

# A Classification of Opening Posts in Commercial Software Help Forums

Xiaole Kuang<sup>1</sup>, Bo Yi<sup>1</sup>, Shengdong Zhao<sup>1</sup>, Jianann Chow<sup>1</sup>,  
Tovi Grossman<sup>2</sup>, George Fitzmaurice<sup>2</sup>

<sup>1</sup>National University of Singapore  
{xiaole, yi-bo, zhaosd,  
jianann}@comp.nus.edu.sg

<sup>2</sup>AutoDesk Research  
{Tovi.Grossman,  
George.Fitzmaurice}@autodesk.com

## ABSTRACT

The opening posts in software help forums reflect the users' confusion/expectations regarding to the software learning and usage. In this study, we present a classification of 1200 opening posts in two commercial software help forums (Adobe Photoshop & Cakewalk Sonar) based on three dimensions: type, scope, and topic. We find that the most common opening posts are related to error/stuck situations while using the software to finish a specific task. Additionally, users tend to spend more words to describe posts that involve the operating system or installation issue. Our classification presents a new perspective to examine the software feedback from users and motivates design implications for software designers and researchers.

## Author Keywords

Opening posts, commercial software, online help forum.

## ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g. HCI): Miscellaneous.

## General Terms

Human Factor, Design

## INTRODUCTION

A Software help forum is an online discussion site where users can ask questions, share their experiences, and help each other to find solutions. It has great importance to both software companies and users, as it provides an alternative cost-saving way of offering customer support, and also serves as an important channel to help seeking.

The process of seeking help in software help forums typically starts with an opening post which is either a question or a statement related to the software. Understanding the opening posts could: 1) identify users' confusion regarding learning or using the software; 2) indicate the difficulties users may encounter when describing a software problem. The former can provide

insights for software designers, and the latter can be the basis for improving software help tools.

Previous research paid more attention to the study of opening posts in general online help forums such as Yahoo Answers [1] or technical support boards [8]. Although Singh, et al. initially studied the types of opening posts in open source software (OSS) help forums [7], there is a lack of systematic analysis of opening posts in software help forums from multiple dimensions. Additionally, considering the essential differences between open source and commercial software applications [5], we believe that the opening posts in commercial software help forums have their own particularities that warrant a separate study.

In this paper, we investigated the characteristics of the opening posts by collecting and analyzing the sampled data in two official help forums of Adobe Photoshop<sup>1</sup> and Cakewalk Sonar Producer<sup>2</sup>. Both are popular commercial software applications with rich functions and relatively large active user base. Following the qualitative content analysis [9], we classified the opening posts of these two help forums according to three dimensions: *type* (how the opening post is raised), *scope* (the extent of the issue in the opening post), and *topic* (what the opening post is about). By manually coding 1,200 opening posts (600 per forum), in each dimension, the posts were further analyzed in terms of their distribution and average post length under different categories.

Our contributions focus on revealing the insights behind the opening posts raised by users in software help forum. In particular, we postulated that the classification of opening posts can help reveal the users' confusion and expectations of software products which can benefit software designers by identifying potential user requirements. Additionally, the analysis of different categories of the opening posts provides suggestions on how to better allow an easier and more convenient way for users to describe their problems.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

APCHI '12, August 28–31, 2012, Matsue-city, Shimane, Japan.  
Copyright 2012 ACM 978-1-4503-1496-1/12/08...\$15.00.

<sup>1</sup> [http://forums.adobe.com/community/photoshop/photoshop\\_windows](http://forums.adobe.com/community/photoshop/photoshop_windows)

<sup>2</sup> <http://forum.cakewalk.com/tt.aspx?&forumid=5&p=2113>

## METHOD

To develop the classification of the opening posts, we followed the standard steps of qualitative content analysis [9] which is defined as “a research method for the subjective interpretation of the content of text data through the systematic classification process of coding and identifying the themes and patterns. [2]”

### Data Preparation

We chose the official help forums of two commercial software applications: Adobe Photoshop and Cakewalk Sonar Producer. Both forums<sup>3</sup> host active posting activities (19,910 discussed topics for *Photoshop*, 109,083 for *Sonar* as of September 5<sup>th</sup>, 2011). By collecting all opening posts within a 12-month time period (July 2009 – June 2010), the dataset consists of 6,250 opening posts in *Photoshop* and 11,488 opening posts in *Sonar*.

