Social Networks and Data Portability using Semantic Web technologies

Uldis Bojārs¹, Alexandre Passant², John Breslin¹, Stefan Decker¹
¹ Digital Enterprise Research Institute, National University of Ireland, Galway
² LaLIC, Université Paris-Sorbonne / Electricité de France R&D
What’s the problem and how to solve it?
What if I use multiple services and I want to...

- Merge my social networks between various websites
- Invite my friends from a social media website to a new service I’ve just registered
- Move the stuff I have on one service to another (e.g. move all my blog posts, comments, etc. from WordPress.com to “Acme Blogs”)
- Move all my stuff from multiple services to one third-party service or entralise on my own service, e.g. my blog
- See my stuff on a third-party service providing an aggregate view, like FriendFeed, but in an open way
So many social media sites…
Even more services…
It takes a lot of time…

THEN
YOU HAVE TO MAINTAIN
YOUR ACCOUNTS
Filling out your profiles, re-adding your friends…
Uploading posts and content items to “stovepipes”!
Social media sites are like data silos

* Source: Pidgin Technologies, www.pidgintech.com
Many isolated communities of users and their data

* Source: Pidgin Technologies, www.pidgintech.com
Need ways to connect these islands

* Source: Pidgin Technologies, www.pidgintech.com
Allowing users to easily move from one to another

* Source: Pidgin Technologies, www.pidgintech.com
Enabling users to easily bring their data with them

* Source: Pidgin Technologies, www.pidgintech.com
Social networking fatigue

• How many general or niche SNSs are you willing to register and / or interact with?

• “People are getting sick of registering and re-declaring their friends on every site” Brad Fitzpatrick (Aug. 2007)

• Need for a “social graph” with distributed social networks and reusable profiles

• A Bill of Rights for Users of the Social Web (Sept. 07)
  – Ownership
  – Control
  – Freedom

• The Semantic Web can help!
The Semantic Web in brief

- "The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation" - Tim Berners-Lee, James Hendler, Ora Lassila, Scientific American, May 2001

- A common model to describe data in a machine-readable way:
  - RDF (Resource Description Framework)
  - RDF statements are triples (subject predicate object):
    SAW08 isA Workshop .
    SAW08 colocatedWith 11thBIS .

- Common semantics for this data, using ontologies:
  - "An ontology is a specification of a conceptualisation" - Tom Gruber
  - RDFS (RDF Schema)
  - OWL (Web Ontology Language)

- The Semantic Web FAQ:
  - http://www.w3.org/2001/sw/SW-FAQ
The Semantic Web and Web 2.0

- Semantic Web and Web 2.0 could benefit from each other to lead to a better Web, with social and machine-understandable data.
- Many examples:
  - Vocabularies: FOAF to describe people, SIOC to describe their data
  - Semantic Wikis: Semantic MediaWiki, OntoWiki …
  - Revyu.com: A review website based on SW technologies
  - Tagging: The Tag Ontology, SCOT, MOAT
  - Journal Of Web Semantics - Special Issue on Web 2.0
- “I think we could have both Semantic Web technology supporting online communities, but at the same time also online communities can support Semantic Web data by being the sources of people voluntarily connecting things together.” Tim Berners-Lee (ISWC2005 podcast)
Representing people and their relationships

- **FOAF is the main vocabulary used to represent people:**
  - Friend Of A Friend - [http://foaf-project.org](http://foaf-project.org)
  - foaf:Person class:
    - “The foaf:Person class represents people. Something is a foaf:Person if it is a person.”
  - Give yourself a URI:
    - [http://apassant.net/alex](http://apassant.net/alex)
  - Relationships using the foaf:knows property:
    - :John foaf:knows :Alex

- **Extensions using the RELATIONSHIP vocabulary:**
  - [http://vocab.org/relationship/](http://vocab.org/relationship/)
  - All rel:* properties are subproperties of foaf:knows
  - :John rel:worksWith :Uldis
  - RDFS inferencing allows tools to answer queries using foaf:knows when people use rel:* alternatives
Integrating social networks with FOAF

Source: Sheila Kinsella, Applications of Social Network Analysis 2007
A common semantics for existing services

- Existing FOAF exporters for Facebook, Flickr, Twitter…
- Run unified queries using SPARQL
Identity management across networks

