Using Web Annotations for Asynchronous Collaboration Around Documents

JJ Cadiz, Anoop Gupta, and Jonathan Grudin

May 10th, 2000

Technical Report
MSR-TR-2000-44

Microsoft Research
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052

Using Web Annotations for Asynchronous Collaboration Around Documents

JJ Cadiz, Anoop Gupta, Jonathan Grudin
Microsoft Research, Collaboration & Multimedia Group
One Microsoft Way
Redmond, WA 98052 USA
+1 425 705 4824
{jjcadiz, anoop, jgrudin}@microsoft.com

ABSTRACT

Digital web-accessible annotations are a compelling medium for personal comments and shared discussions around documents. Only recently supported by widely used products, "in-context" digital annotation is a relatively unexamined phenomenon. This paper presents a case study of annotations created by members of a large development team using Microsoft Office 2000—approximately 450 people created 9,000 shared annotations on about 1250 documents over 10 months. We present quantitative data on use, supported by interviews with users, identifying strengths and weaknesses of the existing capabilities and possibilities for improvement.

Keywords

Annotation, asynchronous collaboration, distributed work, computer mediated communication, World Wide Web

1 INTRODUCTION

Highlighting and writing comments in the margins as we read is a natural activity. These annotations are often personal notes for subsequent reference. When shared among co-workers they also support communication and collaboration. With paper documents, such sharing is hindered by the need to exchange physical copies.

The extremely wide adoption of the Internet and World Wide Web opens up significant new opportunities. Not only has it become easy to publish documents on the web for friends and co-workers to read, we can also build rich annotation systems for distributed, asynchronous collaboration. "In-context" annotations can be tightly linked to specific portions of content in a document—accessible from a web browser anytime and anywhere—with threads visible in the document, access control to regulate viewing and editing, and a notification subsystem to inform relevant people when new annotations are added. Although research systems with similar capabilities have been proposed and built (as noted below) widely used commercial systems have only recently become available. The literature contains little on the use of web annotations by large workgroups, a gap this paper begins to fill.

Microsoft's Office 2000 is one of the first commercial products to support web annotations for workgroups as described above. In this paper, after providing a brief overview of Office 2000 web annotations, we focus on a case study of how a large product group used the annotation system. We analyze 9,239 annotations made by approximately 450 members of the group on 1,243 documents between May 1999 and February 2000. We also interviewed several team members to better understand how the system was used.

The paper is organized as follows. After presenting related work in the next section, Section 3 gives a brief overview of the Office 2000 annotation system. Section 4 sets up the context of the Case Study—the workgroup, job roles, their task, and our methodology. Section 5 presents data regarding system usage, including types of annotators, usage over time, and use of notifications. Section 6 discusses factors that influenced use, including orphaning of annotations, staying aware of changes, public nature of annotations, responsiveness of users, and richness of annotations. We conclude in Section 7.

2 RELATED WORK

Previous research has shown that annotating text is an important companion activity to reading, with annotations used for manifold purposes. In an extensive field study of annotations in college textbooks, Marshall [13, 14] found that annotations were used for purposes that included bookmarking important sections, making interpretive remarks, and fine-grain highlighting to aid memory. O'Hara and Sellen [18] found that people use annotations to help them understand a text and to make the text more useful for future tasks. Annotations are often helpful for other readers as well, even when they are not made with others in mind [12, 13].

Computer-based annotations can similarly be used for a variety of tasks. For example, Baecker et al. [1] and Neuwirth [16] state that annotations are an important component in collaborative writing systems, where "collaborative writing" refers to fine-grained exchanges among co-authors creating a document. In the study reported here, the focus is on a later stage in the document generation process when a relatively complete draft of the document is posted on the web and annotations are used to get coarser-grain feedback from a larger group of people (beyond the original authors). Differences in tasks affect the relative value of features, which we expect to see reflected in the use of the annotation system we studied.

2.1 Annotations in Commercial Products

Virtually all commercial document-processing packages (e.g., Microsoft Word, Lotus Notes) support some form of annotations. Microsoft Word provides an "insert-comment" command, with comments shown using an interface similar to footnotes. Similarly, one can track changes made to the document, which are displayed to co-authors who can accept or reject changes. These notes and changes are stored within the document file and are not available for collaborative access over the net: One must give the file to a co-author. Lotus Notes allows discussions around document over a network, but comments are linked to the document as whole, and not to individual sentences or paragraphs. These systems are thus not collaborative in the sense defined in Section 1, and are not considered further here.

More recently, several companies have created client-server systems that provide the ability to annotate any page on the World Wide Web [10, 17, 19, 20, 22]. These systems allow people to attach sticky notes or comments to web pages, which are visible to other people who have downloaded the same plug-ins. One company, Third Voice, drew considerable initial attention with its software, but has been hindered by concern that their system allows undesirable graffiti to be posted on major web sites. Overall, these products have not been directed at corporate workgroups, the focus of our study.

2.2 Annotations in Research Settings

Research systems have also supported digital annotations. Quilt, PREP, and Comments provided annotation functionality for co-authors [11, 16]. Quilt supported both text and voice annotations, provided controlled sharing of annotations based on roles, and used email to notify team members of changes. However, to the best of our understanding, these systems had limited deployment.