### Classification and Coding

Base on Grounded Theory [4], using an initial random sample of 200 opening posts from our dataset (100 for each forum), four researchers carefully developed the coding categories to classify the opening posts. The entire process was an iteration of coding sample posts, testing inter-coder agreement, and revising the coding categories. The categories were tested and revised by the four researchers until the Cohen Kappa values for the coding of samples between two researchers was higher than 0.8.

After the coding categories were finalized, we randomly sampled another 1,200 opening posts (600 each) from the same dataset. Using the developed categories, four objective coders [9] who were not involved in the previous categorization development were recruited to manually code these opening posts. The 1,200 posts were divided into 6 groups (200 per group). For each group, two coders first independently categorized the posts; then the Cohen Kappa value was calculated to evaluate the inter-coder agreement, last the conflicted posts with inconsistent coding results between the two coders were resolved through discussion. Among the 6 groups, the average Kappa value between each two coders was 0.81.

## RESULTS AND ANALYSIS

We focused our analysis on posts related to software learning and usage. Of the 1,200 sampled questions, a number of opening posts that were not relevant (3.3% for Photoshop, 8.17% for Sonar) were identified and excluded from further analysis. The topics of these excluded opening posts ranged from discussing sales price or promotional offers, to purely social greetings to other forum users.

The remaining opening posts are classified by three dimensions independently: type, scope, and topic. We then examine these posts according to their distribution

(measured in percentage) and their average post length (measured in number of words per post [3]). We regard post length as a soft indicator of how complex or information-rich the issue is. In this analysis, we used the one-way ANOVA test to compare the average post length.

### Type of Opening Post

For the first dimension (type), we are concerned with the language composition; how the opening post was described. In particular, four different types are identified:

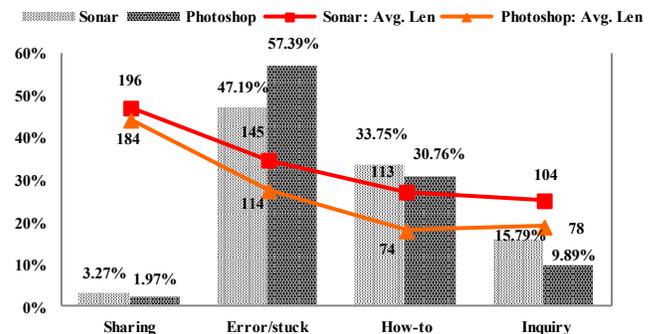
**Sharing:** The opening post does not ask for help, but starts a discussion by sharing a personal experience, e.g. “*I just find a way to make Notion 3 and Sonar 8 to work together. Here is what I did... how do you guys think about it?*”

**Error/Stuck:** The opening post seeks help regarding a situation that prevents users from carrying out a function or operation, e.g. “*CS5 suddenly stops working right after it opens. Any thought?*”

**How-To:** The opening post seeks help, but it is not about a problem or error encountered within a function or feature. Instead, it inquires about a specific procedure, such as how to finish a certain task or use a specific function, e.g. “*How do I show two images on my monitor simultaneously?*”

**Inquiry:** The opening post seeks help by asking for information or exploring the functionality of the application, e.g. “*Is Sonar 8 compatible with XP 64?*”

Figure 1 reveals the distribution of opening posts and the average post length across the different types. Overall, it shows that the most common type of opening posts is *Error/Stuck* (>47%). In Singh et al.’s study about OSS forums, they reported the most common type of questions is *How-To* (34 out of 80, 42.5%) [7]. Although further confirmation needs to be done, one possible reason is that the community-updated nature of OSS [5] increases the flexibility of software functions. This may motivate more *How-To* questions in users’ learning/exploring process.



**Figure 1. The distribution of opening posts and average post length for different types**

Considering the average length for the different types of opening posts, it suggests that the longest type is *Sharing*

<sup>3</sup> In this paper, *Photoshop* and *Sonar* refer to both the forum and the application.

( $p < .001$ ). A sharing opening post normally describes the personal experience or opinions with the intention to start a discussion with other community members. In such posts, users need to explain the step-by-step procedures they have performed, which will invariably lead to more detailed and longer descriptions.

### Scope of Opening Post

For the second dimension (scope), we capture the extent of that the raised software issues may affect (e.g. strictly within the application or beyond it).

