



The Future of Social Networks on the Internet

The Need for Semantics

John Breslin and Stefan Decker • Digital Enterprise Research Institute, Galway

"I read somewhere that everybody on this planet is separated by only six other people. Six degrees of separation between us and everyone else on this planet. The President of the United States, a gondolier in Venice, just fill in the names ... It's not just big names – it's anyone. A native in a rain forest, a Tierra del Fuegan, an Eskimo. I am bound – you are bound – to everyone on this planet by a trail of six people."

—John Guare¹

Everyone on the Internet knows the buzzword *social networking*. Sites such as Friendster, Facebook, Orkut, LinkedIn, Bebo, and MySpace, as well as content-sharing sites that also offer social networking functionality (including YouTube, Flickr, Upcoming, del.icio.us, Last.fm, and 43 Things) have captured the attention of millions of users and millions of dollars from venture capitalists. (For more on the topic, see the articles on social search starting on p. 13 in this issue of *IC*.) Compete.com states that, as of November 2006, the 10 most popular domains accounted for about 40 percent of all page views on the Web,² and nearly half of those views were from the social networking services (SNSs) MySpace and Facebook.

SNSs usually offer the same basic functionalities: network of friends listings (showing a person's "inner circle"), person surfing, private messaging, discussion forums or communities, events management, blogging, commenting (sometimes as endorsements on people's profiles), and media uploading. With such features, SNSs demonstrate how the Internet continues to better connect people for various social and professional purposes. Yet, fundamental problems with today's SNSs block their potential to access the full range of available content and networked people online. A possible solution is to build semantic social networking into the fabric of the next-generation

Internet itself – interconnecting both content and people in meaningful ways.

Where Does the Trend Come From?

From the beginning, the Internet was a medium for connecting not only machines but people. Email, mailing lists, Usenet, and bulletin boards allowed people to connect and form online social networks, typically around specific topics. Although these groups didn't explicitly define social networks, the ways people acted and reacted did so implicitly. The early Web continued the trend. More recently, sites such as Friendster and LinkedIn have brought a different notion of online communities by explicitly facilitating connections based on information gathered and stored in user profiles.

In addition to relationship management, social networks are sometimes used for viral marketing,³ although recent results indicate that this might be less effective than often assumed. For example, Karin Knorr-Cetina reports that "the additional purchases that resulted from recommendations are just a drop in the bucket of sales" and that "marketers should take heed that even if viral marketing works initially, providing excessive incentives for customers to recommend products could backfire by weakening the credibility of the very same links they are trying to take advantage of."⁴

Current Social Networking Sites Are Boring...

Social networks exist everywhere around us – at workplaces as well as within families and social groups. They're designed to help us work together over common activities or interests, but anecdotal evidence suggests that many SNSs lack such common objectives.⁵ (See, for example, www.russellbeattie.com/notebook/1008411.html.) Instead, users