The more recently developed CoNotes system from Cornell [7, 9] allows students to discuss homework assignments and web handouts. It provides a web-based front-end for annotations that can be anchored at pre-designated spots. A study by Davis and Huttenlocher examined the use of CoNotes by 150 undergraduates in a computer science course. Students used annotations to discuss questions and assigned problems. The authors provide evidence that CoNotes use improved performance and established a greater sense of community among students. Although CoNotes was used in other courses, larger scale study results are not available.

Another recent system is MRAS from Microsoft Research [2]. It focuses on annotations for streaming video content on the web. For example, videos of classroom lectures can be annotated with questions and answers. It allows controlled sharing based on annotation sets and user-groups, it supports text and audio annotations, and it uses email for notification. A recent report [3] discusses its use in two offerings of a course for corporate training and makes feature recommendations. Students liked the freedom of on-demand access coupled with the ability to have "in-context" online discussions. Instructors spent less time answering questions than in live teaching, but were concerned by the lack of personal contact. The study reported here involves a system focused on text annotation in a different task context.

In addition to MRAS, other research prototypes have supported both text and audio annotations, and researchers have examined the differential impact of text and audio from author and reviewer perspectives [2, 4, 15]. In general, they report that although audio allows an author to be more expressive (e.g., intonation, complexity of thought), it takes more effort by reviewers to listen to audio comments (e.g., the inability to skim audio). The system used in this study supports only text annotations, so the issue is not directly addressed. However, we do report interview feedback suggesting that richer annotation types would be helpful.

The Anchored Conversations system [5] was presented at CHI 2000. It provides a synchronous text chat window that can be anchored to a specific point within a document, moved around like a post-it note, and searched via a database. Annotations arise not out of asynchronous collaboration, but during synchronous collaboration, and all annotations are archived. A laboratory study of six three-person teams is reported, with more studies planned.

In summary, although there appears to be an agreement on the potential value of annotations and several existing systems that support annotations, we found relatively few research papers on large-scale use of annotations. This research complements the prior literature by reporting on the use of annotations by several hundred people over a ten month period.

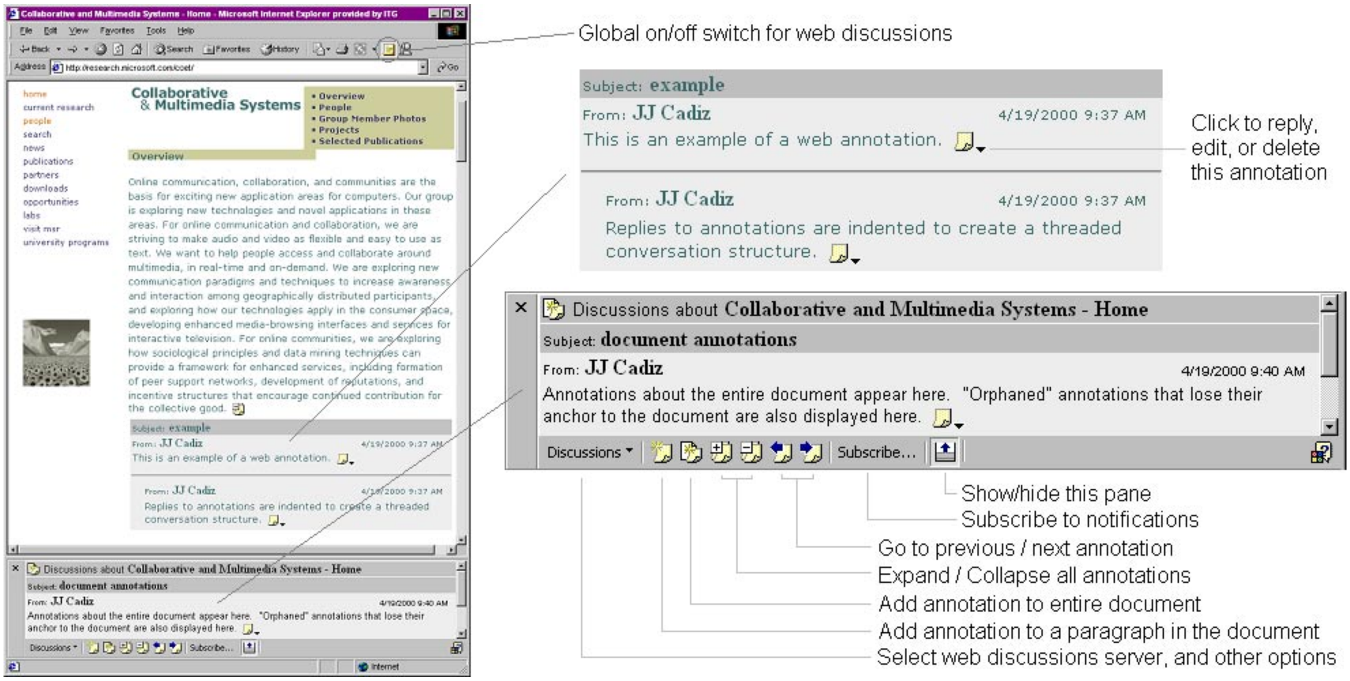
3 THE ANNOTATION SYSTEM

The recently released Microsoft Office 2000 includes a feature called "web discussions," which allows team members to make annotations to any web page.

3.1 System Overview

The annotation system uses a client/server model (Figure 1). The client is the web browser, which receives data from two servers: the web server and the annotations server.

