

Mail2Wiki: Low-Cost Sharing and Organization on Wikis

Gregorio Convertino*, Ben Hanrahan*, Nicholas Kong*, Thiebaud Weksteen†, Guillaume Bouchard†, Cedric Archambeau†, Ed H. Chi*

Palo Alto Research Center (PARC)*, Xerox Research Centre Europe†
Palo Alto, CA 94304, USA | 38240 Meylan, France

{firstname.lastname}@parc.com*, {firstname.lastname}.xrce.xerox.com†

ABSTRACT

Sharing and organizing content using corporate wikis remains prohibitively costly for most knowledge workers in organizations. This problem has led to low adoption rates, compromising the utility of such tools. To reduce these costs and increase the benefits we introduce two proof-of-concept prototypes. We characterize the lifecycle of production and consumption of shared knowledge in a community of professionals that use wikis as CIorg tools. That is, we analyze breakdowns or gaps in the current lifecycle, identify interventions that can improve it, and specify requirements for different roles involved in it. Then, we illustrate the design of the two prototypes that support users in feeding information into wikis and curating this information for better exploitation. First, we present an email plugin that allows knowledge workers to easily feed new content into their wiki sites directly from the email client (see workshop theme “Feeding CIorg”). Second, we present a MediaWiki plugin that allows wiki curators to update and organize the content shared on the wikis via a combination of machine-based functions and human input on interactive visualizations (see workshop theme “Exploiting CIorg”). The two prototypes were motivated by the same vision and were developed using an iterative and participatory design approach (see workshop theme “Design Methods”).

Author Keywords

Corporate wikis, curation, email, enterprise2.0, knowledge organization, knowledge sharing, visualization, web2.0.

INTRODUCTION

The ability for an enterprise to share knowledge efficiently is a key advantage in the current knowledge economy. However, much of the sharing across knowledge workers continues to occur via email, which remains the central content management tool ([15], [32], [12], [46])

While highly versatile, easy to use, and central to many workflows, email keeps useful information siloed in personal inboxes, which makes it difficult to locate and reuse knowledge across the organization. Also, as more media are piped into email, individual workers are increasingly experiencing the email overload problem [46].

On the other hand, while wikis are becoming increasingly common in organizations (similarly to other enterprise2.0

tools [31]) they are affected by poor adoption rates ([27], [21], [9]). This prevents them from reaching the critical mass that is needed to make them a valuable tool for more efficient knowledge sharing. In fact, recent adoption studies of corporate wikis ([27], [21], [12]) point to the high interaction costs for contributing to wikis and subsequently maintaining and organizing the wiki content. These costs are among the key factors impeding a broader adoption of wikis in organizations. In this paper we focus on corporate wikis, but the pro-adoption solutions proposed may also apply to the other enterprise 2.0 tools, such as corporate blogs, forums, microblogs, and social networking sites [31].

In regards to the high costs for contributing to the current wikis, researchers have observed that it takes significantly less effort to informally email information to a small group of colleagues than structuring a wiki entry ([27], [12]). Sharing content on a wiki requires several steps and context switches: copying, switching application, transferring and pasting, formatting, and saving [22]. Researchers have also observed that knowledge workers tend to fall back on email when there are multiple and non-integrated channels for sharing [6]. Overall, the high interaction costs for the users lead to less information being fed into current wikis. We developed an Outlook Email plugin to address this problem.

In regards to the high costs for curating content on wikis (i.e., updating, organizing, and summarizing), some researchers observed that wikis in organizations tend to easily become disorganized and out of date ([10], [21]). Many wiki pages can become rapidly outdated or remain incomplete, while consuming the resources of and confusing those who visit [21]. Moreover, when compared to richly populated public wikis such as Wikipedia, corporate wikis exhibit the additional limitations of having very sparse and heterogeneous data: pages are often incomplete and their content varies widely, including organized lists, project updates, tutorials. This makes automated techniques, such as entity extraction algorithms, more error-prone than when these same techniques are used to support organization in richly populated public wikis such as Wikipedia (see [26]). Overall, the high costs of curation and the limitations endemic to corporate settings result in less exploited and exploitable shared information on wikis. We developed a wiki plugin aimed at reducing the costs for curation. In this plugin we adopted a mixed-

initiative approach [28] to overcome the specific limitations of corporate wikis.

In the following sections we present two contributions related to the workshop themes of Feeding CIorg and Exploiting CIorg. First, we characterize the lifecycle of production and consumption of shared knowledge in a community of workers using corporate wikis and email. We analyze the breakdowns or gaps in the lifecycle when utilizing the current CIorg tools. This analysis informs our design by suggesting what interventions can improve the lifecycle and points to key requirements for specific roles of people involved in the lifecycle. Second, we illustrate the design of the two prototypes: an email plugin that allows end users to easily contribute and pre-organize wiki content directly from email (during micromoments), and a wiki plugin that supports curators or power users who are willing to devote a few moments to update, organize, or summarize previously shared content.

RELATED WORK: TOOLS AND STUDIES

The problem of low adoption rates

Recent studies of adoption for Web2.0 tools in the consumer space (or public web) have described their successes in recruiting large populations of users in terms of power law or Pareto distributions (e.g., [4], [44]). That is, a minority of the users, or power users (the narrow head), make the majority of the contributions; while the majority of the users (the long tail) make the minority of the contributions. This trend in contribution rate is consistent across several Web2.0 communities: Wikipedia editors [29], Usenet discussions [45], and social tagging systems [20]. A key challenge for deployments of these tools in organizations (i.e., enterprise2.0 tools) is making sure that the tools reduce costs or increase benefits to a sufficient extent so that a critical mass of users is led to “inhabit” the head of the distribution.

