



Sidebar: Ad-hoc, Yet Organized, Personal Collaboration

Tyler Close, John Recker, Craig Sayers, Ramamurthy Badrinath

HP Laboratories, Palo Alto

HPL-2008-17

February 26, 2008*

email, wiki, blog,
collaboration, access
control,
personalization

In today's workplace, much information is delivered over the Web, while email remains the workhorse of person-to-person collaboration. Our current mail user agents make it easy to get a web page by clicking on a link in an email, but the reverse link to our email about the web page is missing. To provide this missing link, we developed Sidebar, a plug-in that can run inside both the web browser and the file manager. As the user navigates the Web, or a file system, Sidebar provides a summary of user email that links to the currently displayed document, or folder. In using our software, we've found this basic operation results in many emergent properties that make Sidebar a powerful collaboration tool.

Internal Accession Date Only

Approved for External Publication

Submitted to WWW 2008, 21-25, April 2008

© Copyright 2008 Hewlett-Packard Development Company, L.P.

Sidebar: Ad-hoc, Yet Organized, Personal Collaboration

Tyler Close, John Recker, Craig Sayers, Ramamurthy Badrinath
Hewlett-Packard Labs
1501 Page Mill Road
Palo Alto, CA

{tyler.close, craig.sayers, john.recker, ramamurthy.badrinath}@hp.com

ABSTRACT

In today's workplace, much information is delivered over the Web, while email remains the workhorse of person-to-person collaboration. Our current mail user agents make it easy to get a web page by clicking on a link in an email, but the reverse link to our email about the web page is missing. To provide this missing link, we developed Sidebar, a plug-in that can run inside both the web browser and the file manager. As the user navigates the Web, or a file system, Sidebar provides a summary of user email that links to the currently displayed document, or folder. In using our software, we've found this basic operation results in many emergent properties that make Sidebar a powerful collaboration tool.

Categories and Subject Descriptors

D.4.6 [Operating Systems]: Security and Protection – *access controls*. H.1.2 [Models and Principles]: User/Machine Systems – *Human factors, Human information processing*. H.5.4 [Information Interfaces and Presentation]: Hypertext/Hypermedia – *Navigation, User issues*.

General Terms

Security, Human Factors

Keywords

Email, wiki, blog, collaboration, access control, personalization

1. INTRODUCTION

The Web as global information space is now a reality of everyday life. Much information is directly created for the Web and gateway applications bring much of what's left onto the Web. Corporate portals support browsing of LDAP information; document generators make program source code browsable; and new Semantic Web tools make the contents of any relational database browsable. For any piece of information we may need, there's likely a URL that can fetch it for us, and that we can use to refer others to it. But that last step, of sending referrals and comments to others, and viewing comments made by others, brings many choices for how to propagate, store and revise the content and a host of tools, each built to a particular set of choices. Consequently, we tend to use a blend of many tools in our daily work, selecting one or another based on circumstances and the particular content we're about to compose. For example, for text we plan to revise with co-workers, a wiki is often preferred and some of us maintain a blog to keep others abreast of

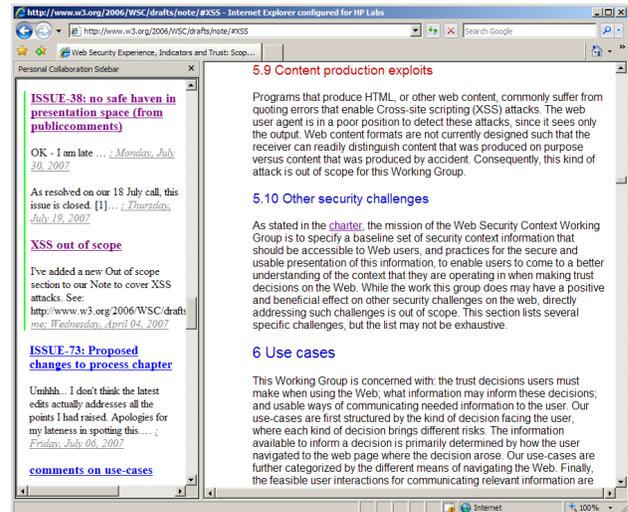


Figure 1: Sidebar at work for the W3C Web Security Context Working Group¹. The browser plugin on the left-hand side summarizes email that links to the currently displayed document. A green highlight runs alongside email referring to the exact section identified by the address bar.

our current activities. However, for various reasons, we often find ourselves back in our email editor.

To improve the productivity of our email collaborations, and better leverage the global information space provided by the Web, we developed Sidebar. The core operation provided by Sidebar is a backlink display containing a summary of user email that links to a currently displayed web page or file system folder. The software can be seen in use in Figure 1, where a W3C editor is using it to track an email conversation about requested changes to a Working Group document. The email displayed in Sidebar comes from the editor's personal email archive. Each message is a normal, run-of-the-mill email written, in fact, by authors unaware of the existence of the Sidebar software. The key enabling detail is that each message links to the document section it is about, as is normal to do when writing about a web page, or other resource on the Web. By searching for these links, a user's personal conversational history about a Web resource can be automatically found and displayed while web browsing. The implementation details for determining this backlink relationship, combined with a large personal email archive, have resulted in emergent properties that make Sidebar a powerful collaboration tool.

