



Provided by the author(s) and University of Galway in accordance with publisher policies. Please cite the published version when available.

Title	Social Networks and Data Portability using Semantic Web technologies
Author(s)	Bojars, Uldis; Breslin, John; Decker, Stefan
Publication Date	2008
Publication Information	Uldis Bojars, Alexandre Passant, John Breslin, Stefan Decker "Social Networks and Data Portability using Semantic Web technologies", 2nd Workshop on Social Aspects of the Web (SAW 2008), in conjunction with 11th International Conference on Business Information Systems (BIS 2008), 2008.
Item record	http://hdl.handle.net/10379/398

Downloaded 2024-03-20T10:30:07Z

Some rights reserved. For more information, please see the item record link above.



Social Network and Data Portability using Semantic Web Technologies

Uldis Bojārs¹, Alexandre Passant^{2,3}, John G. Breslin¹, Stefan Decker¹

¹ DERI, National University of Ireland, Galway, Ireland
`firstname.lastname@deri.org`

² LaLIC, Université Paris-Sorbonne, Paris, France
`alexandre.passant@paris4.sorbonne.fr`

³ Electricité de France R&D, Clamart, France
`alexandre.passant@edf.fr`

Abstract. Social network and data portability has recently gained a lot of interest as one of the issues for social media sites on the Web. In this paper, we will show how Semantic Web technologies and especially the FOAF and SIOC vocabularies can be used to model user information and user-generated content in a machine-readable way. Thus, we will see how data and network information can be reused among various services and applications, at almost zero-cost for developers of such tools.

Key words: Social Media, Semantic Web, Web 2.0, Data Portability, FOAF, SIOC

1 Introduction

Social media sites, including social networking services, have captured the attention of millions of users as well as billions of dollars in investment and acquisition. To better enable a user's access to multiple sites, portability between social media sites is required in terms of (1) identification, personal profiles and friend networks and (2) user's content expressed on each site, whether it is about blog posts, pictures, bookmarks or any type of data. Such portability would allow users to easily exchange content between services, or merge and share their social network between various websites. This requires representation mechanisms to interconnect both people and objects on the Web in an interoperable, machine-understandable, and extensible way. The Semantic Web, which is *an extension of the current Web in which information is given well-defined meaning, better enabling computers and people to work in cooperation* [3], provides those required representation mechanisms for portability between social media sites: it links people and objects to record and represent the heterogeneous ties that bind each to the other. The FOAF¹ initiative [8] provides a solution to the first requirement (1), while the SIOC² project [7] can address the latter (2). By

¹ Friend-of-a-Friend - <http://www.foaf-project.org>

² Semantically-Interlinked Online Communities - <http://sioc-project.org>

using agreed-upon Semantic Web formats like FOAF and SIOC to describe people, content objects, and their connections, social media sites can interoperate and provide portable data by appealing to some common semantics. Moreover, the combination of OpenID and FOAF can be used as a backbone for unique identification and profile definition on social media sites, which can in turn be linked to a user's created content via SIOC.

In this paper, we will discuss the application of these technologies to enhance current social media sites with semantics and to address issues with portability between such services. We will show how FOAF and SIOC can provide smart solutions for data portability amongst various social media sites, allowing one to reuse their data and friends networks from one service on other services, as well as interlinking content from one site to another. We will present theoretical aspects as well as scenarios and implementations of such solutions.

2 Overview of Social Network Portability

2.1 Data Portability History

"Social network portability" is the term used to describe the ability to reuse one's own profile across various social networking sites. Brad Fitzpatrick³ spoke from a developer's point of view about forming a *"decentralised social graph"* [9] and discussed some ideas for social network portability and aggregating one's friends across sites. However, it is not just friends that may need to be ported across social networking sites (and across social media sites in general), but identity and content items as well. Soon afterwards, *"A Bill of Rights for Users of the Social Web"* [11] was authored for social websites who wish to guarantee ownership and control over one's own personal information. As part of this bill, the authors asserted that participating sites should provide social network portability, but that they should also guarantee users *"ownership of their own personal information, including the activity stream of content they create"*, and also stated that *"sites supporting these rights shall allow their users to syndicate their own stream of activity outside the site"*. The Social Graph API⁴ from Google is another related effort that provides methods to query aggregated social graph information from the Web. It currently uses formats like XFN and FOAF, which we will talk about later. More recently, the temporary removal of prominent blogger Robert Scoble from Facebook⁵ relaunched interest in data ownership and portability amongst different social media sites. The DataPortability project⁶ was launched in 2007, with members from various organisations including Facebook, Google and Microsoft coming together to discuss portability issues from technical and legal standpoints. The OpenSocial foundation, recently proposed by Google, Yahoo!