• A need to unify URIs from different services so as to represent one's unified identity

• Linked-data principles are to use owl:sameAs and rdfs:seeAlso:
  – See: [http://www4.wiwiss.fu-berlin.de/bizer/pub/LinkedDataTutorial/](http://www4.wiwiss.fu-berlin.de/bizer/pub/LinkedDataTutorial/)
  – owl:sameAs: Used to identify two resources with different URIs as being the same resource
    • :alex owlSameAs flickr:33669349@N00 .
  – rdfs:seeAlso: “More information about this resource can be found here”, can be used by Semantic Web tools such as Tabulator

• Inference using owl:InverseFunctionalProperty:
  – foaf:mbox, foaf:openid, etc. can be used to identify uniqueness for a foaf:Person

• Unifying aspects of a foaf:Person across networks:
  – All relevant relationships are related to one foaf:Person
  – Social Network unification
Distributed social networking with FOAF
Applications for browsing the social (semantic) graph

- FOAFnaut, FOAF Explorer, etc.
- FOAFGear: thanks to common semantics, only 100 lines of code: [http://apassant.net/home/2008/01/foafgear/](http://apassant.net/home/2008/01/foafgear/)
Semantic social networks tools and services

- Browse / re-use your social graph in personal applications
- Merge identities with pre-defined rules
- Tools:
  - Beatnik
  - Knowee
  - SPARQLpress
  - Nepomuk
Combining FOAF and OpenID

• Link to your FOAF profile from your OpenID URL, so that services can get your machine-readable profile when you log-in:

```html
<head><link rel="meta" type="application/rdf+xml" title="FOAF" href="foaf.rdf" /></head>
```

• FOAF + OpenID Scenario
  – Bob creates an account on Networkr, a new social networking website, using OpenID
  – Networkr retrieves the FOAF URI thanks to an auto-discovery link
  – From the FOAF file, it identifies if there are any people already subscribed to Networkr who are listed in Bob’s defined relationships
  – If that is the case, Bob can add them as “local connections”, share data with them, etc. without having to once again search for / add his friends
  – Specific rules:
    • If I know X from Flickr, he / she can see my pictures on Networkr
Social Media Contributions and the Semantic Web
Social Media Contributions

• Lots of user-created content posted on the Web:
  – Blog posts, wiki pages, bulletin board threads
  – Called « Social Media Contributions » or SMC

• Distributed content
  – Blogging platform, photos-sharing website, social bookmarking service ….

• A need for common semantics to
  – Provide a single model for any SMC, wherever it comes from
  – Enables the use of SPARQL queries instead of proprietary APIs
  – Interlink data and find relationships between content
  – From documents to resources, from WWW to GGG
Modeling Social Media Contributions

- **SIOC - Semantically-Interlinked Social Communities**
  - [http://sioc-project.org](http://sioc-project.org)
  - A ontology to represent the activities of online communities on the Web
  - More than 40 applications, mainly open-source
  - W3C Member Submission, June 2007
    - [http://www.w3.org/Submission/2007/02/](http://www.w3.org/Submission/2007/02/)

- **Namespace**: [http://rdfs.org/sioc/ns](http://rdfs.org/sioc/ns)
  - Five top-level classes: User / Role / Space / Container / Item
  - A “SIOC Types” module for Social Web content
  - Action: A user posts an item in a container

- **A Semantic Web citizen:**
  - Reusing and interlinking existing ontologies
  - Not reinventing the wheel (connects to DC, FOAF, etc.):
The SIOC ontology

- The main classes and properties are:
Sample export of SIOC data from vBulletin

<foaf:Document rdf:about=""/>
<dc:title>SIOC profile for "boards.ie"</dc:title>
<dc:description>A SIOC profile describes the structure and contents of a community site (e.g., weblog) in a machine processable form. For more information refer to the &lt;a href=&quot;http://rdfs.org/sioc&quot;&gt;SIOC project page&lt;/a&gt;</dc:description>
</foaf:Document>