In this paper, we compare the features of popular collaboration tools, discuss some innovative features and present Sidebar, a collaboration tool that merges personal email with the Web. Section 2 contains our discussion of collaboration tool features, enumerating those provided by the combination of a Sidebar enabled browser and a mail user agent. The section also

¹ Email author names are omitted from Figure 1. These normally appear just before the date in the grayed out hyperlinks.

sheds some light on why we tend to spend so much time in our email editors. In Section 3, we provide an overview of the application architecture for Sidebar. This design can be deployed one user at a time, requires no new shared infrastructure and doesn't need a critical mass of users before providing valuable functionality. We have developed a beta version of Sidebar for Internet Explorer, Firefox and the Windows File Explorer, which has been in use within HP Labs for several months. Section 4 discusses the origin and effect of the emergent properties that make Sidebar a powerful collaboration tool. In Section 5, we discuss features we plan to implement in the next version of Sidebar. A survey of related work is provided in Section 6.

2. DESIDERATA

In this section, we discuss features we think are crucial to supporting effective collaboration. Some of these features can be found in popular tools, while others are less common. These two groupings are presented in separate sub-sections.

2.1 From wiki, blog or email

A lot of useful collaboration happens using currently popular collaboration tools, such as wiki, blog and email. The wiki is especially adept at facilitating collaboration on a structured presentation of material, such as is found in a traditional reference manual. For example, both the initial Portland Pattern Repository and the modern phenomenon, Wikipedia, excel at developing a taxonomy of related artifacts with supporting text, where both structure and text reflect a consensus view. At the other end of the consensus spectrum, the blog provides an author with a soapbox from which a particular viewpoint, or subject area, can be explored and presented. In a blog, user effort is more directed at maintaining an ongoing dialog, than presenting a reference work. Email provides a durable store for content exchanged over long running, pair-wise, or group, relationships. In many environments, email is the preferred technology for supporting hand-off of material from one participant to another, and keeping an archive of a collaboration's progression.

Each of these tools is tailored to a particular kind of collaboration and so there are differences in the supported feature set. With Sidebar, we aim to leverage each of these technologies in collaborations that span the different focus areas of each tool. For example, Sidebar provides support for email conversations about Wikipedia pages and blog posts. In this section, we discuss the combined feature set and what each of the existing tools contributes to the pot.

2.1.1 Tagging

Popularized by bookmarking sites like del.icio.us and now included in some blog software, tagging marks content as being relevant to a list of topics. When navigating a topic map, content can be found under any topic it has been tagged with. Consequently, a contributor can add new content once and tag it with all the topics it should be displayed under. In contrast, a wiki contributor must separately add their content to each of the pages for which their comment is relevant. In a wiki, content lives inside a created organizational structure; whereas with tagging, the organizational structure and content are separate, allowing content to be reused in multiple places. Furthermore, since content in a wiki is dependent upon the containing topic, the content cannot be added until the topic has. Knowing where best to put a topic in a given hierarchy is tricky work on its own. In our own observation

(and use) of wikis, pages are frequently created at the wrong level in the page hierarchy and can remain poorly connected to the rest of the wiki. Front-end loading content organization before content contribution can be a significant impediment to use for many users.

2.1.2 Attribution

Knowing the author of a text often affects our interest in reading it, our interpretation of the statements made and the follow-up actions we may consider. In the workplace, where different authors have different authority, authorship can be particularly relevant. Email has long provided for attribution, with varying levels of reliability, and a blog is similarly associated with a particular author. In a wiki, authorship can be very difficult for the reader to discern and so crucial filtering and decision making may not happen.

2.1.3 Group editing

In a wiki, the subjugation of content to structure and the loss of attribution is not without reason. A wiki enables the creation of a coherent document that is a mesh of content contributed by many authors. A polished document is not often one of the outputs of an email thread, or blog storm.

2.1.4 Alerts

Certain topics may be of special interest to a reader. When message traffic is low, checking for changes is tedious and easily forgotten work for the user. A monotonous task is best left to a software agent which notifies the user when something new is found. By default, a mail user agent may notify the user of every received message; however, this behavior can often be tamed with the use of email filters. An RSS reader dutifully takes on the work of polling a blog for changes. Some wikis offer RSS feeds for changes to a page, or email alerts.

2.1.5 Notifications

When contributing content, it may be important for a user to know that particular readers will be notified of the contribution. For example, if a supervisor asked the user to help with a particular topic, the user may wish to be assured that the supervisor will be notified of the contribution. Email supports such notification through the "CC:" feature. It is often not feasible to know who is monitoring a blog or wiki.