The annotation server resides on a company's intranet and consists of a SQL Server database that communicates with



Collaborative and Multimedia Systems - Home - Microsoft Internet Explorer provided by ITG

Address http://research.microsoft.com/coet/

Collaborative & Multimedia Systems

Overview

Online communication, collaboration, and communities are the basis for exciting new application areas for computers. Our group is exploring new technologies and novel applications in these areas. For online communication and collaboration, we are striving to make audio and video as flexible and easy to use as text. We want to help people access and collaborate around multimedia, in real-time and on-demand. We are exploring new communication paradigms and techniques to increase awareness and interaction among geographically distributed participants, and exploring how our technologies apply in the consumer space, developing enhanced media-browsing interfaces and services for interactive television. For online communities, we are exploring how sociological principles and data mining techniques can provide a framework for enhanced services, including formation of peer support networks, development of reputations, and incentive structures that encourage continued contribution for the collective good.

Subject: example
From: JJ Cadiz 4/19/2000 9:37 AM
This is an example of a web annotation.

From: JJ Cadiz 4/19/2000 9:37 AM
Replies to annotations are indented to create a threaded conversation structure.

Discussions about Collaborative and Multimedia Systems - Home
Subject: document annotations
From: JJ Cadiz 4/19/2000 9:40 AM
Annotations about the entire document appear here. "Orphaned" annotations that lose their anchor to the document are also displayed here.

File Edit View Favorites Tools Help
Back
Search
Favorites
History
Go
home
current research
people
search
news
publications
partners
downloads
opportunities
labs
visit msr
university programs
• Overview
• People
• Group Member Photos
• Projects
• Selected Publications
Discussions
Subscribe...
Internet
Global on/off switch for web discussions
Click to reply, edit, or delete this annotation
Show/hide this pane
Subscribe to notifications
Go to previous / next annotation
Expand / Collapse all annotations
Add annotation to entire document
Add annotation to a paragraph in the document
Select web discussions server, and other options

Figure 2: A web page that has been annotated. Annotations can be made to paragraphs within the document or to the entire document. The row of buttons at the bottom of the browser is used to manipulate annotations.

web browsers via WebDAV (the Web Document and Versioning Protocol). After the browser downloads a web page, it checks the database for annotations. Annotations that it finds are inserted at the appropriate places on the web page. Annotations are linked to the text they annotate by storing two pieces of information with every annotation: the URL of the document, and a unique signature of the paragraph to which they are attached. Thus, the annotation system does not modify the original HTML file in any way.

With this implementation, annotations can be made to any web page, including one outside the company's intranet. However, only those people with access to the same annotation server can see each other's annotations.

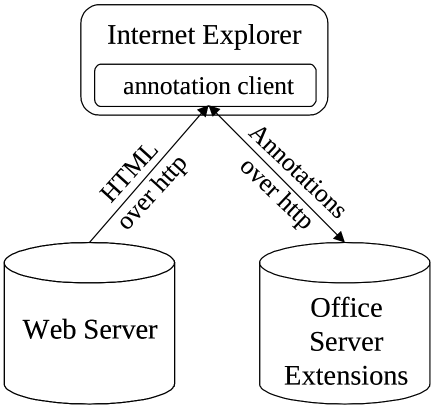


Figure 1: The high level architecture of the Office 2000 annotations system. The Office Server Extensions are implemented on top of a Microsoft SQL Server.

3.2 User Interface

An annotated web page is shown in Figure 2. Annotations are displayed in-line with the original web page. Replies are indented to create a threaded conversation structure.

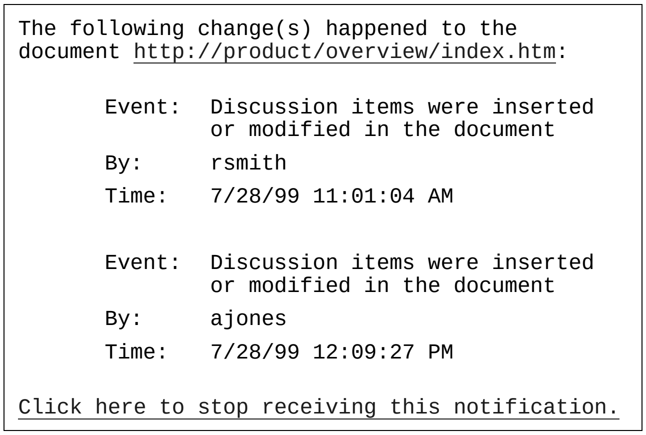
To create an annotation, a user clicks a button at the bottom of the browser. The web browser then displays all possible places where an annotation can be made. The user clicks one of these and a dialog box appears, into which the user types the subject and text of the annotation.

To reply to an annotation, a person clicks the icon at the end of the annotation. An annotation author can edit or delete it by clicking this same icon. Users can expand, collapse, or filter the set of annotations by person or time period using buttons at the bottom of the browser.

With the "subscribe" button, a user can request to be sent email when annotations have been modified or made to a document. With these notifications, users do not have to check a document repeatedly to see if anything has changed. People can choose to have the notifications sent for every change, or the changes can be summarized and sent on a daily or weekly basis. An example of a change notification email is shown in Figure 3.