Recent results suggest that several Web2.0 tools [37] deployed in corporate intranets, including wikis and blogs, are affected by low contribution rates. These rates range between about 2-15% and vary widely depending on factors such as the specific tool, the organization, and the culture. The reported rates of contribution reported for *corporate blogs* in large companies such as IBM [16] and Microsoft [17] is about 3%. About 2-4% of the IBM workforce visited the Beehive social networking site monthly, while about 15% of the workforce had registered over two years [18]. Brzozowski and collaborators [9] found that the participation rates in social media at HP varied widely depending on the tool, country (10% in UK vs. 1.9% in Japan and Mexico), job function and group, and the manager’s activity in the system. Concerning the tool, Brzozowski found that discussion forums were more popular than blogs and wikis (10 and 28 times more popular, respectively) among HP employees [9].

A number of researchers have focused specifically on *corporate wikis*. Studies conducted by MITRE [27], IBM [5], and Microsoft [21] suggest that while many employees report that they use wikis, very few of those actually contribute. Danis and Singer [14], who studied a corporate wiki in a 900-member research organization over two years, found that the wiki had good rates of views and no vandalism, in contrast to Wikipedians. However, they also found that the viewers of this corporate wiki were reluctant to step into the role of editor and modify others’ content, except in special circumstances. Similarly, Grudin and Poole [21], who studied wikis at Microsoft, found that numerous wikis had been created, mainly to support small groups, but most were quickly abandoned. They identified three challenges for adoption and long-term sustainability of corporate wikis: (1) positioning the wiki in an existing information ecology and corporate culture, (2) the high costs for content organization and flexibility over time, and (3) aligning expectations between managers and individual contributors. In this paper, we focus on addressing the first two challenges, by proposing (1) an analysis of the current lifecycle of production and consumption of shared knowledge via email and wikis and (2) two prototypes for low-cost sharing and curation of content on wikis.

Finally, Palen and Grudin [38] have reviewed prior studies of adoption for traditional collaborative systems. They observed that the bottom-up factors of having evangelists and peer pressure were key facilitators in prior success cases of adoption of these systems. For example, in the case of the electronic calendar, the peer pressure was channeled by the interface and by being integrated with email, as this reminded non-users of the use by others and the benefits that they might be missing. A similar facilitating condition for adoption is created by our first prototype.

Lowering the costs for contributing to wikis

Holtzblatt et al. [27] and Convertino et al [12] found that a key barrier preventing wiki adoption is the high cost of contributing content to the wiki. As a result, knowledge workers tend to revert to sharing via email, their primary habitat for work [15]. In fact, for prospective users of enterprise wikis, “informally emailing information to a small group of people [...] requires significantly less effort than structuring a wiki entry and providing supplemental context (p. 4666)” [27]. However, the centrality of email in many workflows exacerbates the problem of email overload [46]. Consequently, useful knowledge remains siloed in individual email inboxes, making it difficult to locate and reuse knowledge across a large organization.

Since current knowledge workers encounter much of their information inside email [Ducheneaut, Whittaker], several systems have been proposed to build additional collaborative functions around e-mail. Systems such as Meshin [1], XOBNI [3], and Salsa [43] achieve this by

providing insights about the worker's inbox from outside sources. These tools are all *centripetal* to email, as they pull in relevant information from various sources or past activity logs while relating it to emails in the inbox. In contrast to these solutions, we propose an Outlook plugin that, by providing direct access to shared wikis, is *centrifugal* to email. Our prototype not only embeds sharing functions directly in email, but also helps with the process of offloading content from email to the appropriate shared wiki space. That is, the vision is to make valuable content, currently trapped in email, more widely available in order to enable it to be collectively curated and exploited later on.

Lowering the costs for updating and organizing wikis

Researchers have also pointed to the fact that it is too costly for users to organize and maintain corporate wikis [Buffa, Hoffmann]. Many wiki pages become rapidly outdated, remain incomplete, while consuming the resources of and confusing those who visit (e.g., [21]). Moreover, for many users, reorganizing information in the current wikis is “a daunting task” [21].

The design of our wiki plugin is aimed at reducing the cost to update and organize wiki pages through various functions. First, to help updating each wiki page, our system recommends dynamic content, pulled from email and RSS feeds. We borrow this idea from other tools that presents other related content in context. For example, targeted advertisements (e.g., Google Ads) recommend contextual dynamic content given the content of a “static” webpage (e.g., an e-mail). Second, to assist with the organization of content in the wiki page, we enabled the users to manipulate the content in the page through views or visualizations that are peripheral to the page. The design of this component draws on ideas from tools that help users to visually organize content in wikis, such as Vispedia [11]. We also take inspiration from data analysis tools, such as Jigsaw [42] and Entity Workspace [7], and recent research prototypes that support organization in wikis via semantic web tools, such as the semantic Mediawiki extension (<http://semantic-mediawiki.org>), Wikulu [25], and Woogie [23]. In particular, SAVVY Wiki [34] helps to organize fragmentary knowledge items across wiki pages. Third, to mitigate the inevitable errors in the semantic information extracted by automatic algorithms from unstructured data, we use mixed-initiative [28] as our general approach. That is, the results of automatic functions are combined with user actions on such results (e.g., corrections or selections). This approach has been used in past systems that allow individuals to categorize data and refine machine-inferred metadata [35]. However, only a few tools support these functionalities in collaborative contexts, such as Hoffman's work on infoboxes in Wikipedia [26]. We are not aware of any prior tool that has implemented the mixed-initiative approach to enhance corporate wikis, nor are we aware of any tool that uses multiple peripheral visualizations to

implement this approach.

Several tools have focused on making contributions to a shared web-based repository easier. Posterous [2] is a consumer tool that makes blogging easier by allowing users to email contributions to a server that publishes the content. Mail2Tag [36] is a shared email repository that is searchable and persistent. Users can email their content to tags, which enables folksonomy-style organization. Our two prototypes, although closely related, differ from these prior tools by directly integrating wikis with the email client (i.e., our email plugin) and enabling low-cost curation both from the email client (i.e., early organization into existing pages and sections via the email plugin) and from the wiki (i.e., curation of the wiki page via the wiki plugin).