2.1.6 Controlled content distribution

Many collaborations are not open to everyone, but restricted to a particular group. Even within a group, there may be important distinctions in access to information. Sometimes a side-conversation can be useful even when the topics discussed are not secret. For example, raising a tangential topic to the whole group may derail the group's conversation. In email, such side-conversations are common and are expressed by the user through the addressee list. A wiki typically assumes that all content is accessible to all wiki users. Some wikis support access controlled pages, though the semantics of a web containing publicly accessible links to private pages can be awkward. Typically a blog does not provide a way to direct a blog post to a subset of the blog's readers.

2.1.7 *Controlled content contribution*

Effective group collaboration is a long-standing and difficult challenge. Though it seems perverse, adding more participants, even highly qualified ones, can reduce the effectiveness of a collaboration. Sometimes, restricting the number of participants is the only practical way to cope with this phenomenon. Spam is an extreme case of harmful participation. In email, participation in a thread can be limited through control of the addressee list, though access to the inbox is notoriously open. From the early kill file, to current spam filters, many tools exist to mitigate this effect. A wiki typically treats write access in the same way as read access, and so similar issues arise. A blog restricts blog post creation to the blog's authors, though the blog comment area suffers from the same issues that email and wikis do. Currently, tools for filtering email, and handling email spam, are more mature than corresponding tools for the wiki or blog.

2.1.8 *Offline access*

Never mind the flying cars; somehow the new millennium didn't even bring us pervasive Internet access. Even as Internet access becomes more widely deployed, firewalls and policies prohibiting network use by outsiders, create new online, and yet still offline scenarios. And, of course, our expectations for where Internet access should be available continue to expand. Having your data with you, with the ability to both read it and write it, remains a useful feature. Email is born of this necessity. Typically, a wiki can neither be edited, nor even read, while offline. An RSS reader will often provide offline read access to blogs.

2.2 **Missing in action**

Some important collaboration features have yet to find a foothold in the mainstream toolset. This section discusses innovative features we aim to provide with Sidebar.

2.2.1 *Borrowed structure*

The success of the Web means not only does every piece of information likely have a URL, but that URL is also part of a hypermedia web that organizes that information along with related information. With a plethora of available organizational structures, it should be rare that a new one is needed. It should be possible to organize comments according to an existing structure. For example, it should be possible to reuse a hypermedia web of actors/directors/movies as a navigational structure for a separate database of comments by fans, or reviews by a critic, or planning by entertainment industry professionals. Instead, every wiki must create its own page hierarchy of related concepts; each blog develops its own folksonomy of tags; and email remains largely unorganized.

2.2.2 *Personalized summaries*

A good deal of the comments I've received from others were addressed to me specifically, rather than to a public forum. Though I find these comments interesting, many of them probably wouldn't survive the editing process on Wikipedia, and maybe their authors wouldn't want them to. Nevertheless, I still wish to see Marc Stiegler's comments when looking at the Wikipedia page for the "[Anne](#)" episode of "Buffy the Vampire Slayer". Every user has their own personal email archive, and so every user has text that is relevant to them, but maybe not to others. These comments should also be included in the summaries shown when navigating an organizational structure. Similarly, there are some comments I

find uninteresting. Just as I can delete those comments from my email archive, I should be able to remove them from the summaries created for me. Instead, every wiki page tends towards a homogenized consensus view, or remains in flux between warring users.

2.2.3 *Neighborhood view*

Many topics are closely related and so when viewing one, comments made about the related ones may also be relevant. The accessibility of information can be much improved by including comments about related topics in the summary produced for the main topic. For example, when viewing a web page about a movie, it may be useful to include comments about the lead actors, in addition to comments about the movie. In contrast, comments in a wiki are typically bound to the topic page in which they reside.

2.2.4 *Email centric*

Email is the most widely adopted technology for exchanging substantive comments. Despite whatever other tools a given user may have adopted, a potential new collaborator will expect to be able to use email to exchange comments. Convincing the potential collaborator to first adopt a new technology, or gain access to a new server or database, will be a hindrance to forming the new relationship. To best enable collaboration, email should be used as both the transmission and storage medium for comments.

In addition to user adoption issues, many workers have company or government regulations covering how their communications must be maintained. The infrastructure and process involved in complying with these regulations is substantial. Creating a new compliant infrastructure is a significant undertaking. It is more efficient to leverage the features provided by the existing email infrastructure. Email already provides us with a decentralized content store and users are already trained in how to use it.

3. **IMPLEMENTATION**

The Sidebar implementation currently requires Microsoft Windows. It depends on several interacting components: a browser-based sidebar extension for the Windows Internet Explorer and file explorer, and the Mozilla Firefox browser, a Microsoft Office Outlook add-in, and a web server providing several services including a RDF data store and a SPARQL [SPARQL] endpoint.