4 A CASE STUDY: SOFTWARE DESIGN

In early 1999, a large team began using the Office 2000 annotations system in designing the next version of their product. This team has well over 1000 employees, and most members are distributed across several floors of multiple buildings on Microsoft's Redmond, Washington campus.



```
The following change(s) happened to the
document http://product/overview/index.htm:

        Event:  Discussion items were inserted
                or modified in the document
        By:     rsmith
        Time:   7/28/99 11:01:04 AM

        Event:  Discussion items were inserted
                or modified in the document
        By:     ajones
        Time:   7/28/99 12:09:27 PM

Click here to stop receiving this notification.
```

Figure 3: An email notification of annotation activity.

4.1 The Task

The product team primarily used the system to develop specification documents, or "specs." Prior to writing the code for a new feature, the feature is described in a spec. Specs are usually Microsoft Word documents or, in this case, web pages. A spec typically covers one feature or a set of related features, such as a spelling checker. Over one thousand specs were used in the development process studied. Although annotations were also made to other types of documents, they were primarily used with specs, thus we focus on this use.

4.1.1 Job Roles

The majority of team members have one of three job roles: program manager, tester, or developer. Program managers design features and drive the development process. Developers write the code to implement the features. Testers act as the quality control agents in the process, ensuring that program managers create high quality specifications and developers write code that works according to the specifications. A program manager "owns" several specs and is primarily responsible for their development, while testers drive the spec inspections. A more detailed view of software development practices at Microsoft is provided by [6].

4.1.2 Using Annotations to Develop Specs

Once a program manager is comfortable with a draft of a spec, it is published on the web and people are notified that it is ready for comments. Because this product indirectly affects many people in the company, specs draw several comments from people outside the product team.

People can read the spec and discuss it through Office 2000's annotations. Program managers may respond to comments and modify the spec accordingly. Group members also discuss specs via phone, email, and face-to-face conversations. Eventually, a formal "spec inspection" meeting is held to discuss unresolved issues. The goal is to bring the spec to a point where everyone will "sign off" on it, at which point developers can begin writing code.

4.1.3 Spec Development Without Annotations: The Spreadsheet Method

Annotations are not the only way a team discusses specs; the team in question was developing specs long before the annotation system existed. In addition, not all groups use the annotation system: others use a combination of annotations and other methods.

Prior to the existence of this system, one system in particular was used for commenting on specs. This method is still used by some groups within the product team. This method has no formal name, but we will refer to it as "the spreadsheet method."

With this method, a program manager publishes a spec and team members print the spec so that each line is labeled with a line number. All comments are entered into a spreadsheet and refer to the spec using the line numbers. Spreadsheets full of comments are sent to a tester who compiles the comments into a single spreadsheet, which is then sent to the spec owner. Using this method, all comments are anonymous. Sometimes the spreadsheet method is used by itself, and sometimes it is used in conjunction with the annotation system.

4.2 Study Methodology

To study this team's use of the annotation system, we downloaded a copy of their annotation server's database. The database included annotations from as early as January 1999, but the system was not widely used until May. Thus, we limited our study to the ten month period of May 1st, 1999 to February 29th, 2000. Prior to analysis, 103 blank annotations (annotations with no words) were deleted. We have no information on the extent to which people read annotations (apart from responses).

From the annotation database, we selected ten people to interview based on usage patterns. We interviewed four of the five people who made the most annotations, three people who used the system moderately, and three who used the system for a while and then stopped. All interviews took place in January and February 2000. Nine of the ten people work in Redmond, Washington; the other works in Silicon Valley. All ten worked for the product group we studied. Five were testers, four were program managers, and one was a developer.

5 SYSTEM USAGE

In the following sections, we discuss the usage of the system. We examined the annotators, the documents that were annotated, and the use of the notification system.

5.1 Annotators

First we examined the nature and continuity of system use. Developing specs using annotations represented a change in work practice, and use was discretionary, not mandatory.

Annotator Statistics	Heavy Users	Light Users	One-time users	All Annotators
Number of annotators	155	145	150	450
Avg number of annotations per person	47.5	9.3	3.6	20.5
stddev	58.6	7.8	4.4	39.9
median	32	7	2	8
Avg number of documents annotated	10.5	3.2	1.3	5.1
stddev	9.7	2.5	1.2	7.1
median	7	3	1	2.5
Avg number of days an annotation was made	10.6	2.8	1.0	4.9
stddev	7.7	0.8	0.0	6.2
median	8	3	1	3
Average number of words per annotation	26.6	32.7	38.9	28.2
stddev	33.7	40.1	50.5	36.2
median	18	24	28.5	20

Table 1: Statistics describing the behavior of annotators.

Overall, about 450 people made at least one annotation during the ten month period. Table 1 shows the annotator statistics. The high variability in use motivated us to classify users based on number of days in which they created annotations. Some people only made comments once or twice, while others used annotations consistently for several months. We created three groups: one-time users, light users (created annotations on two to four days), and heavy users. (A day when a person made one annotation is treated as equal to a day when a person made twenty annotations.) Figure 4 shows the histogram of the number of days that annotators made an annotation, demarcated into the three groups.

One-time annotators only contributed on one day. These annotators tried the system and either decided not to use it again or have had no reason to use it again. 33% of all annotators are in this group, accounting for 5.8% of the annotations in the data set. Table 2 shows that over half of the one-time commenters were not on the product team.

Light users are people who made at least one annotation on two to four different days. 32% of annotators are light users, and 14.6% of annotations came from this set.

The remaining 79.6% of all annotations come from the 32% of annotators labeled **heavy users**, who made annotations on five or more different days.

