

A Topic Sharing Infrastructure for Weblog Networks

Sébastien Paquet
Institute for Information Technology
National Research Council of Canada
Moncton, NB, E1A 7R1, Canada
Sebastien.Paquet@nrc-cnrc.gc.ca

Phillip Pearson
Myelin Electronics
Christchurch, New Zealand
pp@myelin.co.nz

Abstract

Weblogs have recently emerged as a popular means of sharing information on the Web. While they effectively foster the networking of participants on a one-to-one basis, so far they have been lacking the capacity of allowing the establishment of many-to-many communication relationships. This paper describes recent work on facilitating group-forming processes and the sharing of content among weblog authors with shared interests. We have designed, implemented and tested the Internet Topic Exchange, a system that enables weblog posts to be shared among open groups in the form that we call topic channels. After nearly a year of operation, more than 200 topic channels have been created; several of them have been very active and have brought together many participants. This suggests that our approach to enabling weblog authors and topical content to cluster while retaining the advantages of personal publishing is a viable one.

1. Introduction

Personal weblogs [1,2] enable anyone to easily publish and archive their thoughts at very little cost. They are quickly becoming a very popular and dynamic medium for public discussion because they are personal spaces and because they prominently feature hyperlinks that connect individuals and ideas together.

The space of weblogs is currently growing quite rapidly, with thousands of new weblogs being created every day [3]. Newcomers to the blogosphere are confronted with the important problem of finding other webloggers with whom they share interests; the complementary problem is for established webloggers to find new voices that they will find interesting.

Another frequent challenge facing webloggers is to keep up with the conversations around topics or current events of interest when the related ideas are spread across many different weblogs.

We have developed the Internet Topic Exchange (ITE) as a means to address these two issues. The ITE is a

system that combines several Web technologies to allow webloggers to easily associate topics with posts and share these posts by putting them into common pools called topic channels. The idea is for topics to serve as rallying points and for the associated channels to facilitate group-forming processes among webloggers.

As contributing to channels requires additional effort, the challenge here is find a way to ensure that the effort involved be as small as possible, and that potential participants benefit from contributing, preferably in the short term. Our hypothesis was that, for many webloggers, becoming more visible to like minds would be worth the extra effort required in contributing to topic channels.

In what follows we discuss related work, describe the ITE, and present observations on how has been used since its inception in January 2003. The conclusion explores possible future work and questions to explore.

2. Related work

Internet-enabled group discussion is nothing new - mailing lists and USENET forums are popular early examples. In a similar vein, services like QuickTopic.com offer easy to setup discussion spaces that are accessible both via email and the web.

What weblogs offer that is new is the opportunity to directly tie contributions to a personal record, which enables (among other things) participants to get to know each other in a more multidimensional way because they are not subject to the constraint of on-topicalness.

Stated simply, our central idea is to allow related posts to be combined into topic channels. Related contributions that pointed in this general direction are Mower's blogPlex [4], and Gammel's blog network thesaurus [5].

The LiveJournal personal diary system implements "communities" in a manner similar to what we introduce here. However, it works in a closed environment and doesn't feature a wiki component.

A central component in our system is the HTTP-based TrackBack mechanism [6], which allows (among other things) channels to be set up to accept links to weblog posts and publish them together. Several isolated TrackBack-based websites focusing on particular topics have been built, e.g., KMPings [7], Blogpopuli [8],

LazyWeb.org [9], and the Mac OS X conference channels [10]. Setting up channels like these required a certain amount of expertise; with the Topic Exchange, anyone can now very easily create a new channel.

3. The Internet Topic Exchange

3.1. Key components

TrackBack is a framework for peer-to-peer communication and notifications between web sites. The core idea behind TrackBack is that of a TrackBack ping, a request that indicates a relationship between two Web resources. A proper TrackBack target is a resource that can accept TrackBack pings; it is represented by a standard URI called a TrackBack Ping URL.

