data models that describe essentially similar notions, e.g. "friends" or shared "media items". Taking advantage of the conceptual similarity between these core SNS entities implies that the aggregation of the social functionality and data can provide the basis for a unique platform on top of which third parties can deploy new added value services, seamlessly using the underlying SNSs APIs.

In order to achieve that, the proposed solution attempts to develop an SNS-like API that will manage to aggregate data and functionality from underlying SNSs that deliver their functionality through web APIs. We call this API "SocIoS API" and it is the core component of the SocIoS integrated solution[1]. SocIoS aims at providing the tools to build applications that leverage on the dynamics and content that reside in the SNSs. These tools are a collection of services that make use of SNS metadata in order to organize information hosted in SNSs, each service for its own purpose. SocIoS platform allows the creative "mix-and-match" of these services to application workflows which are ready to be consumed.

This approach needs to be attacked on two sides: Firstly, by defining a new object model that captures all the conceptually similar notions, one can actually define the SNS API space. Each new SNS API that is added under this formal specification, needs to comply with the new object model. That is, a subset of the new SNS object model needs to be able to wrap into the common objects mentioned above. Such attempts have been undertaken by projects like GNIP® [1], hootsuite® [2] and Stroodle® [3], however, the most complete approach has been provided by OpenSocial [4].

Secondly, each time a new application service is created by combining available ones based on their specification, a new domain is created, extending the existing. This means that the specification has to be flexible so as to incorporate any possible new definition. Such generic-purpose modeling can be served by the Web Service Modeling Ontology (WSMO) [5] and the Semantic Annotation of Web Resources (SA-REST) [6].

Our method fosters ideas from both approaches: it defines the core objects of the SocIoS API domain through an ontology which helps to model these objects to the underlying SNS APIs using software wrappers (formally called SocIoS adaptors). To build this ontology we start with the conceptual analysis of proof of concept application scenarios in order to capture the main domain concepts. Next, we study a number of SNSs in order to identify the API objects that are semantically equivalent to the domain concepts. This helps us to place the domain concepts in the SNS context and examine their relationships and position in the object model hierarchy. The combination of these approaches leads to a conceptualization of the application domain that is formally captured in the presented SocIoS ontology.

The objective of this ontology is to provide the necessary flexibility of the SocIoS object model so as to be extended as necessary when new SNS APIs or application domains are to be included later in its lifetime. The model is captured using the ontology specification language (OWL) [7] that only includes the information that is necessary for an object model. We directly derive the initial SocIoS object model from this ontology.

The document is structured as such: Section 2 provides the related work in specifications and APIs that have similar goals as ours as well as ontologies from which our approach can adopt concepts so as to be extendible. Section 3 presents the domain concepts of proof of concept applications. In Section 4 the derived domain concepts

---

[1] www.sociosproject.eu

are projected to existing objects of a group of an appropriate selection of SNS APIs. Section 5 gives the proposed ontology as a result of this analysis and connects the classes to existing semantic web entities. Finally, Section 6, summarizes the key points that yield from this work.

2   Related Work

There have been a number of approaches to build a meta-API so as to implement functionality federation and data marshaling from numerous SNSs. However, as stated above, the majority comprise proprietary solutions which limits the extensibility of the platforms to non-supported APIs. Such examples are GNIP®, hootsuite® and Stroodle®.

Even though we study a number of SNS APIs so as to identify the dominant objects that they commonly use, our solution uses mainly OpenSocial as a reference specification. OpenSocial defines a common API for social applications across multiple websites. As it is a product of collaboration of a broad set of members of the web community it naturally encompasses the most dominant SNS concepts. However, OpenSocial is not meant to be used as a meta-API or an aggregator.

By defining the universe of discourse of the proof of concept applications, we assume that we will manage to capture the common denominator of the core SNS applications. This universe can be used at a higher layer by any service. However, here, we need to make a clear separation. With this work we do not want to model the real world concepts that appear in the SNS as a result of their use. Various efforts have been put on that aspect with the most prominent work being the one of Mika, 2007 [8]. Instead we are focusing on conceptual entities that are manifested in the object models of the existing SNSs. By using agreed-upon Semantic Web formats to describe people, content objects, and the connections that bind them together, SNSs can interoperate by appealing to common semantics [9].