³ Founder of the LiveJournal blogging community

⁴ <http://code.google.com/apis/socialgraph/>

⁵ <http://scobleizer.com/2008/01/03/ive-been-kicked-off-of-facebook/>

⁶ <http://dataportability.org>

and MySpace also aims to provide APIs to let developers write social applications that access data and networks from various social media websites.

However, to enable a person's transition and / or migration across social media sites, there are significant challenges associated with achieving such portability both in terms of the person-to-person networks and the content objects expressed on each site. Social media sites should be able to collect a person's relevant content items and objects of interest and provide some limited data portability (at the very least, for their most highly used or rated items). We will refer to these items as one's social media contributions, or SMCs. Through such portability, the interactions and actions of a person with other users and objects (on systems they are already using) can be used to create new person or content associations when they register for a new social media site. Rather than requiring proprietary APIs to access this data from each service, we think that uniform representation mechanisms are needed to represent and interconnect people and objects on the Web in an interoperable, extensible way.

2.2 The Semantic Web and Data Portability

The Semantic Web provides such representation mechanisms: it links people and objects to record and represent the heterogeneous ties that bind us to each other. By using agreed-upon Semantic Web formats, like RDF with existing or new ontologies, to describe people, content objects, and the connections that link them together, social media sites can interoperate by appealing to common semantics. Developers are already using Semantic Web technologies to augment the ways in which they create, reuse, and link content on social media sites, and some of them already provide exports from social networking sites in such machine-readable formats. In the other direction, social media sites can serve as rich data sources for Semantic Web applications. As Tim Berners-Lee said in the ISWC 2005 podcast, Semantic Web technologies can support online communities even as *"online communities ... support Semantic Web data by being the sources of people voluntarily connecting things together"*⁷. Such semantically-linked data can provide an enhanced view of individual or community activity across social media sites (for example, *"show me all the content that Alice has acted on in the past three months"*). Thus, we do not consider Web 2.0 and Semantic Web as opposing candidates, but rather we believe that they can be combined with each other to provide a Social Web where data can be exchanged and interlinked no matter where it comes from[1].

In the next section, we will describe how the Semantic Web, and especially FOAF, can be used to define one's profile and can act as a unique entry point for personal data across different social media sites. We will also place emphasis on how it can be used to define not only personal information, but also decentralised social networks, and how a user could re-use this information within Semantic Web compliant social media websites. The second part of this paper will overview the data portability aspect, thanks to the SIOC ontology that provides a way to

⁷ <http://esw.w3.org/topic/IswcPodcast>

describe all data entries for a given user wherever they come from. We will see through an example how it helps to move data from one platform to another. Finally, we will conclude with various thoughts regarding links between social media sites, the Semantic Web, social networks and data portability.

3 Social Network Representation with FOAF

3.1 Identity and Networking Management Across Social Media Sites

While many social media sites allow people to define their social networks, only a few of them permit users to export their networks so that they can be reused across other applications. Moreover, when this is the case, users have to rely on some specific APIs, which means writing ad-hoc tools for each data provider. The FOAF project provides a way to represent social network data in a shared and machine-readable way, since it defines an ontology for representing people and the relationships that they share. While some sites already offer FOAF export, such as LiveJournal⁸, MyBlogLog⁹ and Hi5.com¹⁰, there are many other social media sites that do not directly expose their data in RDF. However, developers have created different tools to achieve this goal. For example, user profile information is available in RDF thanks to exporters for Flickr¹¹, Facebook¹² or Twitter¹³. In the latter, this complements machine-readable social network descriptions already embedded via microformats in their pages.

Using FOAF, people and relationships can be modeled using these principles:

- each person is represented as a `foaf:Person` instance and may be assigned URI(s), their unique identifiers on the (Semantic) Web;
- each person has various properties, such as a name (`foaf:name`), nickname (`foaf:nickname`) or birthdate (`foaf:birthday`);
- people can be related to each other using the `foaf:knows` property.

