The Interoperability of Lightweight Semantics for Social Networks

Matthew Rowe
Web Intelligence Technologies Lab
Department of Computer Science
University of Sheffield, UK
m.rowe@dcs.shef.ac.uk

I. Introduction

This position paper describes the current state of play regarding the use of lightweight semantics on the web to describe social networks, and the interoperability between such formalisations. Our position is that various specifications exist to describe social network information within existing HTML, however there is little crossover and mapping between such formalisations. Cohesion is required between the different working groups responsible for such technologies, spanning both academic and commercial organisations. This paper presents the background area regarding the use of lightweight semantics to describe social networks. Section two explains the problems associated with the current lightweight semantics. Section three presents possible applications that we believe to be achievable based on the success of interoperating lightweight semantics for social networks.

A. Background

Within the Semantic Web community the requirement for accessible machine readable metadata has lead to the creation of steering committees concerned with the formalisation of such information. Formal standards have been agreed on, and knowledge acquisition has reached the one billion triple milestones. We believe the formalisation of social network information can be divided into two distinct areas of specifications: Heavyweight and Lightweight.

1) Heavyweight: Heavyweight specifications describe social network information using explicit semantics. FOAF (Friend of a friend) [Brickley and Miller, 2004] offers a heavyweight specification to capture knowledge depicting a given person's social network. The person in question is able to describe their digital identity through semantic properties such as name, current_location, etc. Social network information is described by establishing a relation between the given person and each person in his/her social network using the *foaf:knows* relation. Similar work bas been carried out within the SIOC project, but with an emphasis on online communities. The SIOC ontology [Breslin et al, 2005] models online interactions and user roles within forums and weblogs.

2) Lightweight: Lightweight specifications are the converse of heavyweight specifications in the nature of their implementation. Web pages written using standard HTML commonly feature social information that is human readable, but not machine readable. Lightweight semantics add machine readable tags to such information to leverage exportation by software agents and automated machine processes. One of the most popular examples of lightweight semantics are Microformats [Khare, 2006]: Microformats allow existing information to be marked up as knowledge using existing XHTML techniques. For example, the hCard Microformat describes a contact card as example 1 shows, by enabling existing information to be marked up semantically. Pages containing Microformats can then be parsed using GRDDL (Gleaning Resource Descriptions from Dialects of Languages) [Connolly, 2007] and XSLT (Extensible Stylesheet Language Transformations) [Clark, 1999] to extract their internal knowledge.

Example 1: Example VCard

```
<div class="vcard">
<span class="n">
<span class="given-name">Matthew</span>
<span class="family-name">Rowe</span>
</span>
<a class="url"
href="http://www.dcs.shef.ac.uk/~mrowe">Homepage</a>
</div>
```

Social network information is described using the XFN (XHTML Friends Network) microformat within XHTML through the link structure of the page. Given a homepage containing links to web pages belonging to friends, the XFN microformat can be included within each of those links to add semantics to the relation. As example 2 demonstrates, the relationship type is described using the *rel* attribute within the link, multiple types can be used to describe the relationship. Other XFN values can used to cover a range of relationship categories such as geographical location, professional, family and friendship.

Example 2: Example XFN

```
<a href="http://music.blogspot.com"
rel="me">Music Blog</a>
<a href="http://www.myspace.com/mattroweshow"
rel="me">Myspace</a>
<a href="http://www.shef.ac.uk/~fabio/"
rel="colleague">Fabio Ciravegna</a>
```

```
<a href="http://www.shef.ac.uk/~paul/"
rel="friend">Paul Richmond</a>
<a href="http://www.shef.ac.uk/~jdb/"
rel="colleague friend">Jonathan Butters</a>
```

The second notable technology using lightweight semantics is RDFa [Adida and Birbeck, 2008], the extension of the existing Resource Description Framework (RDF) [Manola and Miller, 2004] to allow semantic markup to be included within XHTML. RDFa differs from Microformats by allowing the expression of any semantic type within XHTML providing the correct ontology is referenced. Namespaces are defined within the XHTML to load the necessary ontologies, and concepts within the ontology are assigned to elements within the Document Object Model of the web page. This is similar to the representation used within RDF. Example 3 details how the FOAF specification is used to embed semantic markup within existing HTML without altering the presentation of the information itself. In this instance the FOAF specification is used to describe an instance of type *foaf:Person* within the page, and describe the name and email properties associated with the instance. Although this example does not feature such an expression, it is possible to also describe social network information attributed to a given person using the FOAF specification as described in section 1.1 using the *foaf:knows* property.