We also examined users' job roles, given the likelihood that their roles affect their annotation behavior. Table 2 shows the annotators categorized by 1) one-time users, light users, and heavy users; 2) program managers, developers, testers, or other job roles, and 3) whether annotators worked on the product team or not. Employees outside the team could create annotations by using the team's discussion server. Although these employees did not work directly for this product team, their work was often directly related to the product (for example, product support specialists).

Table 2 shows that about two-thirds of the annotators worked for the product group. A majority of the heavy and light users worked for the product group. One somewhat surprising finding was that even though program managers, developers, and testers have considerably different jobs, the number of annotators was fairly equivalent across these roles, implying that all types of team members tend to be

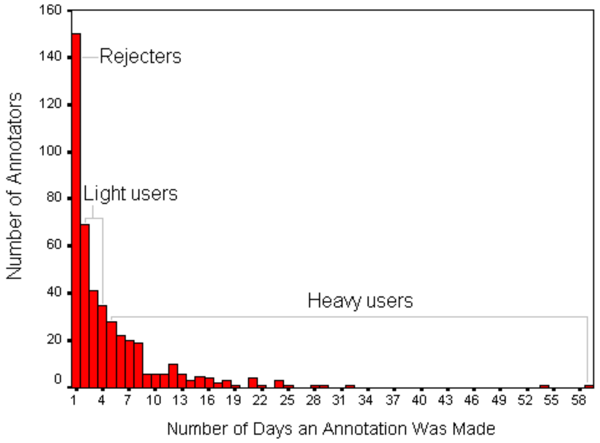


Figure 4: Histogram of annotators based on the number of days they made at least one annotation.

		Pgm Mgr	Dev	Test	Other	**Total**
In product group	Light	19	34	23	13	**89**
	Heavy	39	31	39	13	**122**
	One time	17	21	16	19	**73**
	Total	**75**	**86**	**78**	**45**	**284**
Not in prod group	Light	15	6	4	31	**56**
	Heavy	7	1	2	23	**33**
	One Time	18	9	4	46	**77**
	Total	**40**	**16**	**10**	**100**	**166**
Grand Total		**115**	**102**	**88**	**145**	**450**

Table 2: All of the annotators broken down into user type and job role.

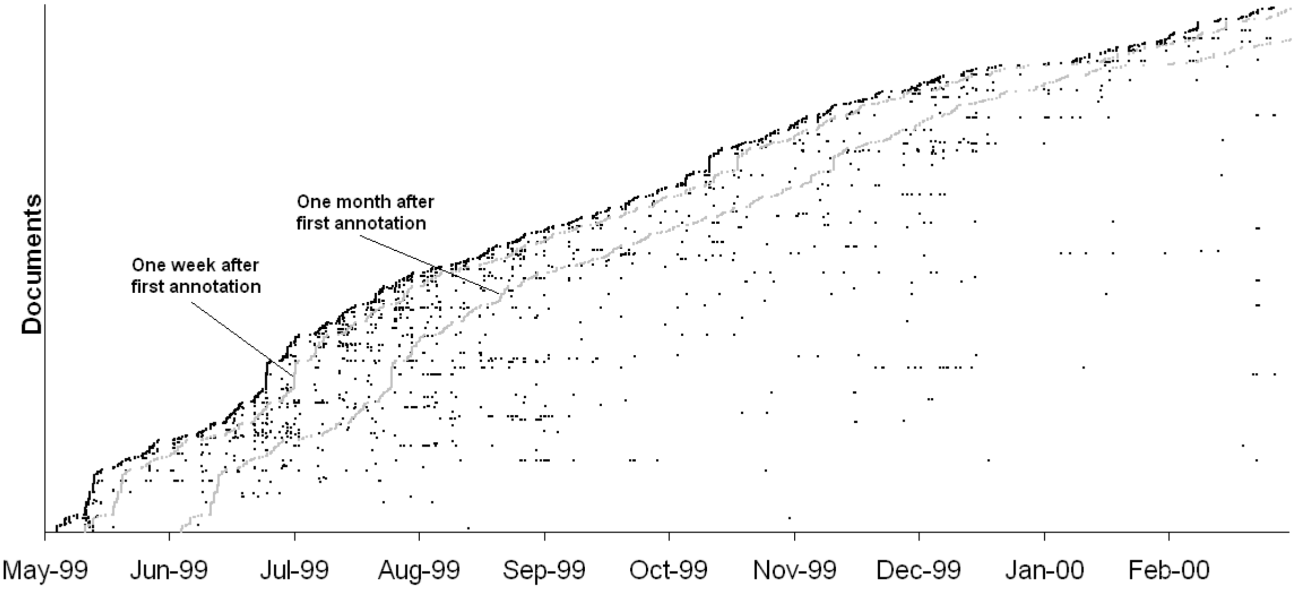


Figure 5: Annotations made on documents over time. Each row represents one document. Each point represents a day on which at least one annotation was made to the document. Two gray lines are drawn to provide a sense of how many annotations are made to a document one week and one month after the first annotation is made.

involved equally with the spec development process.

5.2 Documents

In addition to studying how often people made annotations, we also examined the extent to which documents are annotated. Table 3 displays the document annotation statistics. Most documents had relatively few annotators, and the number of annotations for each document was highly variable.