For example, the following snippet of code represents one of the author's profiles created from Flickr using FOAF:

```
flickr:33669349@N00 a foaf:Person ;  
  foaf:name "Alexandre Passant" ;  
  foaf:mbox_sha1sum "528b95cc44060ceea571d7498a9fd2c7e3ca8a4c" .  
  foaf:knows flickr:32233977@N00 .
```

Leveraging Semantic Web representations of people and social networks using widely-adopted ontologies such as FOAF allows us to use generic RDF parsers

⁸ <http://livejournal.com>

⁹ <http://www.mybloglog.com/>

¹⁰ <http://hi5.com>

¹¹ <http://apassant.net/blog/2007/12/18/rdf-export-of-flickr-profiles-with-foaf-and-sioc/>

¹² <http://www.dcs.shef.ac.uk/~mrowe/foafgenerator.html>

¹³ <http://sioc-project.org/node/262>

and SPARQL[10] (a RDF query language which recently became a W3C recommendation) to browse and reuse data. Thus, end users can use the same tools to parse their network wherever it comes from. To that extent, FOAF simplifies the process of writing tools for developers of social-networking frameworks, especially since many open-source tools are available for most platforms^{14 15}.

3.2 Merging and Querying Social Networks

As introduced previously, FOAF allows us to describe personal profiles, but it can also be used to represent relationships between people. Since various sites can export a FOAF representation of users and social networks using their own URIs schemes as shown in the previous RDF snippet, there is still a need to merge and consolidate distributed profiles. In order to consolidate URIs, network owners may rely on Semantic Web best practices that suggest the use of the following properties to represent the identity of existing objects [4] (1) `owl:sameAs` is used to identify that two resources are the same in spite of different URIs and (2) `rdfs:seeAlso` is used to let crawlers and Semantic Web browsers such as Tabulator [2] know where to find additional RDF statements about the resource. Many Semantic Web tools also follow Linked Data guidelines¹⁶ and try to dereference instance URIs, thus providing another way to find additional RDF data.

Using these properties in a distributed and open multi social-network context allows people to interlink and unify the various URIs that represent themselves. To do so, people can reference a main FOAF URI which can be described via a hand-crafted or automatically generated FOAF profile which links to other existing profiles (and also to interlink distributed social networks from various platforms), as the following snippet and Fig.1 describes:

```
:me owl:sameAs flickr:33669349@N00 ;
    owl:sameAs twitter:terraces ;
    owl:sameAs facebook:foaf-607513040.rdf#me .
```

Providing such an entry point allows any RDF-compliant tool to browse one's complete social network in a simple way, i.e. retrieving relationships from Flickr, Twitter or Facebook (1) with standard libraries and SPARQL queries and (2) without having to crawl the Web for data since everything can be accessed from one FOAF file. As an example, we provide a simple script that renders a users' complete social network in a user-friendly Flash interface¹⁷. This tool only requires the main URI of the user, and thanks to the interlinkage properties described before, it retrieves other URIs and related social networks to render it, as shown on Fig.2. This application, which requires only a few lines of Python and SPARQL queries to parse the complete network, clearly shows the benefits of using common semantics to describe networks on social media websites.