LIFECYCLE: CONTRIBUTE, CURATE, CONSUME

In recent studies of online communities, the earlier problem of enabling greater adoption of the technology is subsumed by the problem of enabling participation in the (technology-mediated) community. For example, participation in a Facebook-based community assumes the use of the Facebook technology. Therefore, our understanding of the conditions that facilitate adoption and sustainability of a discretionary technology for sharing, such as wikis, can be informed by prior studies of participation in communities of people that share information.

In an analysis of various online communities supported by web2.0 tools on the public web, Preece and Shneiderman [39] derived the Reader-to-Leader framework, which describes how the members of these communities can be roughly categorized in the classes that reflect different levels of social participation: readers (or lurkers), contributors, collaborators, and leaders. They use their framework to identify what changes in the design of the technology can encourage people to move to higher levels of social participation. For example, designers can promote greater participation by lowering the threshold in interfaces for readers to easily make small contributions (e.g., no login) or giving visibility to contributors', collaborators', or leaders' work via a reputation system.

Drawing on the Reader-to-Leader framework, in this section we characterize the community of workers in an organization who sharing knowledge via corporate wikis and email. We refer to the different levels of participation observed in terms of *informal roles* that the members can take in relation to the wiki-based sharing.

Roles in corporate wikis

The combination of prior research on wikis and our analysis of about 15 corporate wikis in two organizations have pointed to four main roles that workers with access to corporate wikis can take: curators, simple contributors, lurkers or readers, and unengaged workers or non-users.

The role of *curator*, or a power user, refers to a few

members in the organization who aggregate and package information on wikis in a more consumable manner. Their motivation is for the benefit of the rest of the community, or to comply with their formal job role in the organization. For example, similarly to administrators in Wikipedia (e.g., [29]), a knowledge manager regularly updated and maintained the wiki for a community of Bid Managers at Xerox [12]. This worker had a role analogous to that of the curators in an enterprise file-sharing service that collect, organize, and share files for others (see study of Cattail at IBM [33]). This role is most similar to the *leader* role from the Reader-to-Leader framework.

The *simple contributor* is a role that we observed in several cases. Those are users who make isolated contributions to the wiki. The user makes no changes to the structure of the document. We observed instances of this role especially when a *curator*, such as a manager, had previously created a skeletal page for the employees to then contribute content into specified sections.

The role of *lurker or reader* refers to individuals who do not contribute new content but are simply consumers of information for personal benefit (see also lurkers in IBM's corporate social network [18]).

In our research we also observed the role of *unengaged user*. This refers to those individuals who often see no utility in systems like wikis, although they might rarely happen to consume some curated information.

Overall these four roles considered together and in action, provide a view of the "ecosystem" in which knowledge is contributed, curated, and consumed via corporate wikis and email. In Figure 1, we outline the roles where the unengaged and lurkers are at the bottom producing raw information outside of the wiki, the middle layer is occupied by the simple contributors who sift through this raw information and transform it into semi-structured information, and at the top layer the curators utilize the semi-structured data as a source of information to create the final document. As the level increases in the diagram the number of individuals that occupy that role decreases (generally, exponentially).

Breakdowns observed and requirements derived

There are a few evident breakdowns in this chain with the current tools. First, the threshold for lurkers and simple contributors to easily contribute content is too high. For example, there may be multiple unrelated wikis that all require a login and password every time an edit is to be made, and they are not integrated with core work tools. This is not only an obstacle for lurkers or readers to become contributors but also for unengaged users to become lurkers or readers since wikis are not very easy to access nor visible. Second, the work of curators is poorly supported in that seeking new related content and updating prior content are costly to do in current wikis. This makes it less likely

for contributors to become occasional curators. At the same time the larger number of outdated pages (see [21]) reduces the incentives for becoming lurkers or readers and thus the volume of attention and credit handed by these to contributors and curators. On the other hand, when wikis are not easily accessed or visible, the incentive for curators is also lower (a possible reason to be willing to curate is to be visible in the community, probably the main form of reward in this type of discretionary systems). Third, the work of curators is also made difficult by the lack of tools that allow easy reorganization of the content. For example, there is no support for resorting a page with a chronological list (references, links to professional tools, tasks) into a list with the same items grouped by topic.

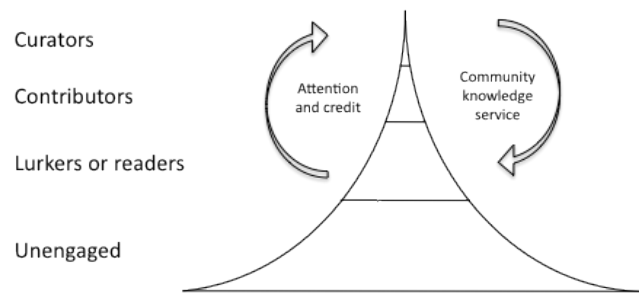


Figure 1. Lifecycle: attention and credit are given to individuals with higher involvement in return for their community knowledge service. The population decreases (exponentially) as the level of participation increases.

To mitigate these breakdowns our work focused on two goals. First, we focused on facilitating the *simple contributor* in contributing semi-structured content while encountering the raw content fed to them. Second, we focused on facilitation the *curator* in organizing and making sense of the information contributed to the wiki and information outside of the wiki (e.g., related information recommended next to the wiki page, see below). This second goal has the added benefit of making the curated page content more attractive to lurkers and contributors. By focusing on these two goals we expect to induce positive domino effect that can mitigate the breakdowns mentioned above: more contributors are likely to lead to more material, more interesting content, and thus more incentives to contribute, and more visibility and rewarding.

We identified three requirements to allow easy contribution of information from email to a shared wiki space. First, reduce the number of steps and context switches in moving between email and the wiki. Second, enable early organization while the user is still engaged with the content to contribute. Third, enable users to expand their knowledge of the wiki through providing information about additional relevant content.

Similarly, there are several requirements needed to support the second goal of curating and updating the shared repository. First, assist curation by providing easier

mechanisms that invite the user to make corrections and edits quickly and easily. Second, address stagnation by making it much easier to view and bring in more current relevant content. Third, enable low cost reorganization of content while accounting for the limitation of sparse and heterogeneous data, (e.g., by using automated techniques via leveraging mixed initiative approaches).