To this end we analyze the existing work in the area of formal specifications of the SNS domain so as to investigate if and how there are ontologies that already capture notions that our approach would like to define.

2.1   Friend of a Friend (FOAF)

The Friend of a Friend (FOAF) project [10] is creating a Web of machine-readable pages describing people, the links between them and the things they create and do. FOAF is a descriptive vocabulary expressed using the Resource Description Framework (RDF) and the Web Ontology Language (OWL). Its purpose is to allow people define their profile, content and relationship to other people and contents so as to interconnect all web resources in a distributed manner.

FOAF's data specification is close to what our approach wants to achieve as it manages to capture basic abstract notions of a social network manifested using web resources, such as a "Person" and its defining properties. However, this proximity in the objectives could be characterized as coincidental given that the concept is rather

different. FOAF is not meant to capture concepts that stem from the technical nature of SNS APIs that SocIoS attempts to unify. In other words, FOAF is meant to capture even intangible concepts, something that is not allowed in SocIoS. Our approach needs a representation that depicts tangible (even in digital format) entities.

2.2   Semantically-Interlinked Online Communities (SIOC)

The SIOC (Semantically-Interlinked Online Communities) Core Ontology [12] provides the main concepts and properties required to describe information from online communities (e.g., message boards, wikis, weblogs, etc.) on the Semantic Web [13]. As with FOAF, SIOC is very close to meeting SocIoS requirements in terms of data specification compatility. The major difference is that SIOC revolves around web communities emphasizing on forums missing some of the SNS concepts.

2.3   Linking Open Descriptions of Events (LODE)

Linking Open Descriptions of Events (LODE) is an ontology for publishing descriptions of historical events as Linked Data, and for mapping between other event-related vocabularies and ontologies [14]. Thus, the compatibility to SocIoS revolves around the notion of Event and the interplay between real world and SNS world events. As in the majority of the SNSs, the event depicts a calendar item that is planned to take place, in SocIoS we are interested in real-world events that generate intense and unusual in-world activity.

3   Application Domain Concepts

As a starting point to our analysis, we employ three application scenarios that stem from real end-user requirements. One scenario is related to journalism and the use of SNS users as a source of information to a multimedia dossier article and two are related to TV commercial production and particularly to the use of social media for location scouting and casting extras. The requirements of these applications are defined by leading companies in this line of business, and in particular, a news broadcasting agency and a TV production company. More details about these applications can be found on the website: www.sociosprject.eu.

All three applications relate to the exploitation of abstract social networks created on the fly based on demographic and social criteria (e.g. credibility, influences, etc), for finding and retrieving information in a reliable, timely and cost-efficient manner. These applications were analyzed so as to define the SocIoS domain of discourse and to have a reference conceptual classification of our universe.

The conceptual analysis of the abovementioned scenarios leads to the following conclusions:

The SNS **users** form a central concept. Their **activity** is important to the end-users who want to monitor it. From that activity the news broadcaster can collect details about events or the producers find locations or people of interest at any particular

moment they are desire. For that reason, SNS users can be classified in **groups** so as to allow the end-users better monitor their activity. Typically, the problem of all social web applications is **reliability**. Reliability must be measured somehow and hints be provided to the end-users. Furthermore, the basic functionality has to support journalists or producers to look for **multimedia items** which are owned by SNS users. Multimedia items can steer end-users' interest if they directly fit to their criteria or if they emerge as central objects in occasions of intense and unusual activity, i.e. an **event**. Typically, an event is taking place at a **location** and this affects greatly the definition of events of interest of the end-users. Finally, communication between SNS users and end-users is achieved through the basic communication channel, which is the **message**.