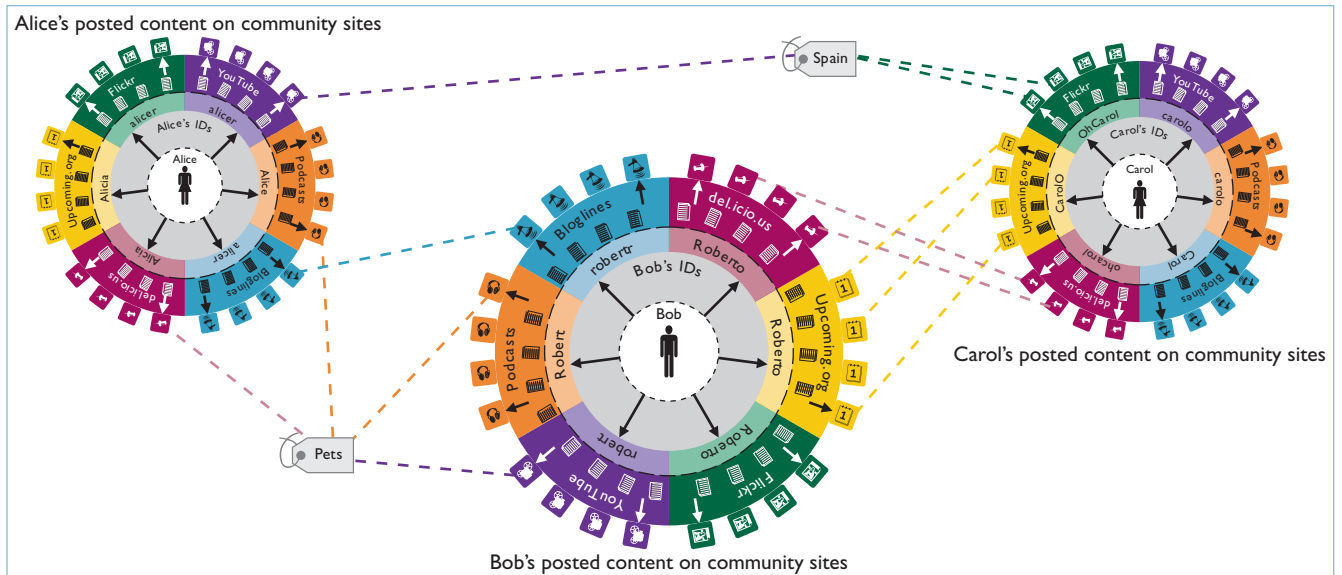


Figure 1. Object-centered social network. Users form social networks (using their possibly multiple online accounts) around the content items they act on — here, on the Web 2.0.

often connect to others for no other reason than to boost the number of friends they have in their profiles. Many more browse other users' profiles simply for curiosity's sake. These explicitly established connections become increasingly meaningless because they aren't backed up by common objects or activities.

The act of connecting sometimes becomes a site's primary (only) activity. In fact, some sites act simply as enhanced address books; although potentially useful for locating or contacting someone, they provide little attraction for repeat visits. This is a flaw with the current theory. As Jyri Engeström, cofounder of the Jaiku.com microblogging site, put it, "social network theory is good at representing links between people, but it doesn't explain what connects those particular people and not others."⁴ Indeed, many social networking sites are becoming increasingly boring and meaningless.

Another problem is that the various SNSs don't usually work together. You thus have to reenter your profile and redefine your connections from scratch when you register for each new site. Some of the most popular SNSs probably wouldn't exist without this

sort of "walled garden" approach, but some flexibility would be useful. Users often have many identities on different social networks. Reusable profiles would let them import existing identities and connections (from their own homepage or another site they're registered on), thereby forming a single global identity with different views (using systems such as OpenID, for example; www.openid.net).

Providing Meaning to Social Networks

Engeström has argued that social networking sites' longevity is proportional to their *object-centered sociality*⁴ — that is, the degree to which people are connecting via items of interest related to their jobs, workplaces, hobbies, and so on (www.zengestrom.com/blog/2005/04/why_some_social.html). Similarly, Ken Jordan and colleagues advocate *augmented social networks*, in which citizens form relationships and self-organize into communities around shared interests.⁶

One way to develop object-centered sociality on the Web is via people's actions around the content they create together, comment on, link to, or for which they use similar annotations.

Adding annotations to items in social networks (using topic tags, geographical pinpointing, and so on) is particularly useful for browsing and locating interesting items and people with similar interests. Content items such as blog entries, videos, and bookmarks serve as the lodestones for social networks, drawing people back to check for new items and for updates from others in their network.

Figure 1 illustrates an object-centered social network for three people. Bob and Carol are connected through bookmarked Web sites that both have annotated, as well as through events they're both attending. Alice and Bob have matching tags on media items, and they subscribe to the same blogs.

As online connections between people become increasingly intertwined with real-world interests, social networking methods are moving toward simulating real-life social interactions: rather than randomly approaching each other, people meet through things they have in common.

Semantics and Social Networks

Although object-centered social networks can fix one problem (that of

Your Social Graph

Social network portability describes the ability to reuse user profiles across various social networking sites and applications. Considerable ongoing work is examining how to implement such functionality.