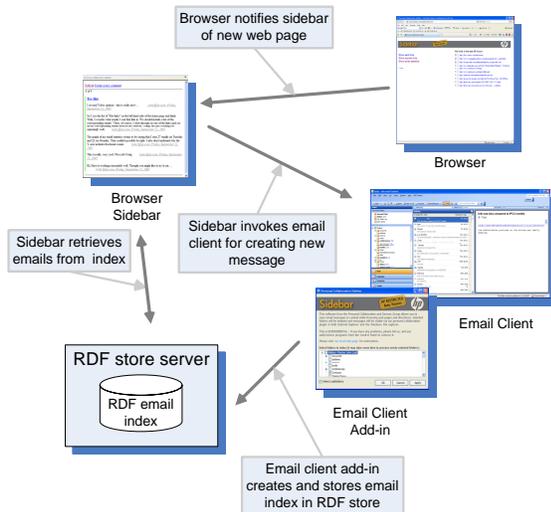


Figure 2: System architecture. Using a SPARQL query, the Sidebar browser plugin searches for related email messages described in an RDF store. An Outlook add-in synchronizes the RDF store with the user's email archive.

3.1 The sidebar

The user-facing component of the Sidebar architecture displays emails relevant to a specified document, or filesystem folder. There are two implementations of the email display: a Microsoft COM object for Microsoft applications and a Mozilla Firefox XUL & Javascript based extension. Irrespective of the low level details of the sidebar implementation, the architecture is the same: the email display is a plugin which renders an HTML and Javascript web page served by the Sidebar server. The sidebar plugin registers an event handler with the envelope application, to receive notifications about a change of target URL. On event receipt, the sidebar re-renders its web page. The sidebar also monitors HTTP re-directs to display emails relevant to both any initial URL as well as intermediate URL's supplied by the underlying service. In these cases, the URL supplied to the web page for email selection is the union of the initial and all intermediate URL's.

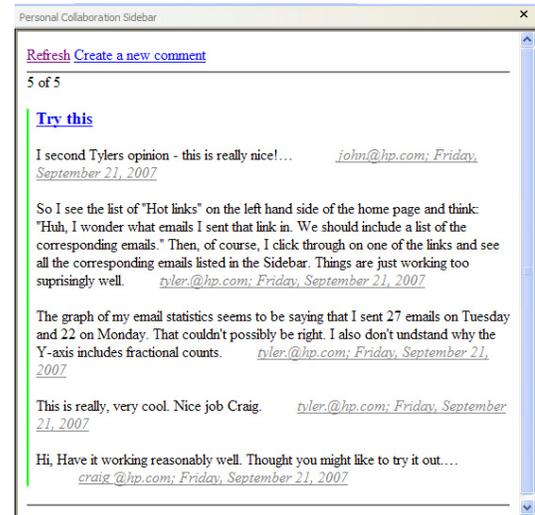


Figure 3: Sidebar. Email messages with the same subject are grouped together in a section. A section presents an abstract of the body of each email, along with a link to the original message. The hypertext for this link provides the author's name and the date the email was sent. The "Create a new comment" link opens a new email editor window, pre-populated with a link to the current document.

The task of the sidebar web page is to query the Sidebar server for email related to the target URL of the envelope application: a web page in the case of a browser, or a folder in the case of the file explorer. The web page uses Javascript to issue a SPARQL query to the Sidebar server to search for messages with "dcterm:references" objects matching any member of the current URL set and process the returned JSON formatted result. Message description summaries for matching emails are then sorted both by subject and chronologically formatted into HTML and displayed within the sidebar.

The sidebar display includes a number of HTML links. At the top of the sidebar are "Refresh" and "Create a new comment" links. The refresh link simply re-generates the sidebar display, potentially discovering newly indexed e-mails. The "new comment" link references a URL that, when invoked, opens a separate email client authoring window from the user's preferred mail user agent with the email body pre-populated with the URL topic for the sidebar. Each sidebar email summary is post-fixed with a link which references a URL that, when invoked, causes the user's preferred mail user agent to open the full email in a separate window.

3.2 Microsoft Outlook Add-in

The task of the Microsoft Outlook add-in is to translate email, contact, and calendar event descriptions from user selected Outlook folders into RDF based descriptions in our RDF data store. It must furthermore respond to folder change events to insure that the RDF data store mirrors the state of the outlook folders.

3.2.1 RDF

Using RDF for the email index worked well for this project for a number of reasons. SPARQL includes a W3C standard HTTP binding [SPARQLP] and Joseki [JOSEKI] a complete Java SPARQL endpoint servlet implementation which provided both a low effort standard based data store server and a convenient

access point for a browser-based client. Terms from the Dublin Core Metadata Initiative [DUBLINCORE] matched closely with a number of fields extracted from the email store. For example, items such as the email author, subject, creation dates, and link URLs are all described with Dublin Core attributes, furthering our goal of re-using standards where possible. In future, we also plan to integrate content from sources other than email by mapping them to the same ontology. The graph structure inherent to RDF made it possible to trivially extend the implementation to include calendar and contact information without rebuilding existing indices. Furthermore, SPARQL naturally supports performing queries that join terms from these different element descriptions.