Example 3: Example RDFa

```
<div typeof="foaf:Person"
xmlns:foaf="http://xmlns.com/foaf/0.1/">
<p property="foaf:name">Matthew Rowe</p>
<p><a rel="foaf:mbox"
href="mailto:m.rowe@dcs.shef.ac.uk">email</a></p>
</div>
```

This use of lightweight semantics, particular RDFa allows powerful expressive ontologies to be employed within web pages to describe social information using formal semantics. Both RDFa and Microformat information can be exported from within the page using GRDDL and XSLT, Google Social Graph[1] has also been developed to utilise existing semantic metadata for exportation, and now new technologies such as Yahoo's Search Monkey[2] intend to make use of lightweight semantics to provide a richer search experience.

II. PROBLEMS

A. Lack of Interoperability

The main issue that we found from our investigation is that currently web standards for describing social networks offers little room for interoperability. We believe that this problem could have ramifications for the future of web standards if no agreed format or mapping is finalised between the existing specifications. To contextualise this, one can focus on the differences between the usage of RDFa and Microformats. In figure 1 exactly the same information is expressed using RDFa and Microformats, yet the need for cohesion between these specifications becomes apparent as the overlap between concepts is noticed. For example, a mapping between hCard and FOAF would ensure efficient knowledge capture from web resources. This problem is currently being investigated by the Diso[3] working group to provide an abstract mapping between alternative social network standards.

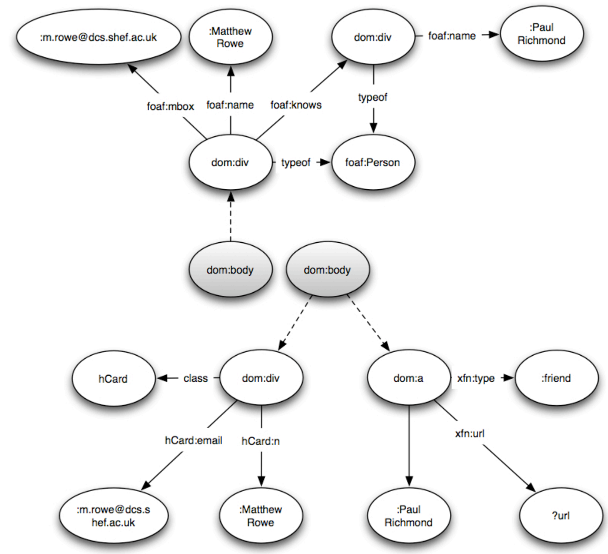

Figure 1. Social Network Information defined by both RDFa and Microformats

B. Identification

Identification is a key issue when concerning interoperability of social networks. As figure 1 demonstrates the content of web pages could differ in their semantic structure, yet describe the same knowledge. Without the email property in each graph it would be impossible to effectively deduce that the person in question was indeeed the same without using inferences based on the available information. The need for identification is apparent in this instance as a unique URI for each given person would alleviate the need for reasoning. We believe the current state of play involving OpenID[4], particularly in the FOAF specification, is a move in the right direction and we have begun work on a social graphing service to combine decentralised social web accounts to a single identity. In the following section we present applications that would take advantage of this approach.

III. APPLICATIONS

A. Persistant Digitial Identity

Based on the advancement of interoprability between social network standards one could imagine a web comprised of citizens, not consumers. Each citizen has their own unique digital identity that is effectively an identity card. This identity card contains important information about their social web presence, notably abstract concepts such as reputation and

[1] http://code.google.com/apis/socialgraph/
[2] http://developer.yahoo.com/searchmonkey/
[3] http://code.google.com/p/diso/
[4] htt://www.openid.net

trust. The identity card would be linked to an OpenID used to assign a unique URI to the owner of the card. Any consumer requesting access to information on the card would require authorisation from the OpenID provider, upon successful authorisation the consumer would be granted access to the required information.