Based on this short summary, we conclude to the a small number of common domain concepts, i.e: User, Activity, Group, Reliability, Multimedia item, Event, Location and Message.

In what follows, we explain how we formally capture and extend these notions based on the actual vocabulary used by the object models of the underlying SNSs.

4   SNSs API Analysis

Conceptually, the SNS universe as part of the social web, can be seen as a single social network, in which people are socially connected to each other and the basic social interaction is the sharing of content (text, visual and audio). In practice, this is achieved, if we consider the section of the groups of all the common basic entities of the underlying SNS. This task by itself is rather complex given the highly diverse nature of SNS applications: Communication, Collaboration/authority building, Multimedia, Reviews and opinions, Entertainment and Brand monitoring [15]. The entities "living" in them are generally different with the exception of few basic ones (like the notion of people).

This Section explains in a practical way what the main objects in the SNS API domain are which fit to the descriptions of the domain concepts as they generated from the two scenarios of Section 3. This analysis is pivotal in the definition of the intermediate ontology that will enable the SNS APIs integration and cross-platform application support.

For the analysis we used a set of social media platforms selected based on criteria such as: SNS Popularity, API Popularity, Scope, Openness and Maturity. Following these criteria, the following APIs were studied: OpenSocial, Facebook, Twitter, FlickR and Youtube. A summary of the results of this study are presented on Table 1.

A number of conclusions can be extracted from Table 1. Firstly, these notions differ in the naming as well as in the hierarchy in the formal description (others are first-class objects, while others are mere object properties). Therefore the relationship is not always one-to-one, nor the objects are the exact semantic match.

Table 1. Domain concepts and semantically equivalent objects in selected SNS API (N/S stands for "not supported")

Concept	Social Networking Site API objects				
	OpenSocial	*Facebook*	*Twitter*	*FlickR*	*Youtube*
User	Person	User	Users	people	users
Group	Group	Group and Friendslist	Lists	groups: members	N/S
MediaItem	MediaItems	Photo and Video	Tweets	photos	media
Message	Message	Message	Tweets and direct_ messages	activity: userComments	gd: comments
Activity	Activity	N/S	Timeline	activity	activity
Event	N/S	N/S	Trends	N/S	N/S
Reputation	N/S	Review: Rating	Users: suggestions	stats	ratings
Location	Person/ MediaItem: Location	Album: Location	geo	places	yt: location

Furthermore, not all the concepts are supported by all the SNS APIs (marked as "N/S"). Thirdly, some objects are implicitly related to the concepts, e.g. a Tweet object is semantically equivalent to a Message. Another conclusion is that not all conceptually similar objects are of the same class (Location is second-class in OpenSocial, Facebook and Youtube and first-class in all the rest). Finally, the notion of Events in the way it is defined in this document is only existent in Twitter.

Having defined and examined the domain concepts under an SNS context we can move to the designing of the SocIoS Core Ontology.
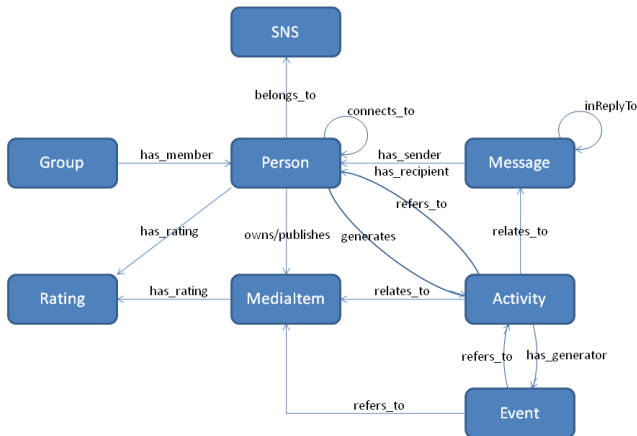
Fig. 1. Overview of simplified SocIoS ontology

5   SocIoS Core Ontology

From the abovementioned analysis, it is clear that the method to conceptually model the domain, is a combination of two efforts: map the underlying SNS object models into a single point of reference and model the SocIoS scenarios universe. As seen in Section 4, other domain concepts can be promoted to classes, whereas others remain as properties. Furthermore, the analysis of the objects and their members also assisted in the identification of the most prevalent data properties.