We envision that the ecosystem described above can be made more sustainable by addressing these requirements with enhanced tools that make it easy for users to share information on wikis (i.e., feeding CIorg) and then curate it to keep it relevant and current (i.e., exploiting CIorg).

TWO PROTOTYPES

As outlined earlier, we developed two prototypes to address the two problems of feeding information into the wiki and curation of the content that has been shared to facilitate later exploitation of this content. The first prototype is an email plugin developed for Microsoft Outlook (called Mail2Wiki) and allows feeding information at low cost into the wiki. The second prototype is a plugin developed for MediaWiki installations (called VisualWikiCurator) and allows updating and organizing the wiki page content at low cost. Both prototypes are described in detail below.

I. Lowering the costs for contributing to wikis

The primary goal of the Outlook plugin is to *lower the cost of contributing or sharing* content, thus easing the transfer of currently siloed content from email to wikis. Other goals of this prototype are to provide a bridge between email and a wiki (e.g., preview wiki content), enable early organization the contributed content as it is being shared, and allow multiple emails to be summarized into a wiki page. We first describe the user interface and then present three main usage scenarios.

Outlook plugin: design

We chose to build an extension of the Microsoft Outlook client because it is the most widely used by workers in enterprises but have also experimented with other platforms such as Thunderbird.

The Outlook plugin adds an additional panel to the reading pane area of the email client interface (Figure 2A). The added panel provides an outline of wiki pages related to the current email content as well as controls to select other wiki pages. This outline is meant to surface the current knowledge organization in the web repository while the user is processing incoming emails (email triage) or composing outgoing emails (see [22], for more details on the interface of the email plugin). Some functions of the system are also accessible via a lightweight version of the system. This does not require the plugin but uses the existing Mail2Tag platform for individual to explicitly email content to tags on the shared repository (see [36]).

The Outlook plugin supports the following four functions:

- The plugin provides direct manipulation to support low-cost sharing of content. This enables the user to select text from an email and drag it first to the desired page and finally into the proper section.
- The plugin supports the early organization of content within the wiki by using the previously described dragging functionality to target the proper section.
- The plugin leverages recommendations that are interweaved with direct manipulation. Recommendations are presented in two ways. Pages that are relevant to the selected email are presented on the proximal view. Once a page is selected, as the user drags the content into the outline the three top sections are highlighted by the other sections fading in slowly (we adopt the ephemeral menu technique used by [19]).
- Our plugin also includes functionality directed to feeding large amounts of information to the wiki by enabling page generation. The user achieves this by selecting multiple emails and either creates an entirely new page or a new section on an existing page (see drag and drop interaction in Figure 2B).

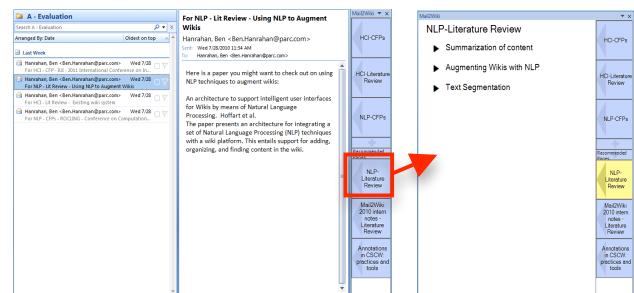


Figure 2A.Plugin: the interface. By clicking on a wiki page listed in the sidebar (red box, left figure), the user can open and interact with the outline of that page (right figure).

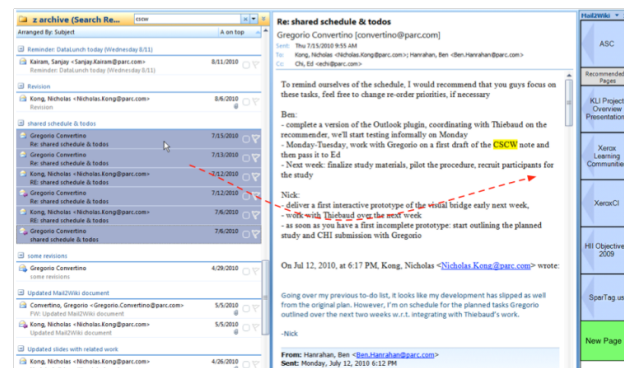


Figure 2B.Plugin: contributing a batch of emails. The red dashed line describes the dragging of these onto the plugin.

Interface without outlook plugin

The overall Mail2Wiki system can be used with or without the plugin. The interface without plugin is supported for users who do not want to install the plugin, are not allowed

to do so by their organization, or use a different platform. Without the use of the plugin these user can still use the system and contribute content by emailing it to a wiki page. This use of the system is similar to how users email content to a tag in the Mail2Tag system [36]. Before publishing the content, the system sends a feedback email message. Through this message the user can asynchronously review and confirm the proposed contribution or cancel it. Through similar asynchronous interactions the user can contribute a single message, a selected part of it, or a batch of email messages, which are summarized to form a new page or to be appended to an existing wiki page.

Outlook plugin: three usage scenarios

1. Contribute part of an email with the plugin

- 1a. Select a part of an email (text)
- 1b. Drag it onto one of the recommended pages or those previously saved as favorites
- 1c. While dragging, view the outline (i.e., list of headings) of the selected page and, within this, the recommended sections
- 1d. Drop the section in the chosen section and, if needed, briefly edit the content before saving
- 1e. Show and hide the content of each section of the page outline

2. Contribute a batch of emails with the plugin

- 2a. Select multiple emails (batch)
- 2a-2e same as 1b-1e above, note that:
 - If the user chooses a recommended/favorite page, then the system merges the batch with it, otherwise it creates new page (and allow user to choose name)
 - If the user chooses a recommended section, then the system merges the batch with it, otherwise it creates new section (and allow user to choose name)

3. Contribute one or more emails without the plugin

- 3a. Select one or a batch of emails as attachment
- 3b. Use the mail2tag formalism to push complete batch to wiki (tag@xeroxshare.com).