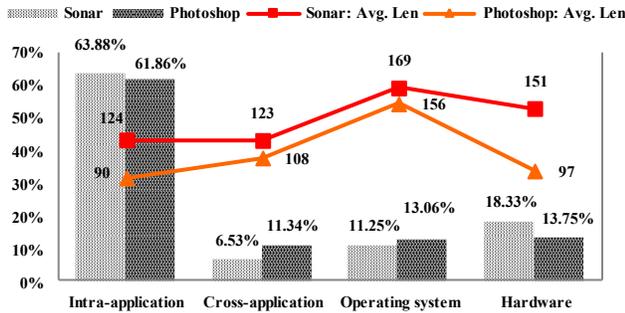


Figure 2. The distribution of opening posts and average post length for different scopes

**Intra-Application:** The scope of the opening post is limited to the application itself, e.g. “*In the history panel, how can I set the source for the history brush?*”

**Cross-Application:** The scope of the opening post involves other software applications, e.g. “*Does Notion 3 (another application) run smoothly with Sonar?*”

**Operating System:** The scope of the opening post involves operating systems, e.g. “*How good is Sonar 8.5 on Windows 7?*”

**Hardware:** The scope of the opening post involves specific hardware, such as graphic cards or scanners, e.g. “*After I installed a new tablet, Photoshop could no longer detect my graphics card.*”

While categorizing the scopes of the posts, we are only concerned with the content of the posts rather than their possible causes. For example, an opening post with the description “*my application crashed suddenly*” is classified as the scope of Intra-Application, although the reason for the crash might be at the system level (*Operating System*). Figure 2 presents the distribution of opening posts (in percentage) and the average post length for the different scopes.

Overall, it shows the majority is about functions and features within the application (*Intra-Application*, >61%), which is not surprising since such forums are places meant for users to discuss the specific application.

However, a significant number of opening posts are actually related to the compatibility issues (the other three scopes), which suggests that the usability of software

usually cannot be determined by its features and functions alone. More interestingly, such compatibility issues, especially with *Operating System*, normally indicate a consideration or even potential threat why customers hesitate to purchase the product, which should attract more attention from the development company.

Further analysis of the length factor reveals that opening posts that relate to the operating system are the most verbose ( $p < .01$ ). This is possibly because that a problem the topic of *Operating System* often suggests a circumstance that is urgent and needs to be solved; users may consider such problem as more critical and attempt to attach all relevant information to attract replies.

### Topic of Opening Post

For the third dimension (topic), we are interested in the subjects discussed (e.g. the events that cause the issue to be raised). Four different topics are observed:

**Installation:** The opening post describes installing or uninstalling something such as a new version of the application or an upgrade of the system, e.g. “*I attempted to upgrade Photoshop 7.0 to 7.0.1. Suddenly, the upgrade program could not find the installation path.*”

**Maintenance:** The opening post is about how to configure the application to make it function correctly or run better, e.g. “*I’m curious how people configure their non-RAID hard drives for Sonar 8.5.*”

**Specific Task:** The opening post describes a particular task that the user intends to accomplish with the software, e.g. “*I’m creating a simple vector in Photoshop. I was wondering if you could lock the vector mask.*”

**General Learning:** The opening post inquires about the functionality of the software, e.g. “*What do you think of the tempo envelope feature?*”

Figure 3 shows the distribution of opening posts (in percentage) and average post length for these topics. It can be observed that *Specific Task* has the highest proportion (>42%).

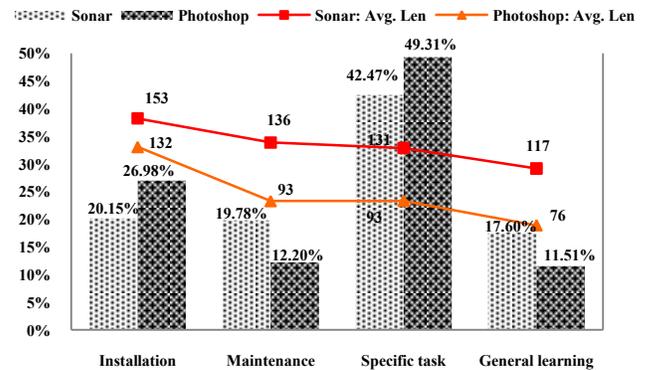


Figure 3. The distribution of opening posts and average post length for different topics

Given the facts that installation is only an initial step for a trial or product and commercial software product typically undergoes stringent installation testing<sup>4</sup> before it is released, it is thus troubling that the topic of *Installation* still happens quite often (>20%). The analysis of post length further shows that the opening posts with *Installation* topic use the most words on average ( $p < .05$ ). Our observation is that such issues are difficult to diagnose, and users typically attach any possible relevant descriptions with the hope that other forum members can help identify possible causes and hence suggest solutions.