Given the above we drafted an ontology that captures the basic context and rules behind the SocIoS domain that we attempted to record. The ontology starts by putting Person and MediaItem as the central entities. Then, it analyzes the relationship that bind these two entities with the rest of the core entities. For practical reasons, a new entity is introduced so as to record the SNS in which a Person is active, publishing and sharing MediaItems. The overview of this ontology is presented in **Fig. 1**.

In what follows we present the classes depicted above along with their properties and relevant definitions that exist in the ontology projects analyzed in Section 4. The latter aims to help the reader who is familiar with these initiatives, better digest the semantics of the SocIoS Core Concepts.

5.1   Person

Properties

- connects_to: Other person entities with whom the person in question shares some kind of relationship (is a friend, follows, etc).
- owns: Photos that are published and owned by this person.
- has_rating: A person has a rating assigned by other people
- belongs_to: The SNS in which the Person exists
- birthday: The birthday of this person.
- displayName: The name of this Person, suitable for display to end-users.
- ethnicity: Person's ethnicity.
- gender: The gender of this person.
- location: Description of a person's location.
- name: The complete person's real name.
- tags: A user-defined category label for this person.

Relevant Definitions

Class: foaf:Person The Person class represents people. Something is a Person if it is a person. The Person class is a sub-class of the Agent class, since all people are considered 'agents' in FOAF.

5.2   MediaItems

Properties

- has_rating: Average rating of the media item on a scale of 0-10.
- created: Creation datetime associated with the media item.

- description: Description of the media item.
- duration: For audio/video clips - playtime length.
- last_updated: Update datetime associated with the media item.
- location: Location corresponding to the media item.
- num_comments: Number of comments on the media item.
- num_views: Number of views for the media item.
- num_votes: Number of votes received for voting.
- start_time: For streaming/live content, datetime when the content is available.
- tagged_people: Array of string (IDs) of people tagged in the media item.
- tags: Tags associated with this media item.
- title: Describing the media item.
- type: The type of media.

Relevant Definitions

Propery:foaf:depiction The basic notion of 'depiction' could also be extended to deal with multimedia content (video clips, audio), or refined to deal with corner cases, such as pictures of pictures etc.

5.3   Activity

Properties

- relates_to: MediaItem that may accompany the activity
- has_generator: An activity may be generated or related to an event
- refers_to: The person who triggered the activity
- body: Specifying an optional expanded version of an activity.
- publishes: Any photos, videos, or images that should be associated with the activity.
- postedTime: Specifying the time at which this activity took place.
- title: Specifying the primary text of an activity.

Relevant Definitions

No equivalent notions in other ontologies.

5.4   Events

Properties

- refers_to: a (physical, social, or mental) object involved in an event. In SocIoS this is related to MediaItems or Activity.
- atPlace: a named or relatively specified place that is where an event happened.
- atTime: an abstract instant or interval of time that is when an event happened.
- circa: an interval of time that can be precisely described using calendar dates and clock times.
- illustrates: an event illustrated by some thing (typically a MediaItem object)
- location: an abstract region of space (e.g. a geospatial point or region) that is where an event happened.

Relevant Definitions

Class: LODE: Event "Something that happened," as might be reported in a news article or explained by a historian. An event consists of some temporal and spatial boundaries subjectively imposed on the flux of reality or imagination, that we wish to treat as an entity for the purposes of making statements about it. In particular, we may wish to make statements that relate people, places, or things to an event.

5.5   Messages

Properties

- has_recipients: Person entities who receive the message.
- has_sender: The person who sent the message.
- body: The main text of the message.
- collections: Identifies the messages collection this message is contained in.
- inReplyTo: Use for threaded comments/messages.
- status: Status of the message. (NEW, READ, DELETED)
- timeSent: The time that the message was sent.
- subject: The subject of the message.