- Brad Fitzpatrick, founder of the LiveJournal blogging community, wrote an article in August 2007 from a developers viewpoint, discussing some ideas for forming a “decentralized social graph” to support social network portability and aggregate users’ friends across sites (www.bradfitz.com/social-graph-problem/).
- In September, Joseph Smarr et al. authored “A Bill of Rights for Users of the Social Web” (www.opensocialweb.org/2007/09/05/bill-of-rights/) on how “social Web” sites can help guarantee users’ ownership and control over their personal information.

org/2007/09/05/bill-of-rights/) on how “social Web” sites can help guarantee users’ ownership and control over their personal information.

- In “The World Is Now Closed” (www.danbri.org/words/2007/09/13/194), Friend-of-a-Friend (FOAF) vocabulary cocreator Dan Brickley elegantly discusses why social networking services shouldn’t define our relationships in absolute terms and that even an aggregate social graph cannot be so clearly defined.
- In parallel with Fitzpatrick’s article, colleague David Recordon established a social network portability mailing list to discuss topics such as social network centralization versus decentralization, FOAF, the XHTML friends network

(XFN), the hCard microformat, OpenID single-sign-on, Bloom filters, ownership of published content, categorizing friends and personas, the OpenFriend format, the Social Network Aggregation Protocol (SNAP), aggregation and privacy, and the Extensible Messaging and Presence Protocol (XMPP).

- Six Apart has since announced that it will let people reuse their own social graph data elsewhere (www.sixapart.com/about/news/2007/09/were_opening_th.html).
- An API to leverage social graph data from Google’s SNS, Orkut, is also due in November (www.techcrunch.com/2007/09/21/google-to-out-open-facebook-on-november-5/).

sites becoming boring), the remaining challenge is how to achieve interoperability among SNSs and, ultimately, content-creation facilities on the Web. As more social networks form around connections between people and their objects of interest, and as these object-centered social networks grow bigger and more diverse, more intuitive methods are needed for representing and navigating the information in these networks – within and across social networking sites. Also, to better enable navigation across sites, interoperability among SNSs is required in terms of both the content objects and the person-to-person networks expressed on each site. This requires representation mechanisms to interconnect people and objects on the Web in an interoperable, extensible way.

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 to describe people, content objects, and the connections that bind them together, SNSs 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 networking and media sites. These efforts include the Friend-of-a-Friend (FOAF; www.foaf-project.org) project, the Nepomuk social semantic desktop (<http://nepomuk.semanticdesktop.org>),⁷ the Semantically-Interlinked Online Communities initiative (SIOC; www.sioc-project.org), and ontology-enhanced wikis such as the Semantic MediaWiki (http://meta.wikimedia.org/wiki/Semantic_MediaWiki). Some SNSs, such as Facebook, are also starting to provide query interfaces to their data, which others can reuse and link to via the Semantic Web (www.openlinksw.com/blog/~kidehen/?id=1237).

The Semantic Web is a useful platform for linking and for performing operations on diverse person- and object-related data gathered from heterogeneous social networking sites.⁸ In the other direction, object-centered networks can serve as rich data sources for Semantic Web applications. This linked data can provide an enhanced view of individual or community activity in localized or distributed object-centered social networks (for

example, “show me all the content that Alice has acted on in the past three months”).

The SIOC initiative is particularly aimed at linking related online discussions taking place on platforms such as blogs, message boards, and mailing lists. In combination with the FOAF vocabulary for describing people and their friends, and the Simple Knowledge Organization Systems (SKOS) model for organizing knowledge, SIOC lets developers link discussion posts to other related discussions, people (via their associated user accounts), and topics (using specific “tags” or hierarchical categories). As discussions begin to move beyond simple text-based conversations to include audio and video content, SIOC is evolving to describe not only conventional discussion platforms but also new Web-based communication and content-sharing mechanisms.