To send a TrackBack ping, the client makes a standard HTTP POST request to the TrackBack Ping URL, and receives a response in a simple XML format. The POST request features four parameters describing the source of the ping and called title, excerpt, url, and blog_name. Only the url parameter is required; the others provide context for the URL being referenced. The normal behavior when a ping is received is for the target to update itself by adding a link to that URL.

Wikis [11,12] are essentially Web sites where every page is editable by any visitor using a simplified HTML-like syntax. Revision histories are available to allow mistakes to be undone and prevent abuse. Wikis have a long history in the software development community as collaborative hypermedia tools.

The XML-based RSS (Rich Site Summary) standard [13] allows recent additions or updates to websites to be described and published in a convenient format usually called a RSS feed. Using a piece of software called a personal news aggregator, people subscribe to feeds of their choosing and are notified when the corresponding sites update. This makes it possible to keep track of many more sites systematically without having to actually visit them in their browser.

3.2. Operation

Figure 1 illustrates how the ITE functions. Our explanation will begin in the right-hand portion of the figure. The ITE has a number of topic channels, each available either as a HTML weblog-style output or as an RSS feed. Users may keep up with the channels' contents by viewing the weblogs or by subscribing to the RSS feeds that interest them. A sample channel's HTML view is shown in Figure 2.

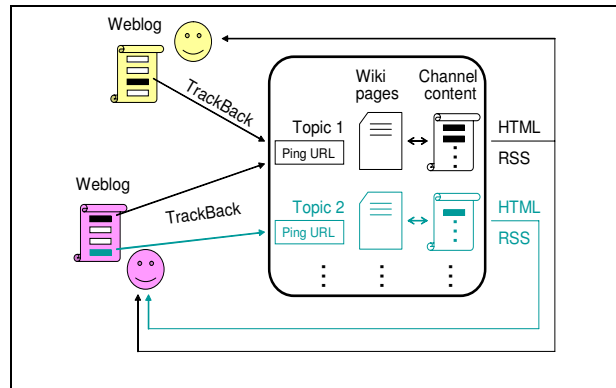


Figure 1. The Internet Topic Exchange.

topicexchange.com:
part of the [Internet Topic Exchange](#)

Social software talk

To post to this channel, [click here](#) or ping the URL http://topicexchange.com/t/social_software/ with your TrackBack-enabled blogging tool.

[About the 'social_software' channel.](#)

Also available in [RSS 2.0 \(XML\)](#).

You are not [logged in](#).

Sun	Mon	Tue	Wed	Thu	Fri	Sat
		1	2	3	4	
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	
Dec ←						

2004-1-1

[collaboration, military style...](#)

from [judith meskill's knowledge notes...](#) (at 2004-01-01 21:30:44, IP 64.71.437.435)

Military Information Technology :: Collaboration Through Technology By Chris Watson ...In 1999, Congress instructed the Department of Defense and the intelligence community to address the lack of interoperability between fielded collaboration tools. The Office of the Secretary of Defense (OSD)...

[Sir Tim of the Web](#)

from [Teledyn](#) (at 2004-01-01 08:19:43, IP 207.228.226.8)

Proving how great honour may still befall ordinary people, Buckingham has bestowed the Knight Commander of the Order of the British Empire for the contribution of "just another program" on Sir...

2003-12-31

[Are enterprises ready for social networking?](#)

Figure 2. A view of the "Social Software" channel.

Each topic channel also has a unique TrackBack Ping URL. When a user writes a post that has to do with a topic that has a channel in the Topic Exchange, he can either let his (TrackBack-enabled) authoring tool ping the corresponding URL, or fill in the form provided on the ITE page for that channel. (The Movable Type weblog tool conveniently enables users to associate TrackBack Ping URLs to particular categories; a ping will subsequently be triggered with every post tagged as being

in the category.) When the Exchange receives the ping, it pushes its content into the appropriate channel.