Figure 5 displays the days on which particular documents were annotated, sorted in order of the first day on which a document was annotated. Two inferences can be drawn. First, the slope of the top-most points shows the rate at which new documents were annotated, and the clusters of points indicate time periods when more annotations were made. A noticeable burst of annotations occurs in July and August of 1999, which is consistent with a significant milestone in the product cycle. In addition, annotations become less frequent as time progresses (especially after January 2000), which is consistent with the product cycle as the focus moves from developing specifications to developing code. A lack of annotations is apparent during holidays in early July, late November, and December.

Document Statistics	
Number of documents annotated	1,243
Avg number of annotators per document stddev median	1.9 1.7 1.0
Avg number of annotations per document stddev median	7.4 19.1 2.0
Avg # days between first and last annotation stddev median	12.6 32.1 0

Table 3: Statistics describing documents that were annotated.

The second, more significant inference is that although most annotations for a document occur near the time of the first annotation, many documents continue to be annotated for several months. One document was first annotated in June 1999, heavily discussed through August, and then discussed again in February 2000. Thus, the number of days that pass between the first and most recent annotation for each document is highly variable.

5.3 Use of Notifications

The highly variable period over which documents are annotated evidences a need for the notification system that tells users when changes are made to document annotations. We examined the extent to which notifications were used.

Of the 1,243 annotated documents, 411 have at least one person subscribed to notifications about that document. The average document has 0.7 subscriptions (standard deviation is 1.4).

269 people subscribed to notifications for at least one document. Of these, the average user subscribed to notifications for 5.3 documents (standard deviation is 10.0). 28% of the notifications are set to notify people whenever a change occurs, 70% are set to send a daily summary (the default), and 2.7% are set to send summaries once a week.

Of the 269 people who subscribed to notifications, 67 never made an annotation, indicating that people may read annotations but not create any. This may be similar to “lurking” in distribution list or newsgroup contexts, although these readers may have contributed comments by phone, email, or in face to face meetings.

6 FACTORS INFLUENCING USAGE

Data from the previous sections indicate that usage of annotations is quite variable. Thus, we turn to our interviews to explore more deeply some factors that influenced how the system was used.

6.1 Technical Orphaning of Annotations

From our interviews, the primary reason that people stopped using the system was annotation orphaning. Because the annotations system anchors annotations by computing a unique signature for each paragraph, the system can fail to match an annotation to the correct location when the text is edited. When this happens, the annotation is "orphaned" and displayed at the bottom of the browser.

Annotation orphaning is understandably frustrating: the power of annotations stems from being context-based, and they are worded with the context assumed. Without the context, many annotations are useless. From the annotator's standpoint, it can be extremely frustrating to take the time to comment on a document, only to see the comments become meaningless through orphaning.

Interestingly, the orphaning problem is not unique to the annotations system. With the line number-based spreadsheet method, spec changes rendered line numbers invalid. (One person mentioned that printing out a spec was a method used to keep program managers from changing a specification document.) Thus, the annotations system did not introduce the orphaning problem: it just failed to fix it and arguably made it worse.

However, even if a better technical approach to preserving document location is found, another problem remains: the annotation context may change so that the annotation no longer makes sense (for example, after a problem has been fixed in response to the comment). The solution to these issues may rely on involving the people who understand the context. When an annotation is being orphaned, its creator could be notified and asked to reposition or delete it as desired. Other solutions suggested by interviewed annotators were to provide a way to mark discussions as "closed" so that others can tell when a thread is no longer pertinent (allowing a 'design rationale' to be preserved without hampering ongoing activity), and giving document owners greater ability to modify annotations on their document.

6.2 Staying Aware of Changes

Some users we interviewed felt they did not need the notification system as they checked the web pages frequently enough. Others, however, expressed frustration with the notification system. These users felt it was often difficult to tell when they should check back to review the document again. There are two components to users' notification needs:

1) Knowing when new annotations have been added to the document

2) Knowing who added them, where in document they were added, and what the annotation says

The first desire was mostly met by the current notification system. The system provided fine-grain control on when notifications were sent (e.g., whenever a change occurred, once a day, once a week). The notification system, however, does not provide control based on who made the annotation (e.g., primary author, one's manager), and this could be used to control notification traffic. We say more on this topic in later subsections.

The second user desire—to be shown specifically who changed what for a document—was only partially met by the notifications system. Notification email (Figure 3) provides a hyperlink to the original document and a list of the people who made annotations, along with when they made annotations. However, this list does not allow a user to easily see what has changed. This is especially important for long documents that are heavily annotated. The easiest way in the current system to get to a specific new annotation is to go to that web page and filter annotations based on date and the person who had made the new annotation. This is, however, a cumbersome process.

Expanding the notification message to include a few sentences of context along with the text of the annotation should go a long way in addressing this concern. With this improvement, if the annotation text were not interesting, the reader would not need to go to the web page. Another improvement would be to have the URL in the notification contain additional anchor information, so that clicking on it would take the user to the precise position in document where the annotation was made [2].

A minor issue with the notifications system was that there is no support for meta-awareness, e.g., knowing if a particular person had subscribed to notifications. As a result, people reported often sending email to inform another person that annotations had been added to a document. The solution to this problem is not easy, however, as there are privacy concerns.

6.3 Responsiveness of Users

Related to the notifications problems was the perception that users' response time with annotations was relatively slow. In our interviews, people said they did not make annotations if something had to be done quickly. They felt that the turn-around time for annotations is not fast enough when a quick response is required. This feeling was supported by an examination of annotation timestamps. Table 4 indicates that the time to answer questions and reply to annotations averages about seven days (with a median of one to two days); annotations in the form of questions are typically replied to more quickly than annotations that are not questions.