Relevant Definitions

Class: sioc: Post A Post is an article or message posted by a UserAccount to a Forum.
    The SIOC Types Ontology Module describes some additional, more specific subclasses of sioc:Post.

5.6   Rating

Properties

- score: The score that depicts the rating
- range: The score range (min, max).

Relevant Definitions

No equivalent notions in other ontologies.

5.7   Group

Properties

- has_member: Person entities who are members of the group.
- title: Title of the group.
- description: Description of group.

Relevant Definitions

Class: sioc:Usergroup A set of UserAccounts whose owners have a common purpose or interest. Can be used for access control purposes. A Usergroup is a set of members or UserAccounts of a community who have a common Role, purpose or interest.

5.8   SNS

The SNS concept captures the notion of the container in which a typical social network is manifested for a person and the mediaitems are shared. Typical examples are Facebook, Twitter, MySpace, YouTube, FlickR.

Properties

- container: The platform SNS, e.g. Facebook, Twitter, etc.

Relevant Definitions

No equivalent notions in other ontologies.

6   Conclusions

This document summarizes the procedure and the results of the conceptualization of the SocIoS domain as an intermediate step to define the SocIoS Object Model (API). The analysis starts from the domain concepts defined by the SocIoS scenarios and continues with the identification of conceptually common data specifications from various APIs that semantically fit to the domain concepts. The result of this work is captured in the SocIoS core ontology that is expected to assist the developers to better comprehend the notions behind the SocIoS object model and extend it as appropriate.

The ontology permits the extension of the SocIoS API towards any direction, either for including new applications or for supporting further SNS APIs. This flexibility is also crucial in cases of strong dependence on third parties because it allows for quick cope with possible changes. Furthermore, we present how the ontology is linked to semantic web specifications, a fact that adds to the extendibility of our proposed solution.

References

1. Gnip, The Social Media API, `http://gnip.com/`
2. Hootsuit, Social Media Dashboard, `http://hootsuite.com/`
3. Stroodle, Your Social Pulse, `http://stroodle.it/`
4. OpenSocial Social Data Specification 1.1, `http://opensocial-resources. googlecode.com/svn/spec/1.1/Social-Data.xml`
5. Web Service Modeling Ontology (WSMO), `http://www.wsmo.org/`
6. W3C SA-REST: Semantic Annotation of Web Resources, `http://www.w3.org/ Submission/SA-REST/`
7. McGuinness, D.L., van Harmelen, F. (eds.): OWL Web Ontology Language Overview. World Wide Web Consortium (W3C) recommendation (February 2004), `http://www. w3.org/TR/owl-features`

8. Mika, P.: Ontologies are us: A unified model of social networks and semantics. Journal of Web Semantics 5(1), 5–15 (2007)
9. Breslin, J., Decker, S.: The Future of Social Networks on the Internet: The Need for Semantics. IEEE Internet Computing 11(6), 86–90 (2007)
10. The Friend of a Friend (FOAF) project, `http://www.foaf-project.org/`
11. Brickley, D., Miller, L.: FOAF Vocabulary Specification v0.98 (August 9, 2010), `http://xmlns.com/foaf/spec`
12. The SIOC project, `http://sioc-project.org`
13. Berrueta, D., Brickley, D., Decker, S., Fernández, S., Görn, C., Harth, A., Heath, T., Idehen, K., Kjernsmo, K., Miles, A., Passant, A., Polleres, A., Polo, L.: SIOC Core Ontology Specification. In: Bojārs, U., Breslin, J.G. (eds.) March 25 (2010), `http://sioc-project.org/ontology#term_Item`
14. Ryan Shaw with contributions from Raphaël Troncy and Lynda Hardman, LODE: An ontology for Linking Open Descriptions of Events (July 7, 2010), `http://linkedevents.org/ontology`
15. From Wikipedia, Article available at, `http://en.wikipedia.org/wiki/Social_media`