In addition, each channel has a Wiki page for members of the interest group to use. Generally this will be useful for relevant background material, e.g. for sharing pointers to the most important or up-to-date reference information on the topic, to other good sources or for linking to other, related topic channels. As an example, Figure 3 shows the Wiki page for the channel illustrated above.

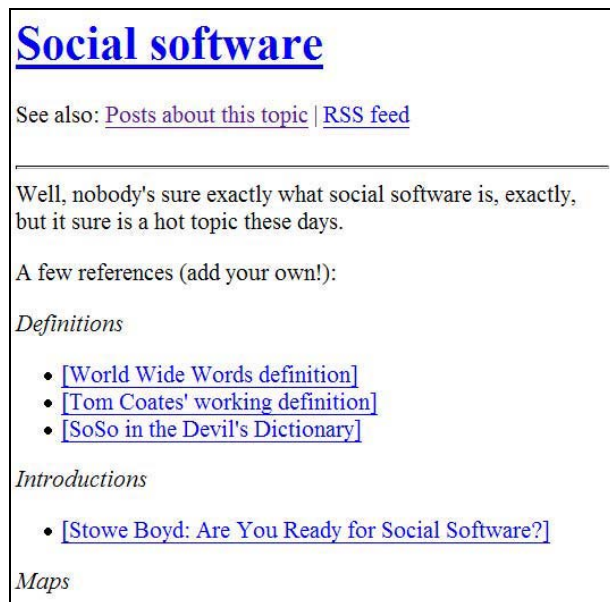


Figure 3. The “Social Software” channel’s wiki page.

Creating a new channel is as easy as visiting a page on the Topic Exchange and providing a name for the channel and an e-mail address to identify the creator.

The current implementation of the ITE was coded in Python. An XML-RPC interface [14] was created, which among other things allows automatic creation of new channels. The Wiki subsystem was implemented by integrating the perl-based UseMod software [15].

The service was set up on Myelin’s servers and launched in mid-January 2003.

3.3. Observations of use

After almost a year of operation, 238 channels have been created. As could be expected, the channels had varying degrees of popularity. At the time of writing the most active channels were:

- “politics”, a channel mainly used for discussing world politics;
- “blog”, a channel used to discuss the weblog medium;

- “social software”, a channel for tracking developments in the field of social software (which encompasses weblogs, wikis, and social networking systems); and
- “google”, a channel for following the action surrounding the Web’s popular search engine.

Figure 4 gives an idea of the activity across the 50 most active topics.

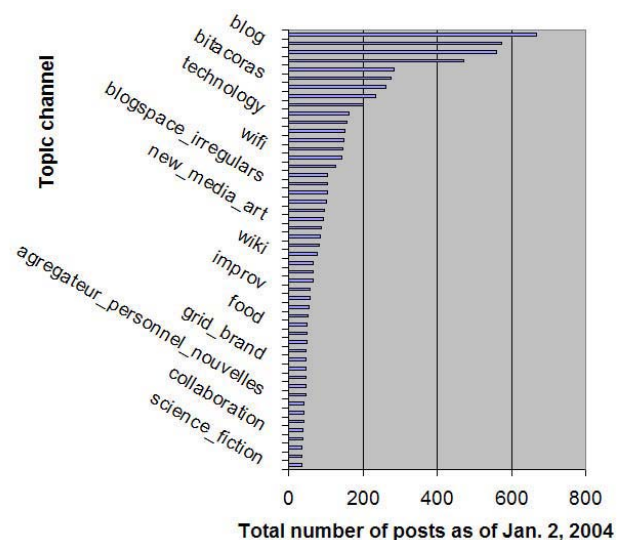


Figure 4. Distribution of posts among the top channels (*not all channel names are shown*).

Although the most successful channels have had many (10 or more) contributors, several others have not (yet?) become popular. Since there’s so little investment involved in starting a group in the ITE, the cost of failure is very low, which invites experimentation. Some topics turn out to be sticky enough to attract participants.