Outlook plugin: evaluation of the interface

We used a 2 (tool) x 2 (page complexity) within-subject experimental design. Participants made contributions both with and without the tool. We also tested both simple and complex pages. A fully crossed design with 2 contributions per condition resulted in 2x2x2=8 contributions per subject.

In the with-tool condition, the four pages were immediately available in the interface. We disabled the recommendation feature to measure the impact of our interaction techniques only. In the without-tool condition, we set the homepage in the browser to the main page of the wiki. From here, the participants navigated to a group page, project page, and finally the target page. This navigational stage was based both on our interviews about current wiki usage and the work of Phuwanturak [40], who reported on the existence of many project pages. In our analysis, we considered the without-tool condition both with and without

this navigational stage. The results of our evaluation are detailed in Figure 3 (see [22], for more details).

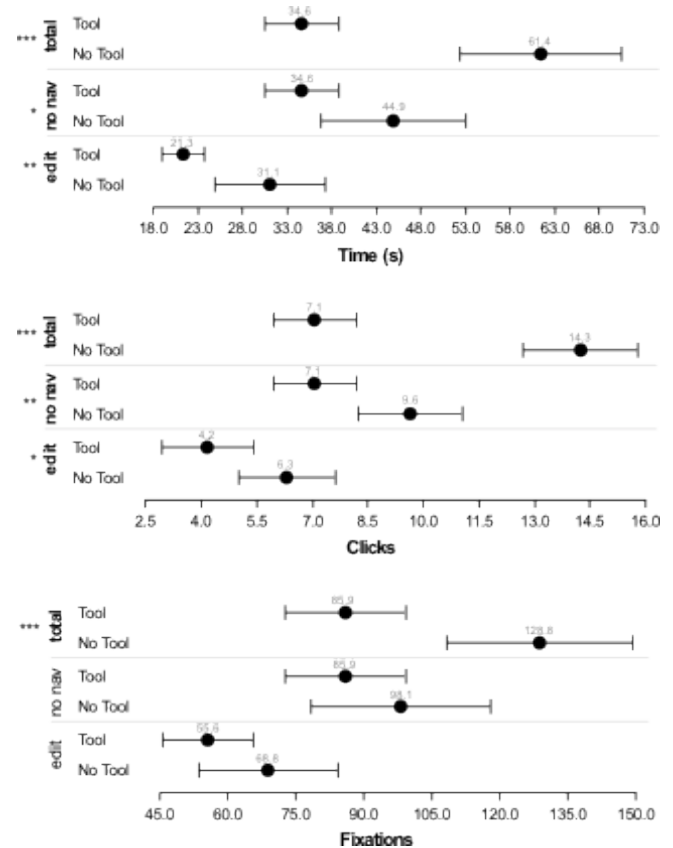


Figure 3. The results of our evaluation. The significance for $p < .05$, $p < .01$, and $p < .001$ is indicated by a *, **, and *** respectively. “Total” includes all three stages, “no nav” excludes navigation, and “edit” includes only editing stage.

The tool had a clear and strong effect on Completion Time (Figure 3, top). Similar to the effects on time, the measure of the number of clicks (Figure 3, middle) indicated that with the tool the users performed fewer clicks across all the stages. The average number of eye fixations over all stages (Figure 3, bottom) was significantly lower with the tool. Other measures did not reach significance.

II. Lowering the costs for updating and organizing wikis

We implemented a MediaWiki plugin or extension, called VisualWikiCurator, designed to *lower the costs to update and organize information on wikis*. This prototype allows users to update the content of wiki pages with information from external data sources and to organize their content.

VisualWikiCurator Design

Our extension embeds two main components into each wiki page. The first component recommends external content, such as e-mails or RSS feeds, that is relevant to the current wiki page for future inclusion and organization (Figure 4, bottom). We call this external content *dynamic content*. The second component extracts metadata from the content and

visualizes this metadata in a “visual bridge” (Figure 4, top). Via the bridge, multiple users can analyze the content from multiple views, correct any erroneous metadata extractions, and organize the wiki page.

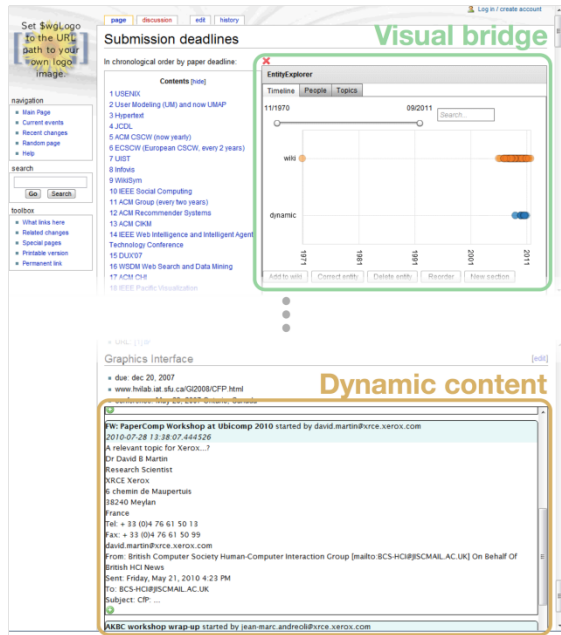


Figure 4. Wiki plugin overview. (Top) The visual bridge, displays alternative views with metadata extracted from the wiki page and dynamic content (a timeline in this figure). (Bottom) The dynamic content (e.g. emails or RSS feeds) is recommended by the system at the bottom of each page.

The recommendation component is intended to lower the cost of updating the wiki. It collects external information from predetermined/prespecified sources, chooses the most relevant pieces of the information based on the current wiki page, and places these pieces alongside the page. Users can easily add new information to the wiki page because of the proximity of relevant information.