<sio:link rdf:resource="http://vbulletin.sioc-project.org/showthread.php?t=1"/>
<sio:num_views>1166</sio:num_views>
<dc:title>quake.ie - WWWBoard</dc:title>
<dc:created>1998-02-12T15:36:23</dc:created>
<sio:has_parent>
    rdf:resource="http://vbulletin.sioc-project.org/sioc.php?sioctype=forum&amp;siod=13"
    rdfs:seeAlso>
    <sio:container_of>
        rdf:resource="http://vbulletin.sioc-project.org/sioc.php?sioctype=post&amp;siod=1"
        rdfs:seeAlso>
      </sio:Post>
    </sio:container_of>
    <sio:container_of>
        rdf:resource="http://vbulletin.sioc-project.org/sioc.php?sioctype=post&amp;siod=2"
        rdfs:seeAlso>
      </sio:Post>
    </sio:container_of>
    <sio:container_of>
        rdf:resource="http://vbulletin.sioc-project.org/sioc.php?sioctype=post&amp;siod=3"
        rdfs:seeAlso>
      </sio:Post>
    </sio:container_of>
  </sio:Forum>
</sio:container_of>
</sio:Thread>
Connecting people and their user accounts

• The sioc:User class:
  – Can be thought of as a virtual representation of any person online, within the context of a given social media website or community
  – A subclass of foaf:OnlineAccount

• foaf:holdsAccount property:
  – “The foaf:holdsAccount property relates a foaf:Agent to a foaf:OnlineAccount for which they are the sole account holder.”
  – Links a foaf:Person to various sioc:User(s)
  – As many sioc:User(s) as required can be linked to a single person
  – One people, various identities

• Users create and manage content:
  – has_creator and has_modifier properties
  – :blogpost123 sioc:has_creator :john
A person and their user accounts

**SiO(S) + FOAF**

- Person
  - holdsOnlineAccount
  - holdsOnlineAccount
  - account_of

- User
  - holdsOnlineAccount
  - Post
  - ChatChannel
  - Weblog

- User
  - Post
  - Weblog
  - MessageBoard

- User
  - Post
  - Weblog
  - MessageBoard

- User
  - Post
  - Weblog
  - MailingList
Using SIOC and FOAF to represent portable data
• Find all content created by someone with a given OpenID URL:
  – Browse someone’s social media contributions posted on various websites using different account names, but for the same person
  – A single SPARQL query, no need to play with different APIs

```
SELECT ?item
WHERE {
  ?person foaf:openid '<$openid>' ;
  foaf:holdsAccount ?user .
}
```
Moving SIOC data between containers

• Exporting data from the original container
  – Use one of the SIOC exporter (or write yours)

• Importing SIOC data is easy:
  – Parse SIOC RDF data (e.g. using ARC2 or RAP for PHP)
  – Convert SIOC data to the content model of the target system:
    • e.g. content and other properties of blog posts and comments
    • Can use SIOC APIs to hold the data model
  – Store data in the target application:
    • The most difficult part 😞
    • Wordpress plugin

Just a Weblog (View site »)

Dashboard  Write  Manage  Links  Presentation  Plugins  Users  Options
  General  Writing  Reading  Discussion  Permalinks  Miscellaneous  Tags  SIOC Import

SIOC Import Plugin - Control Panel

URL of SIOC data to import


Process URL
Summary
Summary

• Data portability is needed to avoid social network fatigue for end-users
  – Many networks, many friends, many contributions
  – Distributed in proprietary data silos

• The Semantic Web provides common models for such needs:
  – Common semantics to represent data wherever it comes from
  – FOAF to represent people and social networks
  – SIOC to represent Social Media Contributions
  – Use standards API and languages (SPARQL) to get your data
Be part of the Social Semantic Web!

- Vocabularies and tools (APIs, producers...) already exist
  - [http://esw.w3.org/topic/SemanticWebTools](http://esw.w3.org/topic/SemanticWebTools)
  - [http://sioc-project.org/applications](http://sioc-project.org/applications)
  - Needs a wider adoption

- Join us and contribute!
  - [http://sioc-project.org](http://sioc-project.org)
  - #sioc on irc.freenode.net
  - sioc-dev on google-groups

- Social Data on the Web
  - Workshop @ ISWC2008
Contacts

• Uldis Bojars
  – uldis.bojars@deri.org // http://captsolo.net

• Alexandre Passant
  – alex@passant.org // http://apassant.net

• John Breslin
  – john.breslin@deri.org // http://johnbreslin.com

• Stefan Decker
  – stefan.decker@deri.org // http://stefandecker.com

• Thanks to Dan Brickley (FOAF) for his valuable comments about this work