### DISCUSSION & IMPLICATION

From the above analysis, it can be observed that the most common opening posts are of type: *Error/Stuck*, of scope: *Intra-application*, and of topic: *Specific Task*. It suggests that users are most likely to start a thread in the forums when they are either stuck in a procedure or encounter an error while executing a particular task. With such knowledge, future help tools for software should devote more effort in monitoring possible error events and automatically capture application-level context.

Moreover, the most verbose case is found to be of type: *Sharing*, of scope: *Operating System*, of topic: *Installation*. In particular, the longest opening post in our sample (3,913 words) is the user shared his/her configuration strategy to optimize the performance of Sonar upon Win7 (operating system). Considering the significant amount of effort a user spends in composing such a lengthy post, this certainly illustrates an opportunity for future help tools to reduce the users' workload for elaborating such details. For example, recording the user's operation sequence at the system level and providing a visualized context collectively with the user's self-description can largely ease the process of sharing. Tools, such as Chronicle [6], that help users record and share their workflow histories could be useful.

Our developed classification provides a new perspective to examine users' feedback about the software products.

For example: an opening post, which has the type of *How-to*, the topic of *Specific Task*, and the scope of *Cross-Application*, normally indicates the user's expectation regarding a desired but unimplemented software feature. E.g. "I was playing with Premiere (Cross-Application). I enjoyed its white balance feature. I want to create a picture for my personal website (Specific Task). How can I achieve a similar effect using Photoshop (How-To)?" Summarizing the users' expectations from such posts can help to develop the users' wish list of the expected features. For a software company, implementing such relevant features can enhance its competitive power and therefore attract potential customers.

Another useful example is when the opening post is of the type *Error/Stuck*, the topic of *Installation*, and the scope of *Operating System*. Such posts typically reflect the compatibility problems between the software and the operating system. *Error/stuck* emerging in such situations might be a concern when deciding whether to continue using the product or not. Summarizing the raised problems in such posts and providing the corresponding step-by-step instructions in the help manual can better stabilize the products and conserve current customers.

### CONCLUSION AND FUTURE WORK

We provide a systematic classification of opening posts in two commercial software help forums. The results suggest that the opening posts in these forums are quite diverse along the dimensions of type, scope, and topic.

Future studies could look at detailed analysis of the interactions between these dimensions which we studied. In the future we would also like to extend our work to classify the responses that are given to opening posts. We hope that our study and results can help researchers and designers gain a better understanding of users' confusion and expectations of the software products. Such insights can aid a better design of future software.

### ACKNOWLEDGMENTS

This research is funded by "Question Answer Social Forum Analysis Project", WBS R-252-000-432-720, Autodesk Canada, US\$20,000 (PI).

### REFERENCES

1. Harper, F. M., Moy, D. and Konstan, J. A. Facts or friends?: distinguishing informational and conversational questions in social Q&A sites. In *Proc. ACM CHI* (2009), 759-768.
2. Hsieh, H.-F. and Shannon, S. E. Three Approaches to Qualitative Content Analysis Qual Health. *Qualitative Health Research*, 15, 9 (2005), 1277-1288.
3. L. Adamic, J. Zhang, E. Bakshy and Ackerman, M. Knowledge sharing and yahoo answers: everyone knows something. In *Proc. ACM WWW* (2008), 665-674.
4. Patricia Yancey Martin and Turner, B. A. Grounded Theory and Organizational Research. *The Journal of Applied Behavioral Science*, 22, 2 (1986), 141.
5. Peeling, N. and Satchell, J. Analysis of the impact of open source software. (2001).
6. T. Grossman, J. Matejka J and Fitzmauriceh., G. Chronicle: Capture, Exploration, and Playback of Document Workflow Histories. In *Proc.*, ACM UIST. (2010), 143-152.
7. V. Singh, M. Twidale and Rathi, D. Open Source Technical Support: A Look at Peer Help-Giving. In *Proc. IEEE HICSS* (2006), 118c - 118c.
8. Yardi S. and E., P. Please help!: patterns of personalization in an online tech support board. In *Proc. ACM C&T* (2009), 285-294.
9. Zhang Y. and Wildemuth, B. M. Qualitative analysis of content. *Applications of social research methods to questions in information and library science* (2009), 308-319.

---

<sup>4</sup> [http://en.wikipedia.org/wiki/Installation\\_testing](http://en.wikipedia.org/wiki/Installation_testing)