The visual bridge is also intended to lower the cost of organization by providing views of metadata extracted from the wiki page and dynamic content and allowing users to organize the page directly from these views. Our system extracts entities (specifically people, places, dates, and organizations) and then provides views of the data. Figure 5 illustrates four of these views: a timeline view for dates, a map for locations, a sortable table for people and organizations, and a topical view based on topic models (see [8] on LDA). The user may be able to more easily identify relevant information to add, or obsolete information to remove, when using one of these views.

As no entity extraction algorithm will have perfect accuracy, we utilize a *mixed-initiative* approach to wiki organization: that is, we combine machine and human intelligence. On the machine side, the system extracts entities and recommends a view to the user based on the number of each type of entity. When the user selects a set of entities and clicks “New section”, the system can move the sections associated with the selected entity into a new section and suggest the content and formatting of a new section. On the human side, the user may correct erroneously extracted or categorized entities, and is also shown a preview of any organizational action the system recommends. The user can then correct any incorrect or unclear formatting or content. The following usage scenario illustrates these features.

VisualWikiCurator: Usage scenario

Here we illustrate an example usage scenario of our wiki extensions. Sonia is browsing a wiki page that lists conference deadlines relevant to her team. On behalf of her collaborators and twenty minutes before a new meeting, she decides to use VisualWikiCurator to group the deadlines that will occur in 2011 into a new section, including any recommended new calls not yet listed in this wiki page. Figure 6 shows a flow chart with her sequence of actions:



Figure 5. Different views in the visual bridge. (a) Timeline view. (b) Location view. (c) Person/Organization view. (d) Topical view.

- [a] Sonia begins work in the timeline view of the visual bridge. She sees that a date that has been assigned a year of 1970, hovers her mouse over the mark to see its content (see yellow tooltip in Figure 2a), and notices that the date was incorrectly extracted: the year should be 2009.
- [b] She clicks the “Correct Entity” button (Figure 2b), changes the date, and confirms.
- [c] Timeline updates with corrected date (Figure 2c).
- [d] She filters down by using the slider at the top, selects a specific interval of dates in 2011 with her mouse (see

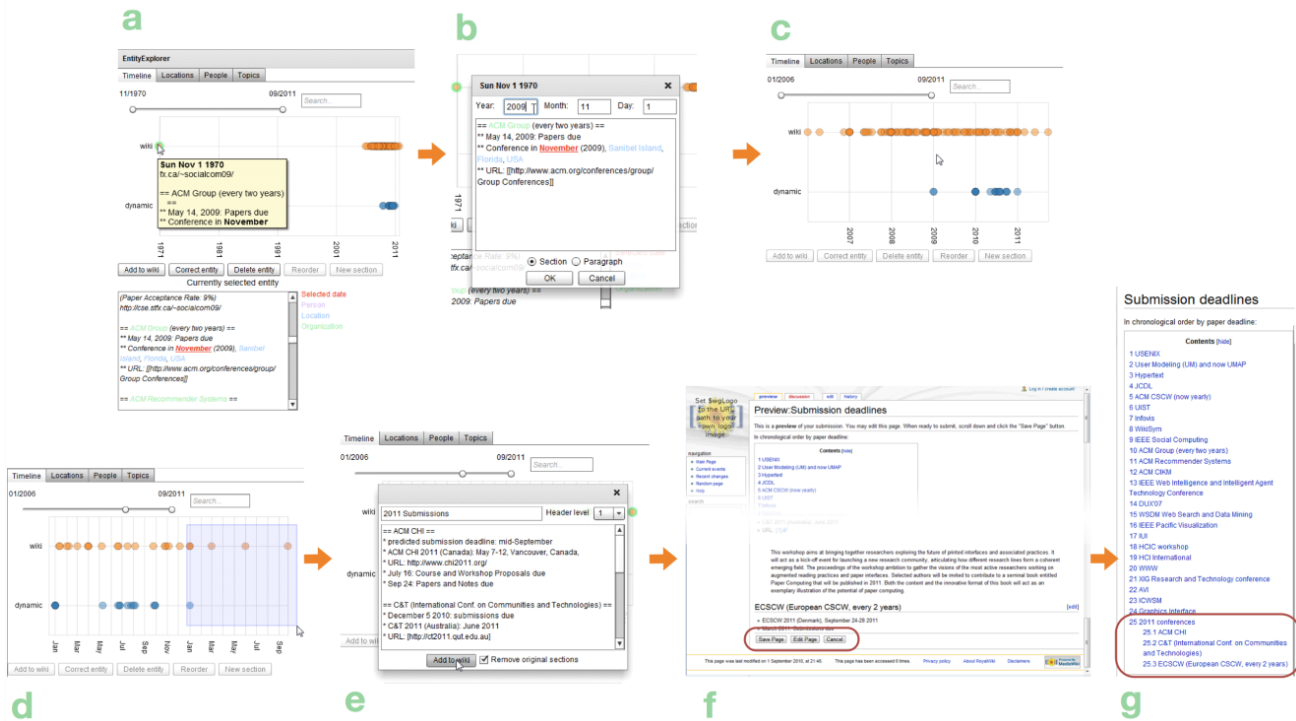


Figure 6. Scenario. A user organizes a wiki page containing conference deadlines. (a-c) The user corrects an entity. (d-g) The user selects entities to create a new section on the wiki page (see [30], for more details on the system)

the blue selection box in Figure 2d), and clicks the “Add new section” button.

- [e] This opens a dialog box prompting for the new section name and a preview of the section content in edit mode (Figure 2e), which she quickly revises.
- [f] Next, the system opens a preview window of the wiki page with the changes (Figure 2f). She scrolls down the page, checks the new section temporarily added at the bottom and clicks the “Save Page” button.
- [g] Finally, the system redirects her to the published wiki page with the new section (Figure 2g)

VisualWikiCurator: ongoing evaluation and development

While developing the extension, we identified common usage scenarios in corporate wikis and solicited user feedback on our features. Following from that process, we are further assessing needs and usability concerns by deploying the wiki extension at two different research centers. In addition, we plan on conducting an evaluation of how effective our extension is in enabling contribution and organization, similar to our evaluation of the Outlook plugin. We will present users with a page and ask them to perform specific organizational tasks (e.g., collect all the conferences occurring in 2011), both with and without the extension.