As Tim Berners-Lee said in a 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” (<http://esw.w3>

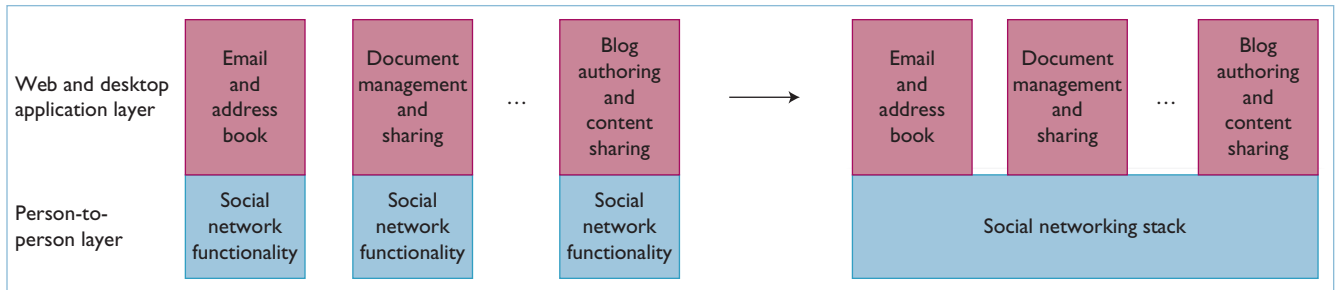


Figure 2. Making social networking a shared component across various desktop and Web applications. Rather than having a fragmented view of one's network in each application, the social networking stack would let users employ all their person-to-person connections in any application.

org/topic/IswcPodcast). Social networking site users are already creating extensive vocabularies and annotations through “folksonomies” (collections of free-text keywords used to tag content items). Because a consensus of community users is defining the meaning, these terms are serving as the objects around which those users form more tightly connected social networks.

Beyond Ego Surfing

So far, SNSs use explicit representations of social networks primarily for visualization and browsing purposes. Yet, some research prototypes show that social networks are actually useful for more than just ego surfing to discover unexpected links in networks of friends. For example, some efforts are under way to examine email filtering and ranking based on social networks.^{9,10} Explicitly represented social networking information can also provide a means for assessing a piece of information's importance and relevance for many other kinds of information filtering (for example, in semantic attention management¹¹) and routing, in general.

Rather than building a separate social networking layer into tools (with all the created maintenance problems), information space and application architects need to fold it into the technology stacks (see Figure 2). Nepomuk does this for the desktop, but given the evolution toward ubiquitous computing and the so-called “Internet of

things,” which will deliver much more information, the Internet infrastructure itself might need to be augmented to include social networking infrastructure to keep users from drowning in an ocean of unconnected and meaningless information. Just as the social semantic desktop Nepomuk provides an operating system layer for representing and exchanging information on the desktop, information creation on the Web and the Internet should take existing connections between content objects and people into account to provide meaning for this information. For example, SNSs might include mechanisms to automate the creation of connections among information items or to route information based on existing relationships between people and content items.

A social networking stack needs to take into account people's relevant objects of interest and provide some limited data portability – at the very least, for their most highly used or rated items. New social networking sites or applications could thus use a person's actions and interactions with other users and objects (exhibiting relevant properties) in existing systems to create new user or group connections when they register. To enable the sharing of existing contacts and to aid in creating new ones, the cross-application social networking stack would require several layers:

1. The *personal authentication and authorization layer* would use Open-

ID, Sxip, or some other single-sign-on mechanism to authenticate that individuals are who they claim to be and that they're authorized to use their social network connections (layer 2) and leverage previously created content items (layer 3).

2. The *social network access layer* would utilize an individual's social networking contacts across various platforms – by collecting FOAF knows relationships from multiple sites, for instance. Access control is required because social connections aren't always bidirectional. For example, Alice could create a connection to Bob in order to view his public content, but Bob might have to approve the connection in the reverse direction if Alice wants to send him a direct message. This layer would ensure that the required directional links exist for various interactions, and it would also verify that the source of the social network information was valid.
3. The *content object access layer* would collect users' relevant content objects and verify that they were allowed to reuse the associated data and metadata in the current application. One way to achieve this would be to use the Semantically Interlinked Online Communities ontology as a representation format, aggregating users' created items (through their accounts) from various site containers. For reputation purposes, this layer would also

The magazine that helps scientists to apply high-end software in their research!