Figure 4 shows a sample email and Figure 5 its corresponding RDF index description

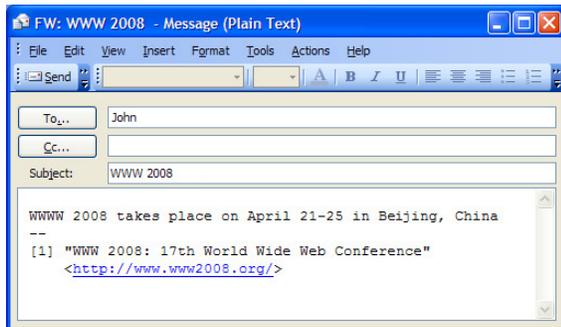


Figure 4: Sample email

```
<?xml version="1.0" encoding="utf-8" ?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:xmles="http://www.w3.org/2001/XMLSchema#"
  xmlns="http://www.hp.com/mms1/sw/email/ns#" />
  <Message rdf:about="x-outlook:3E85BD0A4156A00">
    <dc:creator>john@hp.com</dc:creator>
    <dc:subject>WWW 2008</dc:subject>
    <body>
      WWW 2008 takes place in Beijing, China
      --
      [1] " WWW 2008: 17th World Wide Web Conference"
      &lt;http://www.www2008.org&gt;
    </body>
    <dcterms:references
      rdf:resource="http://www.www2008.org/" />
    <dcterms:dateSubmitted rdf:datatype="xmles:dateTime">
      2007-10-24T21:28:34.84Z
    </dcterms:dateSubmitted>
    <dcterms:dateAccepted rdf:datatype="xmles:dateTime">
      2007-10-24T21:28:35.64Z
    </dcterms:dateAccepted>
  </Message>
</rdf:RDF>
```

Figure 5: Sample email RDF description. Most of the relevant information is expressed in the Dublin Core ontology.

3.2.2 Control Panel

An add-in pop-up control panel allows the user to specify which folders are indexed by the service. Folders specified by the user

through the panel are scanned, and indexed into the RDF data store.

One nice consequence of selecting Outlook as a platform is that Outlook abstracts the email providers for the add-in. For example, although Outlook is often used to access email from Microsoft exchange servers, Outlook also supports other email protocols which allows importing email from popular email providers including Yahoo, Gmail, Hotmail, and AOL.

3.3 The server

A Jetty [JETTY] web server is used to provide three services for our system. For security, requests are restricted to authorized software running on the local machine. The first service is a Joseki servlet which provides a SPARQL endpoint and a persistent RDF data store. RDF data store updates are provided by a second service, which provides a custom interface to allow clients to create, replace and delete RDF records. Lastly, the server serves the HTML and Javascript files used to render email summaries within the sidebar browser.

4. EMERGENT PROPERTIES

The core operation provided by Sidebar is a backlink display containing a summary of user email that links to the currently displayed web page or file system folder. The implementation details for determining this backlink relationship, combined with a large personal email archive, have resulted in emergent properties that make Sidebar a powerful collaboration tool. This section discusses the origin and effect of these emergent properties.

4.1 Web site as email index

The joint popularity of the Web and email results in many email messages containing URLs. When Sidebar is first installed, it indexes the user's existing email archive. Many users have at least one website that they visit frequently and send email about. For example, many of our initial HP labs users have email containing links to the company portal. Others may be heavy users of Wikipedia. Once Sidebar has indexed this email, the web site begins to act as an index into the user's email. Clicking through the pages in the web site refreshes the email summary shown in Sidebar, giving the user quick access to past conversations. These web sites were not designed as email indexes and neither the user, nor his correspondents, attempted to create such an organization. The organization was an emergent effect of using URLs to refer to things in email conversation. Without further implementation effort, this emergent effect provides the desired borrowed structure feature, described in section 2.2.1. The collective effort expended to develop the hierarchy of sites like Wikipedia is reused as an organizational structure for personal email archives.

In effect, the URLs in existing email act as tags, drawn from the folksonomy of the Web. Just as in familiar tagging systems, such as del.icio.us, Sidebar includes an email in a generated summary when viewing any of the web pages whose URL is included in the email. Using this feature, a given email can contribute content to different areas of a particular folksonomy, or to common parts of competing folksonomies. A hypertext folksonomy map, such as provided by a site like del.icio.us, can now act as an email index by including the URL corresponding to a tag in email. Reuse of these existing tagging services provides the desired tagging feature discussed in section 2.1.1.