DISCUSSION

In the previous sections we have presented two contributions that are related to the workshop themes of

Feeding Clorg and Exploiting Clorg. First, drawing on prior studies and our own study of corporate wikis, we proposed an analysis of the lifecycle of production and consumption of shared knowledge in communities of professionals that use current corporate wikis as Clorg tools. We reported on four main user roles that were observed in these studies, analyze key breakdowns in the current lifecycle, and derived requirements for tools that can improve this lifecycle. Then, we have presented two proof-of-concept prototypes that address these requirements. We reported also about prior and ongoing evaluation results about these two prototypes.

Ultimately, by reducing the costs for sharing and curation, the designs of the two prototypes supported the same vision: that is increasing user participation by making it easier for larger number of workers to contribute valuable content, currently trapped in their email, and also to collectively curate it for effective exploitation across the organization. The solutions that we proposed for corporate wikis could be applied to increase participation in other enterprise 2.0 tools, such as corporate blogs or forums that are also suffering of low adoption rates and lack of critical mass of useful content. With this in mind, we point to a few general novel aspects in our prototype designs.

A key novel capability introduced with the outlook plugin is previewing within the email client the structure of the content published on shared remote spaces (wiki pages or

shared documents). Beyond viewing this structure, the user can interact with it to check more detailed information (i.e., the content of the section) and, more importantly, to add his personal content to the most relevant shared page (or documents) and, within that page, the most relevant section. Therefore, in addition to lower the costs for sharing, with the email plugin we proposed a new interaction technique that enables early curation via direct manipulation of the structure or outline of a page or document on a remote shared repository. This is an example of a smart bridge that can be built between the personal information space and multiple related shared spaces.

Another key novel aspect visible across the design of both prototypes is the use of the mixed-initiative approach, wherein NLP and other intelligent techniques are used to assist users in identifying new content to share and organizing existing content (see also [26], [25]). The Outlook plugin recommends possible target locations on the wiki for new e-mails at a page and section level. This extends the user's awareness for relevant public recipients for his selected bits of personal content. The wiki plugin recommends relevant external content for each wiki page, and provides alternate views of the data by extracting and visualizing metadata (in this case, entities).

FUTURE WORK

We plan to extend the Outlook plugin by providing a more concrete connection to the wiki plugin. More importantly however, is to evaluate the complete plugin, including recommendation and page generation functionality, in a naturalistic setting to assess the impact on wiki adoption.

Additionally, we plan to further extend the wiki plugin based on what we learn from our current deployments. Additional features include storing the metadata in order to make metadata changes persistent, and the generation of new pages and sections from the dynamic content. We would also like to further explore the possibilities of mixed-initiative organization. For example, we plan to investigate semi-automatic ways of pulling data from other pages by parsing URLs and parsing the content located at those URLs. Finally, as noted earlier, we plan on evaluating the plugin's effectiveness in assisting in page organization.

Finally, we will be working toward integrating the two prototypes with the Mail2Tag platform and enabling users to access the server-based functions via alternative front-end or interfaces, so that users can start contributing without the Outlook plugin and then be enticed to use more advanced functions via the Outlook plugin and the wiki plugin.

ACKNOWLEDGMENTS

Colleagues and study participants at PARC and XRCE.

REFERENCES

- [1] <http://www.meshin.com/>
- [2] <http://posterous.com>.

- [3] <http://www.xobni.com>.
- [4] Adamic, L. A., Zhang, J., Bakshy, E., and Ackerman, M. S. (2008). Knowledge sharing and yahoo answers: everyone knows something. Proc. WWW'08, 665-674.
- [5] Arazy, O., Gellatly, I., Soobaek J., Patterson, R., (2010). Wiki deployment in corporate settings. In Technology and Society Magazine, IEEE, 28(2) (2009), 57-64.
- [6] Balakrishnan, A. D., Matthews, T., and Moran, T. P. Fitting an activity-centric system into an ecology of workplace tools. In Proc. CHI 2010. ACM Press (2010) 787-790.
- [7] Bier, E., Ishak, E., and Chi, E. Entity Workspace: an evidence file that aids memory, inference, and reading. Proc. of ISI 2006, 466-472, Springer-Verlag.
- [8] Blei D.M., Ng A.Y., Jordan M.I.; Latent Dirichlet Allocation JMLR 3:993-1022, 2003.
- [9] Brzozowski, M. J., Sandholm, T., and Hogg, T. (2009). Effects of feedback and peer pressure on contributions to enterprise social media. Proc. GROUP '09. 61-70.
- [10] Buffa M., Intranet Wikis, in Proc. of the IntraWeb Workshop, Proceedings of WWW 2006, ACM Press.
- [11] Chan, B., Talbot, J., Wu, L., Sakunkoo, N., Cammarano, M., and Hanrahan, P. 2009. Vispedia: on-demand data integration for interactive visualization and exploration. Proc. of SIGMOD 2009. ACM, NY.
- [12] Convertino et al. (2010) Designing for Learning Communities in a Large Enterprise. ACM CSCW 2010 Workshop on Collective Intelligence in Organizations: Toward a Research Agenda.
- [13] Cunningham, H., Maynard, D., Bontcheva, K., and Tablan, V. GATE: A Framework and Graphical Development Environment for Robust NLP Tools and Applications. In Proc. of ACL'02, Philadelphia, US.
- [14] Danis, C. & Singer, D. (2008). A wiki instance in the enterprise: Opportunities, concerns, and reality. Proc. CSCW'08, 495-504.
- [15] Ducheneaut, N. and Bellotti, V. E-mail as habitat: an exploration of embedded personal information management. Interactions, ACM (2001) 8(5): 30-38.
- [16] Dugan, C., Geyer, W., and Millen, D. R. (2010). Lessons learned from blog muse: audience-based inspiration for bloggers. In Proc. of CHI 2010. ACM, New York, NY, 1965-1974.
- [17] Efimova, L. and Grudin, J. 2007. Crossing Boundaries: A Case Study of Employee Blogging. In Proc. of HICSS. IEEE Computer Society, Washington, DC, 86.
- [18] Farzan R., DiMicco JM, Brownholtz B. (2010) Mobilizing Lurkers with a Targeted Task. Proc. of ICWSM '10. May 2010.
- [19] Findlater, L., Moffatt, K., McGrenere, J., and Dawson, J. (2009) Ephemeral adaptation: the use of gradual onset to improve menu selection performance. In Proc. CHI '09, NY, USA, 2009. ACM. 1655-1664,
- [20] Golder S. and Huberman B. A. (2006). Usage patterns of collaborative tagging systems. Jour. of Information