Annotation Statistics	
Number of annotations	9239
Percent of annotations that were replies	32%
Average time to reply stddev median	7.6 days 15.5 1.9
Number of annotations that were questions	2650 (29%)
Number of questions that were replied to	1128 (43%)
Average time to answer questions stddev median	6.2 days 12.9 1.0

Table 4: Statistics describing reply times for annotations. Annotations were classified as questions if a ? appeared anywhere in the annotation.

However, the slow response time was not generally seen as a significant disadvantage, perhaps because other communication channels (such as email) can handle situations requiring a quick response. In addition, although user response time is slow, it is faster than the spreadsheet method: Program managers reported liking the annotation system because they could receive feedback on specs when it was written, rather than waiting a few days for all comments to be compiled into a spreadsheet.

Furthermore, slow response time did not hinder productive discussions from occurring. People mentioned that they liked using annotations to resolve minor issues outside of large group meetings, thus preserving face-to-face meeting time for more important topics.

For future systems, however, we believe an improved notification system could quicken the discussion turn-around time. A lightweight mechanism that reliably indicates when to check a document again—if it made it easy to see what had changed—could decrease response time. This is a complex issue with functionality and interface considerations, requiring further research. Perhaps people should be immediately notified of replies to their annotations, regardless of whether they have subscribed to notifications. Perhaps subscribers should be immediately notified of annotations made by a document's primary owner, which are often responses to issues raised by reviewers. While email works well as a delivery mechanism, it can be overwhelming. Thus, support for automatic filtering or routing, if it maintained visibility of activity in the email interface, might encourage finer-grained notification. Notification mechanisms might also shift based on overall annotation activity, for example by treating a comment after a long period of inactivity differently.

6.4 The Public Nature of Annotations

The fact that annotations are automatically shared and potentially viewable by anyone on the team also affected use of the system. People said they did not make annotations if their comment was "nitpicky" or relatively minor, saying that it is a waste of time for all readers of a document to see comments on grammatical or spelling errors. On the other hand, some users think these comments contribute to overall document quality and regret their absence.

In addition, people said they did not make annotations when they felt a comment could be taken the wrong way, or when they did not want to appear overly harsh. For example, users stated that they did not use annotations when their comments were of the "This is stupid," "Have you thought this through?", or "What were you thinking?" ilk. Phone and face to face conversations were the preferred method of communication for these types of comments.

Finally, it was perceived that people did not repeat a comment on a spec if someone else had already made the same point (note, however, that this assertion is contrary to the report of [21]). If true, this saves time for document reviewers, but is lost information for document owners. With the spreadsheet method, reviewers would make comments without knowledge of other reviewers' comments. Thus, some comments would be replicated, but these comments could be grouped together to give the document owner a sense of how many people thought a particular item was an issue. Prioritizing issues in this way is not possible with the annotations system.

A relatively simple improvement that could address this issue would be the addition of binary agree/disagree buttons to every annotation, similar to that provided by Third Voice [19]. With these buttons, reviewers could express feelings about existing annotations with little effort, and document authors could use this information as a prioritization tool.

6.5 The Richness of Annotations

Users we interviewed said they did not make annotations if they had a high level comment that was relatively difficult to express in text. People also said that they did not use annotations when they were trying to clear up confusion. "I don't use annotations when the person doesn't get it," one stated.

Research by Chalfonte reports that communication media that are 1) more expressive, and 2) more interactive are, "especially valuable for the more complex, controversial, and social aspects of a collaborative task" [4]. Given that plain text is relatively limiting in expressiveness (e.g., no intonation of voice), and annotation responses had long latency (shown in Section 6.3), the behavior expressed in interviews should not be too surprising.

The lack of annotation richness may also affect their ability to support discussions. Examining the annotation threads (Figure 6), we find that annotation discussions are rare and brief. Of the 6,263 threads in the database, 4,067 had only one annotation, 1,717 had two annotations, and only 479 threads had three or more.

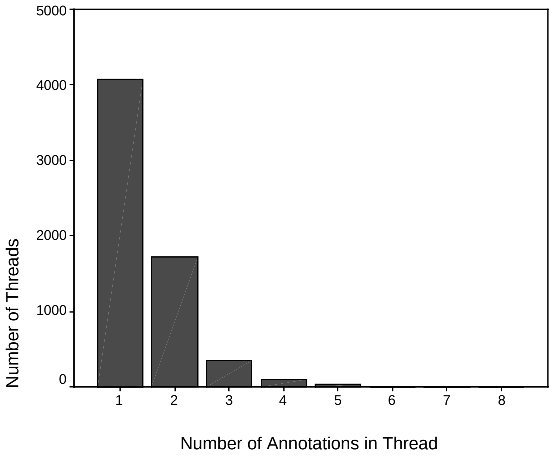


Figure 6: Histogram of the number of annotations in each thread.

To provide a richer communication medium, the annotations system could be modified in two ways. First, as proposed by Churchill [5], programs like MS Messenger and AOL Instant Messenger could be combined with the annotation system to enable synchronous discussions. The instantaneous nature of chat creates a highly interactive medium, which could be appropriate when all parties are available.