4.2 Email as document index, and vice-versa

When Sidebar is indexing the user's email it provides special treatment for a URL having a fragment identifier component. The email containing the link is recorded in the index as referring to both the primary resource, the URL without the fragment, and the secondary resource, the URL with the fragment. Consequently, when browsing a document with many anchors, like a large HTML document, email referring to the document itself or one of its sections is displayed in Sidebar. The email messages are grouped according to which section they refer to and a highlight is placed beside the group that refers to the same location as the browser's main window. For example, in the view of sidebar shown in Figure 1, the browser's navigation window is focused on section "5.9 Content production exploits". In Sidebar, the email messages referring to this section are displayed with a green highlight running along the left-hand side edge. Email referring to some other section is also displayed, but not highlighted. The subject of each email is a hyperlink to the document section referred to by the containing email grouping. As a result, the user can navigate the large HTML document by clicking on email subjects in Sidebar. For example, clicking on the hyperlink shown in Figure 1 with hypertext "ISSUE-73 ..." navigates the browser's main window to the section discussed in the email quoted under the hyperlink. The email messages sent and received about the large HTML document act as an index into the discussed portions of the document.

During development of Sidebar, an author of this paper used this feature extensively in his role as an editor on a W3C Working Group. It is the editor's responsibility to collect feedback from Working Group members and integrate their comments into the group's published document. When a member took issue with a section, they would frequently send an email to the group's mailing list which included a link to the offending section. When following this link, Sidebar would show me if the section had been the subject of past controversy. Scrolling through the links shown in Sidebar would walk me through the parts of the document that had been contentious². The members of the Working Group did not know they were helping me create this document index, since Sidebar was not yet an announced project. This document index was again an emergent effect of the Working Group members using URLs to refer to things in email conversation. This particular emergent property is unusual in that it is a yet higher order effect. A document index is provided by the structure created by the email, which was organized according to the document's internal structure.

4.3 Video as email narrator

The email summary shown in Sidebar is updated every time the browser's main window location changes. A web application running in the browser's main window can cause such an update by modifying the fragment portion of the window's current location URL. An interesting use of this technique is an application that updates the window location URL fragment to correspond to the current scene in a video playing in the main window. Consequently, the email group highlight displayed in Sidebar moves in synchronization with the playing video. The user can thus easily track the part of a conversation that refers to the scene they are currently watching. This narration is an

² The size of the thumb in the Sidebar scrollbar is an interesting measure of the work involved in being an editor for the W3C.

emergent effect of assigning URLs to individual scenes in a video. Sidebar knows nothing of video, only URLs.

4.4 Web page + Sidebar = blog

At the top of the Sidebar display is a hyperlink: "Create a new comment". This link launches a new email editor window in the user's preferred mail user agent. The body of the email is pre-populated with the page title and URL currently displayed in the browser's main window. The email soon becomes part of the Sidebar index and will be displayed when viewing the linked to page. Often the recipient of the sent email will respond, leaving the received text mostly unmolested, and so the response email also contains the URL. Consequently, the response email will also be displayed in Sidebar when browsing the referenced web page. This behavior frequently persists for the duration of the email thread and so the complete history of the discussion is shown in Sidebar when viewing the referenced web page. The resulting functionality is similar to that provided by blog software. The web page acts like a blog post and Sidebar acts like the blog post comment area. In this way, any web page can be treated like a blog post. Unlike a blog, readers can control who sees their comment by determining the addressees of their email. If the comment should be seen by a wide audience, the email may be sent to a mailing list. If the comment should only be seen by a select few, the email is only addressed to the select few. Each blog reader therefore also sees a unique view of the blog, a view composed of only those comments they were made party to and decided to keep in their email archive. The blog-like view is only seen by Sidebar users. Other users see a normal email thread, the same as any other they participate in. This blog-like behavior is an emergent effect of the email thread participants keeping the page URL as a useful part of the email they are composing.

By leveraging email to create a blog comment area, Sidebar acquires many of its desired features by simply inheriting the features of email. Attribution, alerts, notifications, controlled content distribution, controlled content contribution and offline access are all provided in familiar ways by the underlying email infrastructure. By composing a blog comment area from the user's personal email archive, Sidebar creates the personalized summaries described in section 2.2.2. The summary shown in Sidebar when viewing a web page includes content from all linked messages in the user's personal email archive, whether sent to a public mailing list, or addressed to the user alone, and excludes those messages the user chose to delete³. By keeping content distribution and storage in email, Sidebar remains email centric and so inclusive of non-Sidebar users, as discussed in section 2.2.4.

4.5 Personal trends

When using Sidebar, the custom of including URLs in email tends to be amplified and so more and more email messages start to include URLs⁴. As this happens, the Sidebar index starts to contain more and more information about the user's topics of conversation. Included with Sidebar is a web application for

³ Sidebar gets its name from this feature, since it enables private side-conversations within a more public context, similar to the conversations lawyers may have with a judge in a courtroom.

⁴ This phenomenon has been observed in our early use of Sidebar, but has not yet been confirmed by a larger user study.

extracting reports about the Sidebar index. These reports show how long the user has been talking about a topic, what days of the week the topic is normally discussed and what the hottest topics of conversation are. Similar to the emergent effect of company valuation in the stock market, Sidebar trends track the popularity of topics in the user's social sphere.