Peer-Reviewed Theme & Feature Articles

2008

Jan/Feb	SDSS Archive
Mar/Apr	Usable Community Grid
May/Jun	Combinatorics in Computing
Jul/Aug	Computational Astrophysics
Sep/Oct	Computational Provenance
Nov/Dec	High-Performance Computing Education

Top-Flight Departments in Each Issue!

- Books
- Computer Simulations
- Education
- Scientific Programming
- Technology
- Visualization Corner



Subscribe to CiSE online at
<http://cise.aip.org/cise> or
www.computer.org/cise

verify that the authenticated individuals had, in fact, created the items on the sites they referenced.

Various architectural alternatives exist for implementing a social networking stack. The existing DNS system offers one possible architecture, but it creates a central point of control. A P2P approach is another worthwhile possibility to explore.

The availability of a social networking stack would also affect existing networking layers because social routing algorithms would be able to deliver information directly to people for whom the information was relevant – simple examples including email filtering and routing with social networks.

Although their meteoric rise might be past, social networks will remain an important part of the Internet. Yet, we believe that form and deployment will evolve toward object-centered networks and – driven by the need to exploit information assessment methods – direct integration into the technology stack of clients (the desktop) and the Internet itself. □

References

1. J. Guare, *Six Degrees of Separation*, Vintage Books, 1990.
2. "Notable Social Networking Web Sites," *Searcher*, Jul. 2007, pp. 36–37; www.infotoday.com/searcher/jul07/Reid_Grey_Web_sitesTable.pdf.
3. J. Leskovec, L.A. Adamic, and B.A. Huberman, "The Dynamics of Viral Marketing," *Proc. ACM Conf. Electronic Commerce*, ACM Press, 2006, pp. 228–237.
4. K. Knorr-Cetina, "Sociality with Objects: Social Relations in Postsocial Knowledge Societies," *Theory, Culture & Society*, vol. 14, no. 4, 1997, pp. 1–30.
5. M. Irvine, "Breaking Up with Social-Networking Sites," *The Seattle Times*, 9 Oct. 2006; http://seattletimes.nwsources.com/html/living/2003292646_onlinebacklash09.html.
6. K. Jordan, J. Hauser, and S. Foster, "The

Augmented Social Network: Building Identity and Trust into the Next-Generation Internet," *First Monday*, vol. 8, no. 8, 2003; www.firstmonday.org/issues/issue8_8/jordan/index.

7. S. Decker and M. Frank, *The Social Semantic Desktop*, tech. report 2004-05-02, Digital Enterprise Research Inst., May 2004; www.deri.ie/fileadmin/documents/DERI-TR-2004-05-02.pdf.
8. T. Berners-Lee, J. Hendler, and O. Lassila, "The Semantic Web," *Scientific Am.*, vol. 284, no. 5, 2001, pp. 35–43; www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21.
9. D. Fisher et al., "Using Social Sorting to Enhance Email Management," *Proc. Human-Computer Interaction Consortium (HCIC 06)*, ACM Press, 2006; <http://research.microsoft.com/research/pubs/view.aspx?type=Publication&id=1600>.
10. J. Golbeck and J. Hendler, "Reputation Network Analysis for Email Filtering," *Proc. 1st Conf. Email and Anti-Spam*, Microsoft Research, 2004; www.ceas.cc/papers-2004/177.pdf.
11. C. Petrie, "Semantic Attention Management," *IEEE Internet Computing*, vol. 10, no. 5, 2006, pp. 93–96.

John Breslin is a researcher and adjunct lecturer at the Digital Enterprise Research Institute at the National University of Ireland, Galway. His research interests include social software, online communities, and the Semantic Web. Breslin has a PhD in electronic engineering from the National University of Ireland. He is a member of the IEEE. Contact him at john.breslin@deri.org.

Stefan Decker is a professor at and director of the Digital Enterprise Research Institute at the National University of Ireland, Galway. His research interests include the Semantic Web, digital libraries, and the social semantic desktop. Decker has a PhD in computer science from the University of Karlsruhe. His dissertation work was quoted as one of the inspirations for the Darpa Agent Markup Language (DAML) program, which spans the Semantic Web effort. Contact him at stefan.decker@deri.org.