Second, the system might support voice-based annotations, a more expressive medium. Voice annotations have been found to be especially helpful for complex or potentially controversial topics [4, 15], two types of comments people chose not to make with the system studied. However, voice annotations require audio-enabled hardware, are difficult to skim, and have often proven unpopular [2, 8].

7 CONCLUDING REMARKS

The annotations system we studied continues to be used successfully by the product design team. Context-based discussions are viewed as an improvement over previous work practices, and the interviews have suggested various ways in which the system could be improved further.

We see several areas that warrant further research. First, as noted earlier, notifications could benefit from additional research. This is a very general problem for systems to support cooperative work: How to be unobtrusive but accessible, inform without overwhelming, separate higher and lower priority information for different actors at different times? What defaults are appropriate (given that most users choose the defaults, as we found in this study) and how can defaults be changed with little effort?

Second, how to differentially treat and design for various actors (for example, a document owner, an annotation creator, and a respondent) is another area requiring careful consideration and further research. This issue interacts with notifications, and could also play a role in handling finer-grained comments (such as spelling fixes that might be handled through a separate channel).

Finally, our data have only indirectly suggested aspects of annotation readership behavior. This should be explored more thoroughly, along with the complementary roles of phone, email, and face to face discussions that facilitate collaboration around documents.

ACKNOWLEDGMENTS

We thank Mike Morton for helping us obtain the data that made this study possible, as well as Dave Bargeron and Marc Smith for valuable discussions. We also thank the product team members who spoke to us about their experience with the system.

REFERENCES

1. Baecker, R., Nastos, D., Posner, I., and Mawby, K. (1993). The User-centered Iterative Design of Collaborative Writing Software. *Proceedings of the 1993 ACM Conference on Human Factors in Computing Systems* (INTERCHI 93).
2. Bargeron, D., Gupta, A., Grudin, J., and Sanocki, E. (1999). Annotations for Streaming Video on the Web: System Design and Usage Studies. *Proceedings of the Eighth International World Wide Web Conference* (WWW8).
3. Bargeron, D., Gupta, A., Grudin, J., Sanocki, E., and Li, F. (1999). Asynchronous Collaboration Around Multimedia and its Application to On-Demand Training. Microsoft Research Tech Report 99-66.
4. Chalfonte, B., Fish, R., and Kraut, R. (1991). Expressive Richness: A Comparison of Speech and Text as Media for Revision. *Proceedings of the 1991 ACM Conference on Human Factors in Computing Systems* (CHI 91).
5. Churchill, E., Trevor, J., Bly, S., Nelson, L., and Cubranic, D. (2000). Anchored Conversations: Chatting in the Context of a Document. *Proceedings of the 2000 ACM Conference on Human Factors in Computing Systems* (CHI 2000).
6. Cusumano, M., and Selby, R. (1997). How Microsoft Builds Software. *Communications of the ACM*. 40(6).
7. Davis, J., and Huttenlocher, D. (1995). Shared Annotation for Cooperative Learning. *Proceedings of the 1995 Conference on Computer Supported Cooperative Learning* (CSCL 1995).
8. Grudin, J., (1989). Why groupware applications fail: Problems in design and evaluation. *Office: Technology and People*, 4, 3, 245-264.
9. Huttenlocher, D. CoNote: A System For Supporting Collaboration with Shared Documents. Available at http://www3.cs.cornell.edu/dph/docs/annotation/annotations.html
10. Hypernix, http://www.hypernix.com
11. Leland, M., Fish, R., and Kraut, R. (1988). Collaborative Document Production Using Quilt.

Proceedings of the 1988 ACM Conference on Computer Supported Cooperative Work (CSCW 88).

12. Luff., P., Heath, C., and Greatbatch, D. (1992). Tasks-in-interaction: Paper and Screen Based Documentation in Collaborative Activity. *Proceedings of the 1992 ACM Conference on Computer Supported Cooperative Work* (CSCW 92).
13. Marshall, C. (1997). Annotation: From Paper Books to the Digital Library. *Proceedings of the 1997 ACM International Conference on Digital Libraries* (DL 97).
14. Marshall, C. (1998). Toward an Ecology of Hypertext Annotation. *Proceedings of the Ninth ACM Conference on Hypertext and Hypermedia* (Hypertext 98).
15. Neuwirth, C., Chandhok, R., Charney, D., Wojahn, P., and Kim, L. (1994). Distributed Collaborative Writing: A Comparison of Spoken and Written Modalities for Reviewing and Revising Documents. *Proceedings of the 1994 ACM Conference on Human Factors in Computing Systems* (CHI 94).
16. Neuwirth, C., Kaufer, D., Chandhok, R., and Morris, J. (1990). Issues in the Design of Computer Support for Co-authoring and Commenting. *Proceedings of the 1990 ACM Conference on Computer Supported Cooperative Work* (CSCW 90).
17. NovaWiz, http://www.novawiz.com
18. O'Hara, K., and Sellen, A. (1997). A Comparison of Reading Paper and On-Line Documents. *Proceedings of the 1997 ACM Conference on Human Factors in Computing Systems* (CHI 97).
19. Third Voice, http://www.thirdvoice.com
20. uTok, http://www.utok.com
21. Wojahn, P., Neuwirth, C., and Bullock, B. (1998). Effects of Interfaces for Annotation on Communication in a Collaborative Task. *Proceedings of the 1998 ACM Conference on Human Factors in Computing Systems* (CHI 98).
22. Zadu, http://www.zadu.com