¹⁴ http://www.mkbergman.com/?page_id=346

¹⁵ <http://www.w3.org/2001/sw/SW-FAQ#tools>

¹⁶ <http://www.w3.org/DesignIssues/LinkedData.html>

¹⁷ <http://apassant.net/home/2008/01/foafgear>

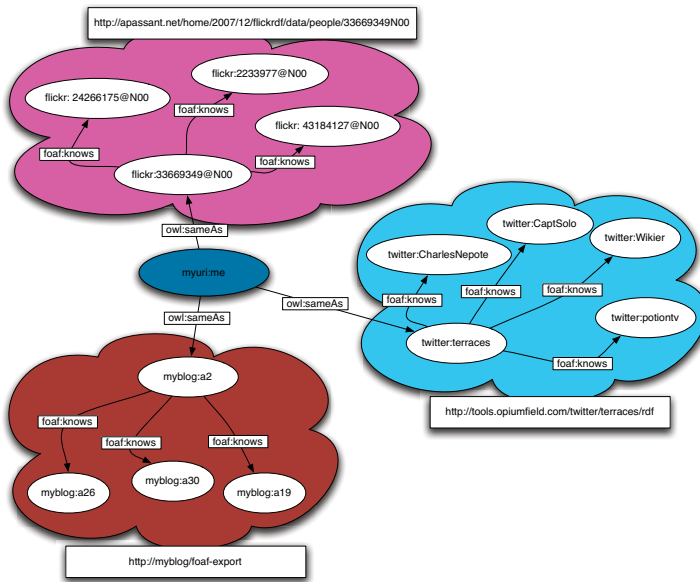


Fig. 1. Interlinking social networks with the Semantic Web

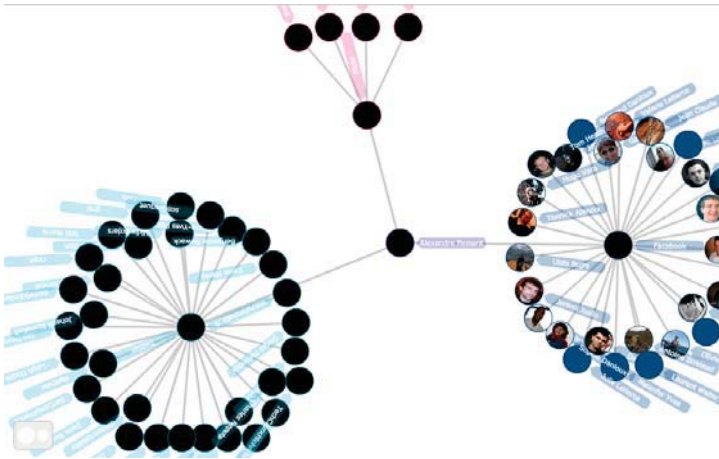


Fig. 2. Browsing a complete social network with a single entry point

Finally, another way to identify uniqueness of someone across various networks is to rely on properties that can uniquely identify him. This is especially useful to merge people among various social networks when they did not explicitly use `owl:sameAs` links. When writing ontologies, OWL offers the ability to define properties as being inverse functional in order to indicate that two RDF descriptions using the same value for this property are "talking" about the same entity. OWL axioms (i.e., `InverseFunctionalProperty`) tell us which

properties can be used in this way, i.e. as indirect identifiers - implementing "reference by description". FOAF uses several properties of this kind, for example, `foaf:mbox_sha1sum` (i.e. scrambled e-mail) and `foaf:openid` (i.e. an OpenID URL). Thus, even if people use different screen names on various websites, as soon as they register with the same email or with the same OpenID URL, they can be uniquely identified across distributed social networks. As an example, the following SPARQL query will retrieve for one user the information about all the people that he or she knows on Flickr and on their weblog (if a contact left some comment there), merging their identities thanks to their `mbox_sha1sum`, whatever their username may be on those platforms.

```
SELECT ?friend ?email
WHERE {
  GRAPH <http://my_flickr_export> {
    :me foaf:knows ?friend .
    ?friend foaf:mbox_sha1sum ?email
  }
  GRAPH <http://my_blog_export> {
    :me foaf:knows ?friend .
    ?friend foaf:mbox_sha1sum ?email
  }
}
```

4 Social Network Portability with FOAF

4.1 Social Networks in Personal Applications

Such semantically-powered social network descriptions can be re-used in existing personal desktop- or web-based tools. For example, Knowee¹⁸ is a web-based application that allows a user to list all of their FOAF URIs, and also includes a microformat parser (which can be used, for example, with Twitter) that features "smushing" capabilities (i.e., identity reasoning), based on user-defined rules as well as on pre-defined ones to uniquely identify people among various social networks descriptions. The network is then browsable using an AJAX-ified user interface. From the desktop point of view, Beatnik¹⁹ provides a semantic address book where you can browse various FOAF profiles and see connections between people. A future development reusing those aspects is the SPARQLPress²⁰ plugin for WordPress, that may be used as a personal social network aggregator based on semantic technologies within this popular blogging tool.

¹⁸ <http://knowee.org>

¹⁹ <https://sommer.dev.java.net/source/browse/sommer/trunk/misc/AddressBook/www/>

²⁰ <http://wiki.foaf-project.org/SparqlPress>

4.2 Reusing Social Networks Across Social Media Site

In order to explain the benefits of such an approach from a portability-between-sites point of view, we will describe the use case of a FOAF-aware social media website.

Bob, a new user, wants to join a (fictional) Networkr service in order to share some pictures and posts with his friends. To do so, he creates an account using his OpenID URL. While the primary advantage with OpenID is that he does not need a new login and password to connect, it allows the site to easily discover his FOAF profile and URI. Indeed, Bob has delegated his OpenID to his own domain name, and added an autodiscovery link in his homepage HTML header to let software agents discover the location of his profile with a single line of code²¹:

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

The system will then retrieve Bob's profile as well as his URI (thanks to the `foaf:openid` property) and read Bob's social networks to check if any people in one of his existing networks are already registered on Networkr. The service will then ask Bob if he wants to consider all those people as friends on Networkr. Since Bob does not want to grant access to everyone about his activities on this website, he decides to check himself who to add from his existing friends. He then adds photos, and decides to restrict access to only those people he had previously added in Flickr. All of these steps have been efficiently achieved by the website since it just has to query a single profile and the related RDF description of Bob's social graph. Moreover, each time he logs in, Networkr again browses Bob's complete network to retrieve updates and change local access rights if needed. Thus, as soon as Bob adds someone as a friend on Flickr, he will gain access to his new picture gallery. Finally, since Networkr is completely open and consider that the data and social graph belongs to the user, it allows Bob to export his new restricted network, which he can then reuse on other websites. Privacy issues should be considered in those uses cases, to allow more complex access rights definition or restrictions, for example if a user wants certain kind of pictures to be seen only by a subgroup of his Flickr network. Identification and trust may also be a problem to display only relevant information when requesting RDF data from Networkr and the recent RDF authentication discussion²² could be considered here.

²¹ <http://wiki.foaf-project.org/Autodiscovery>

²² http://blogs.sun.com/bblfish/entry/rdfauth_sketch_of_a_buzzword

5 Data Portability with SIOC

5.1 Describing Social Media Contributions using SIOC

The SIOC initiative was initially established to describe and link discussion posts taking place on online community forums such as blogs, message boards, and mailing lists. As discussions begin to move beyond simple text-based conversations to include audio and video content, SIOC has evolved to describe not only conventional discussion platforms but also new Web-based communication and content-sharing mechanisms [5].

In combination with the FOAF vocabulary for describing people and their friends, and the Simple Knowledge Organization System (SKOS) model for organising thesaurus-like data, SIOC lets developers link user-created content items to other related items, to people (via their associated user accounts), and to topics (using specific "tags" or hierarchical categories). Through its SIOC Type module, SIOC can represent various types of containers (i.e. `Wiki`, `Blog`, `MessageBoard`) and content items (i.e. `WikiArticle`, `BlogPost`, `BoardPost`). Moreover, this is not limited to textual content because SIOC can be also used to represent content such as `ImageGallery` in the example of the Flickr exporter. Finally, as a good Semantic Web citizen, SIOC reuses and extends existing ontologies such as Dublin Core and FOAF in order to be compatible with RDF data modeled using other existing vocabularies²³.

Various tools, exporters and services have been created to expose SIOC data from existing online communities²⁴. These include APIs for PHP, Perl, Java and Ruby, data exporters for systems like WordPress, Drupal, phpBB and BlogEngine.NET, data producers for RFC 4155 mailboxes, SIOC converters for Web 2.0 services like Twitter and Jaiku, and usage in commercial products including Talis Engage and OpenLink Virtuoso.

All of these data sources provide accurate structured descriptions of social media contributions (SMCs) that can be aggregated from different sites (e.g. by person via their user accounts, by co-occurring topics, etc.). Fig.3 shows the process of porting SIOC data from various sources to SIOC import mechanisms for WordPress and future applications. We will now describe the SIOC import plugin for WordPress.

5.2 Importing SIOC Data, with a WordPress Example

The SIOC import plugin²⁵ for WordPress blog engine is an initial demonstrator for social media portability using SIOC. SIOC import panel in the WordPress administrator user interface (Fig.4) allows a weblog maintainer to import user-created content from another website (described in the form of SIOC data) to their weblog.

²³ SIOC Ontology: Related Ontologies and Vocabularies - <http://www.w3.org/Submission/sioc-related/>

²⁴ <http://rdfs.org/sioc/applications/>

²⁵ http://wiki.sioc-project.org/w/SIOC_Import_Plugin

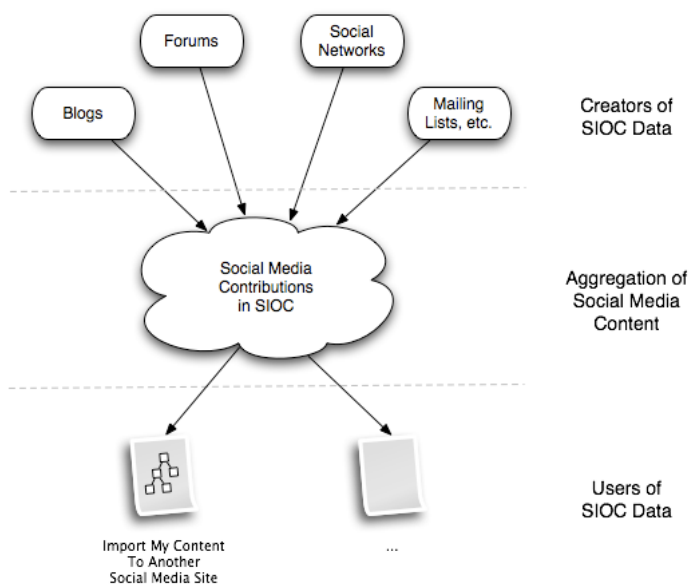


Fig. 3. Porting social media contributions from data providers to import services