5. NEXT STEPS

The combination of inheriting the features of email and benefiting from the emergent effects of Sidebar has provided most of the target feature set. In this section we explain how we plan to acquire the remaining features.

5.1 Summary message

Currently, the text presented in Sidebar reads more like a summary of an email thread than like an edited document. The feel is more blog-like than wiki-like. To enable iterated editing, as discussed in section 2.1.3, we plan to enable user creation of "summary email". A summary email is a normal email with additional meta-data indicating that it is a summary of other email. When Sidebar finds a summary email, it will display its text and omit text from the summarized email. The summarized email will still be available by selecting an option to view the history of the summary email. To edit the text displayed in Sidebar, the user selects the entries to summarize and invokes the summarize command. This command opens a new email editor, pre-populated with the text of each email to be summarized and meta-data indicating that the email is a "summary email". Iterated editing is therefore expressed as the iterated creation of summary email.

Even with the summary email feature, Sidebar will not present a completely coherent document, as is possible with a wiki. Since each user has a unique view composed from their personal email archive, the author of a summary email may be unable to produce a summary of everyone's email on a given topic. Consequently, the text presented by Sidebar will often consist of a few islands of refined text, intermingled with less integrated comments. This presentation is not so different from the typical state of many wiki pages. We expect this level of document editing is sufficient for many collaboration scenarios.

5.2 Extracting topic data from the web page

Searching using a page URL discovers messages referring to that exact page. A natural improvement would be to find messages about the broader topic, rather than just that particular page. Rather than depend on natural language processing, we instead plan to discover topics by relying on authors embedding additional machine-processable information into web pages. For example, multiple fan pages about an actress could all include a common URL. Were this convention followed, comments about the common URL would appear in Sidebar when browsing any fan page about the actress. Nomination of this common URL would result from the same processes that generate a folksonomy. For example, a particular community might converge on using a page at Wikipedia as the common URL. Causing comments made about the most popular page to also appear on their page is incentive for authors to include the common URL. This comment aggregation also works if multiple URLs are used as the common identifier, rather than a single URL. From a Semantic Web perspective, the statement would colloquially be: "This page is about the same thing that these other pages are about". The thing

itself may never acquire its own URL. Similarly, the neighborhood view discussed in section 2.2.3 can be implemented by extracting common URLs that identify not the main topic of the web page, but related topics.

6. RELATED WORK

Wiki Wiki Webs were first introduced by Leuf and Cunningham [WIKIWAY] as an extensible system of interlinked web pages that can be collaboratively edited by anyone. Since then, several efforts have targeted adding semantics to regular wikis. Platypus Wiki [PW] allowed users to annotate information about pages by constructing RDF triples of the form (wiki page, predicate, object). [SWP] describes a system to add semantics to Wikipedia. In this system, one edits a page as one would edit a normal Wikipedia page. Additionally one can attach meaningful relationships to links and define things as belonging to categories (roughly types). A related development is the concept of a personal Wiki. Personal Wikis focus on personal information management. Unlike the client-web server based architecture typical of most Wikis, personal wikis focus on managing local information and typically do not support collaboration. Newton [NEWTON], for example, is a desktop wiki applet for the GNOME2 desktop environment. It is designed to make the creation of richly formatted documents of any type as simple and quick as possible. SemperWiki [SMP] adds semantics to the personal Wiki, allowing users to quickly and unobtrusively write and semantically annotate pages. Sidebar fits somewhere between these various systems. On the one hand, it is not a centralized system optimized for collaborative writing of a single document. On the other hand, it is designed to support personalization of collaboratively created documents, as well as management of personal data.

Bliki [BLIK] also shares some similarities to Sidebar as the result combines features from both blogs and wikis. A Bliki: "like a blog, ... allows me to post short thoughts when I have them. Like a wiki it will build up a body of cross-linked pieces that I hope will still be interesting in a year's time". Sidebar aims to meet this same goal. However, the underlying architecture and messaging system are quite different.

Sidebar fits in the class of applications known as shared Web annotation systems, a class of systems with a relatively long history [HECK]. In its simplest form, a Web annotation [ZOHAR] can be seen as a remark about a document identified by a URI, made by the author of the document or by a third party, with or without author knowledge. In a shared Web annotation system, annotations are stored in specialized servers. Annotations are shared in that everyone having access to an annotation server should be able to consult the annotations associated with a given document and add their own annotations [ANNOTEA]. Wikalong [WIKALONG] is another type of web annotation system. A Mozilla Firefox extension, it embeds a wiki in the browser sidebar whose topic corresponds to the page displayed in the mail window. While these project to varying degrees resemble Sidebar in spirit, they have by in large focused on providing centralized data stores to support their annotations. Selecting email for message store and transport produces a distinctly different result.