- Science, 32(2): 198–208.
- [21] Grudin, J. and Poole, E. S. (2010). Wikis at work: success factors and challenges for sustainability of enterprise Wikis. Proc. of Wikisyms 2009, ACM, NY.
- [22] Hanrahan, B.; Convertino, G.; Kong, N.; Chi, E. H. Mail2Wiki: the wiki enters the workers' email. Manuscript under review.
- [23] Happel H.J, Social search and need-driven knowledge sharing in Wikis with Woogole, In Proceedings of Wikisyms 2009, ACM, New York, NY.
- [24] Heer J., Flare. <http://flare.prefuse.org>, September 2010.
- [25] Hoffart J., Zesch T. and Gurevych I. (2009). An architecture to support intelligent user interfaces for Wikis by means of Natural Language Processing, In Proceedings of Wikisyms 2009, ACM, New York, NY.
- [26] Hoffmann, R., Amershi, S., Patel, K., Wu, F., Fogarty, J., and Weld, D. S. 2009. Amplifying community content creation with mixed initiative information extraction. In Proceedings of CHI '09. ACM, New York, NY, 1849-1858.
- [27] Holtzblatt, L. J., Damianos, L. E., and Weiss, D.. Factors impeding Wiki use in the enterprise: a case study. In Proceedings of CHI '10, ACM, 4661-4676.
- [28] Horvitz, E. 1999. Principles of mixed-initiative user interfaces. In Proc. of CHI '99. ACM, NY, 159-166.
- [29] Kittur, A., Chi, E., Pendleton, B. A., Suh, B., & Mytkowicz, T. (2007). Power of the few vs. wisdom of the crowd: Wikipedia and the rise of the bourgeoisie. Alt.CHI, 2007.
- [30] Kong, N.; Hanrahan, B.; Weksteen, T.; Convertino, G.; Chi, E. H. VisualWikiCurator: Human and Machine Intelligence for Organizing Wiki Content. Under review.
- [31] McAfee, A. Enterprise 2.0: New collaborative tools for your organization's toughest challenges. Harvard Business Press, Boston, MA. 2009
- [32] Matthews, T., Whittaker, S., Moran, T., and Yang, M. Collaboration personas: A framework for understanding designing collaborative workplace tools. In Workshop on Collective Intelligence in Organizations, ACM CSCW 2010, ACM Press (2010).
- [33] Muller, M., Millen, D.R., & Feinberg, J., Information curators in an enterprise file-sharing service. Proc. ECSCW 2009.
- [34] Nakanishi T., Zettsu K., Kidawara Y., and Kiyoki Y., (2009) SAVVY Wiki: a context-oriented collaborative knowledge management system, In Proceedings of Wikisyms 2009, ACM, New York, NY.
- [35] Nardi, B. A., Miller, J. R., and Wright, D. J. (1998). Collaborative, programmable intelligent agents. Commun. ACM 41, 3, 96-104.
- [36] Nelson, L., Nairn, R., and Chi, E. H. Mail2tag: Efficient targeting of news in an organization. In CSCW 2010 Workshop: Collective Intelligence In Organizations, ACM Press (2010).
- [37] O'reilly, T., (2007). What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. Communications & Strategies, 1, p. 17.
- [38] Palen, L. and Grudin, J. 2003. Discretionary adoption of group support software: lessons from calendar applications. In Implementing Collaboration Technologies in industry: Case Examples and Lessons Learned, Munkvold, Ed. Springer., London, 159-179.
- [39] Preece, J. and B. Shneiderman (2009) The Reader-to-Leader Framework: Motivating Technology-Mediated Social Participation, AIS Transactions on Human-Computer Interaction, (1) 1, pp. 13-32
- [40] Phuwanartnurak, A. J. Did you put it on the wiki?: information sharing through wikis in interdisciplinary design collaboration. In Proc. SIGDOC '09, ACM Press (2009), 273–280,.
- [41] Robertson S., Zaragoza H., Taylor M.(2004) Simple BM25 Extension to Multiple Weighted Fields.CIKM'04.
- [42] Stasko, J., Görg, C., and Liu, Z. 2008. Jigsaw: supporting investigative analysis through interactive visualization. Information Visualization 7, 2, 118-132
- [43] Stecher, K., Counts, Cheng, S., L., Williams, S., and Turski, A. Salsa: Leveraging Email to Create a Social Network for the Enterprise. In Proc. ICWSM 2009, aai (2009), 323–326.
- [44] Wilkinson D. M. Strong regularities in online peer production. In Proc. of the 9th ACM conf. on Electronic commerce (EC'08), NY, USA, 2008. ACM. 302–309.
- [45] Whittaker, S., Terveen, L., Hill, W., and Cherny, L. (1998). The dynamics of mass interaction. In Proc. of CSCW '98. ACM, NY, 257-264.
- [46] Whittaker, S. and Sidner, C. Email overload: exploring personal information management of email. In Proc. CHI '96, ACM Press (1996), 276–283.
- [47] Whittaker, S. (2005). Supporting Collaborative Task Management in Email. Human Computer Interaction, 20, 49-88.