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

Fig. 4. Importing SIOC data in WordPress

Data to be imported can be created from a number of different social media sites using SIOC export tools (as described above) which are at the data creation side of the SIOC "food chain" [6]. Since the only requirement for the importer is to "understand" data modeled with SIOC, it makes no difference whether the data comes from another WordPress blog, a vBulletin message board or your latest updates on Twitter.

For example, a SIOC exporter plugin for a blog engine would create a SIOC RDF representation of every blog post and comment, including information about:

- the content of a post (`sioc:content`)
- the author (`sioc:has_creator`)
- the creation / update date (`dct:created` / `dct:updated`)
- tags and categories (`sioc:topic`)
- all comments on the post (`sioc:has_reply`)
- information about the container blog (`sioc:has_container`)

The use of RDF for data representation enables us to easily extend this data model with new properties when they become necessary, even with needs that are not covered by SIOC itself but that one social media site would require to achieve certain task. We thus benefit from a model which is well-formalised but completely extensible.

The import process implemented by the WordPress SIOC import plugin is the following:

- Parse RDF data (using the open-source ARC²⁶ RDF parser)
- Find all posts - instance(s) of `sioc:Post` - which exhibit all of the properties required by the target site
- For each post found, it creates a new post using WordPress API calls

The "proof of concept" implementation of SIOC imported worked with a single SIOC file and imports all the posts contained within it. Fig.5 shows an example post imported into WordPress.

Since SIOC is a universal data format and is not specific to any particular site, this pilot implementation already allows us to move content between different blog engines or even between different kinds of social media sites. However, more functionality is needed for data portability and the next iteration of WordPress SIOC importer uses the `sioc:has_reply` property to identify, retrieve (i.e. fetch additional SIOC RDF files describing these comments) and re-create all the comments associated with the blog posts imported. This approach can be further extended by processing all other kinds of objects and information described in source SIOC data.

5.3 Data Portability for a Complete Social Media Site

We will now describe how a SIOC import tool can be extended to port all user-created content from one social media site to another. By starting from a site's main SIOC profile, we retrieve machine-readable information about all the content of this site - starting with the forums hosted therein, and then retrieving the contained posts, comments, and associated users. This extended SIOC import tool retrieves all SIOC data pages (possibly limited by user-defined filters) and to re-create all the data found in this SIOC page on the target social

²⁶ <http://arc.semsol.org>

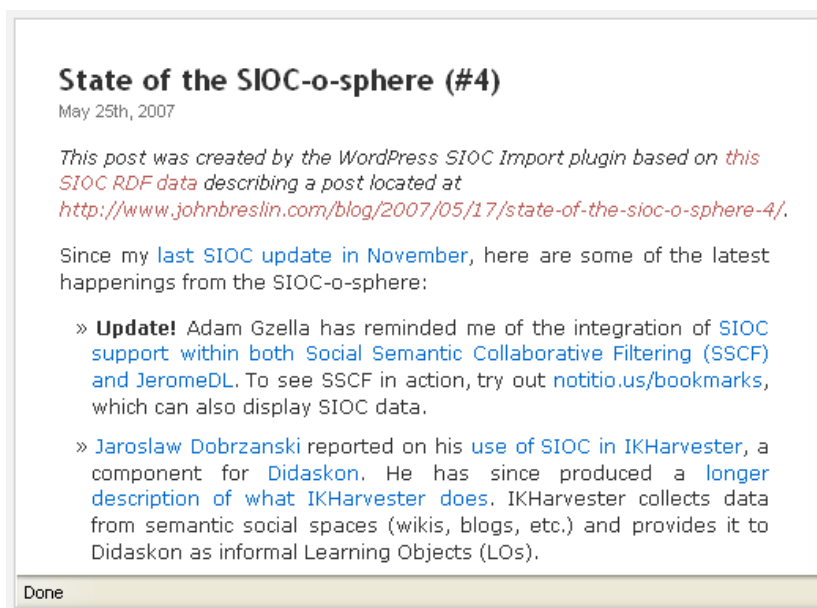


Fig. 5. Imported post in WordPress

media site. Just as can be done with FOAF for social networks, the complete RDF description of a social media site can be found by pointing to a single entry point - a SIOC site profile.