This persistant digital identity would also include a social graph of the owner and weightings between the owner and each member of his/her social network. This would allow the social graph to be transferrable between social web enabled sites. For example, the owner signs up for a new web service that he/she knows his/her friends to also be users of, rather than searching for friends manually, the web service can request access to the social graph of the identity owner. When access is granted, the social graph is returned and the members from the identity owner's social network are automatically added to his/her profile in the service. This process is leveraged by each member of the social network having a unique URI attached to their identity using OpenID. The OpenID would be used in conjunction with OAuth[5] to delegate access to the required service by the consumer, in this instance the social graph requester.

Another application of persistent digital identity allows the identity owner to build up his/her reputation level on one web service and then reuse this reputation level on another service. For example, the identity owner would answer questions on a web based question-answering system on a given topic. The answers that he/she gives are always rated highly and within his/her community the attributed feedback reflects this therefore giving a good reputation on that given topic. The identity owner may wish to sign up to a similar question-answering service and answer questions on the same topic. By using the identity owner's persistent identity it would be possible to transfer the reputation that he/she has built up on the first question-answering service to the second service without the need to start again as a novice and rebuild the reputation level. In this instance the need for identification and interoperability is vital to perform the task. An agreed standard for the formalisation of social network information would encourage sites and services to incorporate such technologies to take advantage of existing information on similar services.

B. Mobile Devices

As the use of mobile devices has risen over recent years, so too has the uptake of third party applications developed specifically for mobile devices. By embedding lightweight semantics within existing web content web browsers would take advantage of this information by producing access to bespoke knowledge services. Mobile devices offer a partcularly novel example of this application due to their GPS enabled technoloy. To contextualise this application, imagine a web user browsing a web site about the city of Sheffield which they are currently visiting. This web page contains RDFa content featuring the Geonames ontology[6] where each landmark discussed on the web page contains a geocoded reference point describing the latitude and longitude of the landmark location. Combined with the GPS technology on the mobile device, it would be possible to automate the process of working out directions and distances to each landmark based on the current position of the user. Additional services could also be offered based on the digital identity of the user. For example if the user is disabled, then only landmarks and sights in Sheffield would be shown that are wheelchair accessible.

A second application of mobile devices would be to perform real-time social networking. The mobile device would load the social graph of the user from his/her digital identity provider. As the user roams his/her town the locations of social network members would be shown on a map. By accessing notification feeds semantically marked up using RDFa and Microformats attributed to each social network member in the local proximity, status updates would be retrieved attributed to each social network member from either a micro-blogging service such as Twitter[7] or a social networking site such as Facebook[8]. The mobile device user would then know exactly what his/her friends were doing and where they were. For example, if a given friend was sat in a coffee shop and their status update read "Relaxing, having a coffee", the mobile device user could reach a decision whether he/she wished to join that person for a coffee. This process could also be automated to suggest locations to go to based on a collection of friends currently residing in that area, all with common inviting statuses. Interoperability would be required in this scenario to allow the aggregation of information between various resources, for example by mapping between status updates given on a micro-blogging site and a social networking service. If work continues to branch out in different directions providing lightweight semantics with very little crossover and cohesion, then the ability to develop applications that utilise such formalisations would be reduced.

REFERENCES

[Adida and Birbeck, 2008] Ben Adida and Mark Birbeck. RDFa Primer: Bridging the Human and Data Webs. http://www.w3.org/TR/xhtml-rdfa-primer/. 2008

[Clark, 1999] James Clark. XSL Transformations. http://www.w3.org/TR/xslt. 1999

[Connolly, 2007] Dan Connolly. Gleaning Resource Descriptions from Dialects of Language. http://www.w3.org/TR/grddl/. 2007

[Breslin et al, 2005] John Breslin, Andreas Harth, Uldis Bojars, and Stefan Decker. Towards Semantically Interlinked Online Communities. The Semantic Web: Research and Applications. 2005.

[Brickley and Miller, 2004] Dan Brickley and Libby Miller. FOAF Vocabulary Specification. 2004.

[Khare, 2006] R Khare. Microformats: the next (small) thing on the Semantic Web? Internet Computing, IEEE. Vol. 10. Issue 1. Pp 68-75. 2006.

[Manola and Miller, 2004] Frank Manola and Eric Miller. RDF Primer. http://www.w3.org/TR/REC-rdf-syntax/. 2004.

[5] http://www.oauth.net

[6] http://www.geonames.org/

[7] http://www.twitter.com

[8] http://www.facebook.com