The Semantic Email work under the Mangrove project [MNGR] tries to automate the processing of email by understanding its content marked up in some fashion and with embedded SPARQL in it. Semantic desktops or semantic browsers like Haystack [HYST] try to integrate email along with other documents within an RDF model. While the RDF email

indices are fundamental to the Sidebar implementation, it is not fundamental to the concept. One could envision a different implementation (although perhaps not quite so elegant) that achieve the same functionality.

7. CONCLUSION

In today's workplace, a lot of collaboration happens over email, leaving behind large but mostly unorganized email archives. Today's Web provides a wealth of organizational structures, as well as technology for bringing other organizational structures onto the Web. By more closely integrating email with the Web, Sidebar brings organization to email and more collaboration to the Web. This combination of features, and the emergent properties of the union, creates a powerful collaboration platform for addressing a wide array of media. Anything assigned a URL becomes a point of collaboration in Sidebar. Linking the assigned URL to other resources creates an organizational structure for these collaborations. The work involved in creating useful organizational structures can be shared across disparate user communities who use it to organize private data sets. Created organizational structures are web sites usable by anyone with a web browser and collaborations are email threads, usable by anyone with a mail user agent. This compatibility includes existing web sites and existing email, meaning useful functionality is immediately provided by Sidebar. A critical mass of users is not required to make Sidebar a successful collaboration tool. Since Sidebar is merely client-side presentation software, transmission and storage of user comments remains with the existing email infrastructure. This design reuses existing custom and technology for access control, leverages past infrastructure investments and facilitates deployment, especially in tightly regulated environments.

8. ACKNOWLEDGMENTS

Special thanks to Vlad Bolshakov who helped create the Outlook plugin, and to all those developers who contributed to the open source libraries on which our solution depends. Thanks also to our colleagues at HP for their advice and encouragement, especially Malena Messarina, Venugopal Srinivasmurthy, and Anupriya Ankolekar.

9. REFERENCES

[ANNOTE] José Kahan, Marja-Riitta Koivunen, Eric Prud'Hommeaux, Ralph R. Swick, *Annotea: An Open RDF Infrastructure for Shared Web Annotations*, WWW10, May 1-5, 2001 Hong Kong

[BLIK] M Flower, *Martin Flower's Bliki*, <http://martinfowler.com/bliki/WhatIsaBliki.html>

[DUBLINCORE] Dublin Core Metadata Initiative, <http://dublincore.org/>

[HECK] Rachel M. Heck, Sarah M. Luebke, Chad H. Obermark, *A Survey of Web Annotation Systems*, http://www.math.grin.edu/~rebelsky/Blazers/Annotations/Sommer1999/Papers/survey_paper.html

[HYST] D Quan, D Huynh, D R Karger; *Haystack: A Platform for Authoring End User Semantic Web Applications*; in Proceedings of the WWW2003 Conference, Budapest, May 2003; Online version: <http://haystack.lcs.mit.edu/papers/www2003-developer.pdf>

[JETTY] Jetty web server: <http://www.mortbay.org/>

[WIKIWAY] B. Leuf and W. Cunningham. *The Wiki Way: Collaboration and Sharing on the Internet*. Addison-Wesley, 2001.

[MNGR] Luke McDowell, Oren Etzioni, Alon Halevy, and Henry Levy; *Semantic Email*; in Proceedings of the WWW2004 conference, New York, May 2004; Online version: <http://turing.cs.washington.edu/papers/www2004.pdf>

[NEWTON] Newton wiki applet: <http://newton.sourceforge.net/>

[PW]R Tazzoli, P Castagna, S E Campanini; *Towards a Semantic Wiki Wiki Web*; 3rd International Semantic Web Conference (ISWC2004), November 2004, Hiorshima, Japan. Extended Version: <http://platypuswiki.sourceforge.net/whatis/documents/platypuswiki.pdf>

[SMP] E Oren, R Delbru, K Moller, M Volkel, S Handschuh; *Annotation and Navigation in Semantic Wikis*; In Proceedings of the ESWC 2006, Online version: <http://www.eyaloren.org/pubs/semwiki2006.pdf>

[SPARQL] Prud'hommeaux, and Seaborne, Andy, editors, "SPARQL Query Language for RDF", W3C working draft, 4 October 2006

[SPARQLP] Clark, K.G, editor, "SPARQL protocol for RDF", W3C working draft, 27 May 2005

[SWP] M Volker, M Krotzsch, D Vrandeic, H Haller, R Studer; *Semantic Wikipedia*; In Proceedings of the 15th international conference on World Wide Web, WWW 2006, Edinburgh, Scotland, May 2006; Online version: <http://www.aifb.uni-karlsruhe.de/WBS/hha/papers/SemanticWikipedia.pdf>

[WIKALONG] <http://www.wikalong.org/>

[ZOHAR] R. Zohar. *Web Annotation - an Overview*, Feb. 1999.