A result of importing the SIOC data for a whole site will be a replica of the original site, including links between objects (e.g. between posts and their comments). Often, a part of the content that a user wants to port is not for public consumption (e.g. if a user is porting some personal information between his accounts). SIOC can be used in this case, but the user will first need to authenticate at the source site and ensure that they have enough privileges to access all the data that need to be migrated.

Another step in social media portability is keeping two sites synchronised (if required): having the same set of users, posts, comments, category hierarchies, etc. In principle, this can be achieved by importing a full SIOC dataset and then monitoring SIOC data feeds for new items added (some SIOC export tools may need to be extended to do this). Implementing this in practice will undoubtedly unfold some interesting challenges, such as real-time synchronisation in two directions, as well as the choice of pushing data or letting services find the data on the Web and update it themselves.

5.4 Perspectives of SIOC Data Portability

An interesting use case for SIOC data portability would be migration between different platforms. For example, this could occur if a person has been using a

mailing list for a particular community, and they then decide that the extended functionality offered to them by a Web-based bulletin board platform is required. Once again, since SIOC can be used to represent various types of containers and items, but using the same format and based on the same high level concepts in the SIOC ontology, moving a mailing-list content to a bulletin board, or a blog post comment to a wiki page can be easily achievable.

While existing web feeds can offer some data portability, the "sliding window" metaphor (providing only last 10-15 items) underlying RSS/Atom contrasts somewhat with the goal of creating a complete archive. SIOC goes further by providing a complete format for representing social media site's data and potentially offering a full archive of site's content. At the same time there remains an overlap in scope here between SIOC and Atom in particular, given that Atom also offers a rich data access and editing protocol.

The discussion-type content items are not the only kind of items that can be ported. The SIOC Types module²⁷ extends SIOC to be able to describe various Web 2.0 and social media content types. Different types of content items (**Sound**, **MovingImage**, **Event**, **Bookmarks**, etc) can be organised in `sioc:Container(s)` and ported in the same way. While the example in the previous section was based on a WordPress implementation, any social media web site could use the same process to import data from one service to another. For example, using a common data format, a user can post on his blog, others can reply in comments, and then the whole discussion can be moved to a bulletin board or imported as a discussion thread associated with a photo on an image sharing site such as Flickr.

In some use cases, data portability may require selective importing of a specific kind of objects. E.g. a user may decide to port information about all videos (`sioc:t:MovingImages`) from his website to Youtube, while moving his `sioc:t:Bookmarks` to a bookmark manager such as del.icio.us. What should be kept in mind is that the latter (bookmarks) can be fully expressed in RDF while the former (videos) have a "payload" that will be separate from RDF data and may require specific handling in order to archive it or to transfer from one site to another.

We mainly discussed porting data from one location or site to another, but we can also consider a wider context - personal "life-stream" or activity stream information. Such streams describe all kinds of activities performed by users, including (but not limited to) creation of content items and social network relations. One interesting application using such streams (and of SMC data in SIOC as one kind of activity data) can be "life-stream" archiving where a person retrieves and keeps an archive of all her activities on different social media sites, which can both keep a permanent personal archive and allow different kinds of personal data applications built on top of it.

²⁷ <http://rdfs.org/sioc/types>

6 Conclusions and Future Work

In this paper we began by introducing the various challenges regarding social network and data portability and how the Semantic Web can be used to help with achieving this goal. We first introduced the use of FOAF to represent identity and distributed social networks, as well as ways to reuse it across various personal applications or various social media sites. We then demonstrated how SIOC data can be used to represent and port the diverse social media contributions being made by users on various sites. Consequently, both approaches can be merged to retrieve and identify all content objects produced by a single user on various social media sites and services, as well as their friends or social network data, as shown on Fig.6, thus leveraging the transfer of both people and content data between Web 2.0 sites using the Semantic Web.