The Wiki pages have mostly been used for short descriptions of the topics they were attached to. There are a few exceptions, however, where the pages have grown into compendium-like resources. Also, one user made the page for the TopicExchange channel into an organized topic directory [16], which was subsequently grown collaboratively by users of the system.

4. Conclusion

At this point in time, the Internet Topic Exchange experiment has indeed enabled a number of loosely joined groups of webloggers to form around shared interests. This shows that some people do have the initiative to start new topics where none existed previously. It also shows that some people will go the trouble of casting some of

their posts to shared channels using the means we have developed, in order to participate in groups.

In order to use this resource, people need to know about its existence. To improve awareness of the ITE and its usefulness, a form of viral self-promotion would be helpful. Currently, the ITE plays a very backstage part. Signs of its existence do not appear in most contributors' weblogs. One way to help this situation would be to automatically have posts that are sent to a channel be identified as such on the originator's weblog, with a direct link to the channel's weblog.

Topic auto-discovery is another interesting avenue. One way this could work would be for weblogging tools to query the ITE with post contents and get back channel suggestions that could be presented to the users. However, there might be a risk in making it too easy to post to channels, as it may dilute their contents.

This work raises a few questions of a more general nature that would be interesting to probe. We cite just two examples here: "Is a service like the ITE more useful in facilitating communication within existing groups, or is it better for supporting new group formation?" and "Can we harness the attractiveness of groups to get people to evolve towards semantic blogging through even more structured content tagging?" [17]. Future work will hopefully illuminate such questions and help weblogs become an even richer Web-based information sharing environment.

5. References

- [1] Aïmeur, E., Brassard, G., and Paquet, S., "Using Personal Knowledge Publishing to Facilitate Sharing Across Communities", *Proceedings of the 3rd International Workshop on (Virtual) Community Informatics: Electronic Support for Communities - Local, Virtual and Communities of Practice*, WWW2003, Budapest, May 2003.
- [2] Lindahl, C., and Blount, E., "Weblogs: Simplifying Web Publishing", *IEEE Computer Magazine*, November 2003, pp. 114-116.
- [3] Sifry, D. "Technorati Growing Pains", *Sifry's Alerts*, November 16, 2003.
<http://www.sifry.com/alerts/archives/000313.html>.
- [4] Mower, M., "More on liveTopics", *Matt Mower's Action Journal*, July 10, 2002.
<http://matt.blogs.it/2002/07/10.html#a193>
- [5] Gammel, D. "Thesauri and Web Logs", *High Context*, May 28, 2002.
http://www.highcontext.com/blarchive/2002_05_28.html#000017
- [6] Trott, B., and Trott, M. "TrackBack Technical Specification", *MovableType.org*.
<http://www.movabletype.org/docs/mttrackback.html>
- [7] <http://www.highcontext.com/kmpings>
- [8] <http://lazyweb.org>
- [9] <http://www.blogroots.com/blogpopuli.blog>
- [10] <http://www.movabletype.org/osxcon>
- [11] B. Leuf, W. Cunningham, *The Wiki Way: Collaboration and Sharing on the Internet*. Addison-Wesley, 2001.
- [12] "Wiki". In *Wikipedia: The Free Encyclopedia*.
<http://en.wikipedia.org/wiki/Wiki>
- [13] Fagan, M. "All About RSS", *FaganFinder.com*.
<http://www.faganfinder.com/search/rss.shtml>
- [14] UserLand Software, "XML-RPC Home Page".
<http://xmlrpc.com>
- [15] Adams, C. "UseMod Wiki".
<http://www.usemod.com/cgi-bin/wiki.pl?UseModWiki>
- [16] http://topicexchange.com/t/topic_exchange/about
- [17] Cayzer, S. "Semantic Blogging", *BlogTalk - European Conference On Weblogs*, Vienna, May 23-24, 2003.
http://www-uk.hpl.hp.com/people/steve_cayzer/papers/030524blogpaper_final.pdf