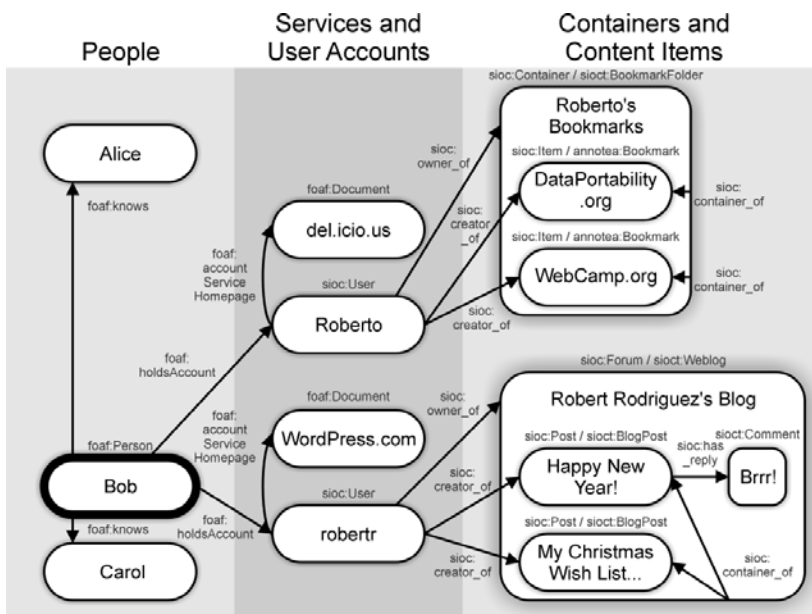


Fig. 6. FOAF, SIOC and Data Portability

For future work, an important issue is who should be allowed to reuse certain data in other sites (as spam blogs are often duplicating other people's content without authorisation for SEO purposes), and what information do people "own" on a social media site and are allowed to port elsewhere. As well as collecting a person's relevant content objects, social media sites may need to verify that a person is allowed to reuse data / metadata from these objects in external systems. This could be achieved by using SIOC and FOAF as representation formats, aggregating content items created by a person (through her user accounts) from various sites, and combining this with some authentication, signature and

trust mechanisms to verify that these items can be reused by the authenticated individual on whatever new sites they choose. When porting content created by others the information about content licenses (e.g. Creative Commons licenses in RDF²⁸) will need to be added to SIOC RDF data and taken into account when porting data.

7 Acknowledgments

This material is based upon work supported by Science Foundation Ireland under grant number SFI/02/CE1/I131. Authors would like to thank Dan Brickley, one of the creators of the FOAF vocabulary, for his valuable feedback and suggestions for this paper.

References

1. Anupriya Ankolekar, Markus Krötzsch, Duc Thanh Tran, and Denny Vrandečić. The Two Cultures: Mashing up Web 2.0 and the Semantic Web. *Journal of Web Semantics*, 6(1), FEB 2008.
2. Tim Berners-Lee, Yuhsin Chen, Lydia Chilton, Dan Connolly, Ruth Dhanaraj, James Hollenbach, Adam Lerer, and David Sheets. Tabulator: Exploring and analyzing linked data on the semantic web. In *Proceedings of the 3rd International Semantic Web User Interaction Workshop*, 2006.
3. Tim Berners-Lee, James Hendler, and Ora Lassila. The Semantic Web. *Scientific American*, 284(5):34–44, May 2001.
4. Chris Bizer, Richard Cyganiak, and Tom Heath. How to Publish Linked Data on the Web. <http://sites.wiwiss.fu-berlin.de/suhl/bizer/pub/LinkedDataTutorial/>, 20 July 2007.
5. Uldis Bojārs, John Breslin, Aidan Finn, and Stefan Decker. Using the Semantic Web for Linking and Reusing Data Across Web 2.0 Communities. *The Journal of Web Semantics, Special Issue on the Semantic Web and Web 2.0 (Forthcoming)*, 2008.
6. Uldis Bojārs, John G. Breslin, Vassilios Peristeras, Giovanni Tummarello, and Stefan Decker. Interlinking the Social Web with Semantics. *IEEE Intelligent Systems*, 23(3), May/June 2008.
7. J.G. Breslin, A. Harth, U. Bojars, and S. Decker. Towards Semantically-Interlinked Online Communities. In *Proceedings of the Second European Semantic Web Conference, ESWC 2005, May 29–June 1, 2005*, Heraklion, Crete, Greece, 2005.
8. Dan Brickley and Libby Miller. FOAF Vocabulary Specification. Namespace Document 2 Sept 2004, FOAF Project, 2004. <http://xmlns.com/foaf/0.1/>.
9. Brad Fitzpatrick and David Recordon. Thoughts on the social graph. <http://bradfitz.com/social-graph-problem/>, 08 2007.
10. Eric Prud’hommeaux and Andy Seaborne. SPARQL query language for RDF. W3C Recommendation, W3C, January 2008.
11. Joseph Smarr, Marc Canter, Robert Scoble, and Michael Arrington. A bill of rights for users of the social web. <http://opensocialweb.org/2007/09/05/bill-of-rights/>, 4 September 2007.

²⁸ <http://creativecommons.org/ns>