An empirical study of chat room discussions for Stack Overflow questions

Wenhan Zhu (Cosmos)

July 10, 2020

Stack Overflow

Over 19 million questions and 29 million answers for programming problems.

Stack Overflow Mechanisms

▲

3

▼

🔖

1

↺

This code has been working for ages. I thought maybe I accidently pressed a key but I cannot seem to see it. I suddenly get the error:

Unable to set the NumberFormat property of the Range class

in the below code:

```
Worksheets("Sheet1").Cells(y + k, x + j).Select
TempValue = Worksheets("Sheet1").Cells(y + k, x + j).Value
Worksheets("Sheet1").Cells(y + k, x + j).NumberFormat = "#
Worksheets("Sheet1").Cells(y + k, x + j).Value = TempValue
```

excel   vba

Figure 1: example question body

Stack Overflow Mechanisms continued. . .

6

The problem as discovered in Chat was the workbook had more than 64,000 formats because of which the user was getting the "**Too many different cell formats**" error message in Excel
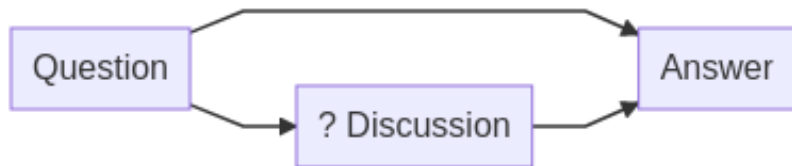
Solution

Topic: You receive a "Too many different cell formats" error message in Excel

Link: http://support.microsoft.com/kb/213904

Figure 2: answer for example question

Question-answering process

Discussions for questions

2   Also what is the value of `y` , `k` , `x` , `j` at the time of error?
    – Siddharth Rout May 29 '12 at 14:55

1   Is your Sheet Protected? – Siddharth Rout May 29 '12 at 14:57

1   Did you check if it is protected? – Siddharth Rout May 29 '12 at 15:03

1   let us continue this discussion in chat – Siddharth Rout May 29 '12 at 15:11

show **3** more comments

Figure 3: discussion in the comment thread

Extended discussions leads to chat rooms

The strange thing is that I havent changed anything. I was considering including more cells with number formatting, but I never implemented it :s

May 29, 23:28

See this link
support.microsoft.com/kb/…

ahhhh. I know why I never got the message earlier

Figure 4: discussion in the chat room

Reverse engineering the chat rooms mechanism

3 back and forth comments between two users

Is there a reason why databases such as `MySQL` and `MongoDB` doesn't work for you? – Cosmos Zhu 11 mins ago

@CosmosZhu I've looked around and found that having `MySQL` or `MongoDB` would mean the electron app can't be "portable" so that it can be distributed easily. – Cosmos547 10 mins ago

What about `sqlite` it stores the database in a file and both `python` and `node` have modules which can interact with a `sqlite` database. – Cosmos Zhu 9 mins ago

@CosmosZhu `sqlite` does seem to work. Do you have other suggestions for a database that's also in JSON form like tinydb? – Cosmos547 2 mins ago

Is there a reason you want to create the database in Python? If using javascript to both create the database then the option to choose a database will be so much easier. – Cosmos Zhu 1 min ago   Edit

@CosmosZhu I've already implemented the processing stage by Python and only want to use it in electron. I guess it makes sense to port it into javascript so that the same database can be shared. Do you have any suggestion for a database for javascript that would serve the need? – Cosmos547 4 secs ago

Please avoid extended discussions in comments. Would you like to automatically move this discussion to chat?
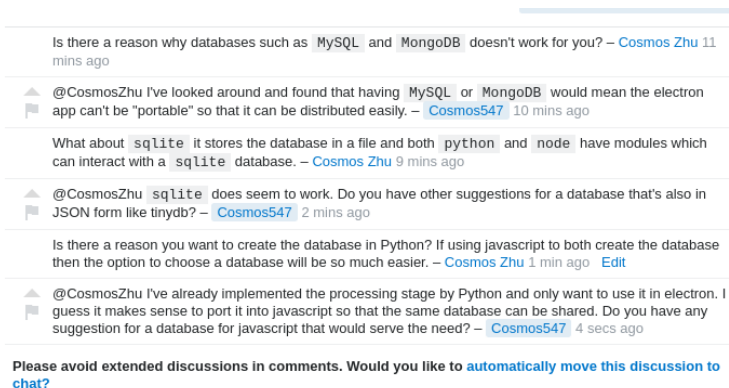
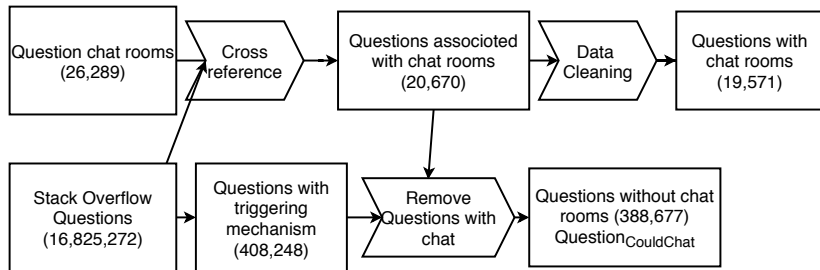Figure 5: triggering the option to create a chat room

Dataset construction

Figure 6: dataset construction

Research Questions

1. How prevalent are question chatted on Stack Overflow?
2. How do discussions migrate from commenting to chatting?
3. How does chatting affect question answering activities?

How prevalent are question chatted on Stack Overflow?

Discussions in comments

33.5 million comments and 1.4 million messages associated with Stack Overflow questions.

Type of questions with comments	percentage with discussions before answers
Questions with answers	75.4% of 8.2 million
Questions with accepted answers	80.0 of 3.9 million

Discussion in chat rooms

408,246 questions have enough conversational comments to trigger the option to create a chat room.

A chat room was not created for 91.3% of the questions.

Type of questions with chat rooms	percentage with discussions before answers
Questions with answers	76.5% of 14,978
Questions with accepted answers	80.0% of 9,996

How do discussions migrate from commenting to chatting?

Discussions move into chat rooms once they are created.
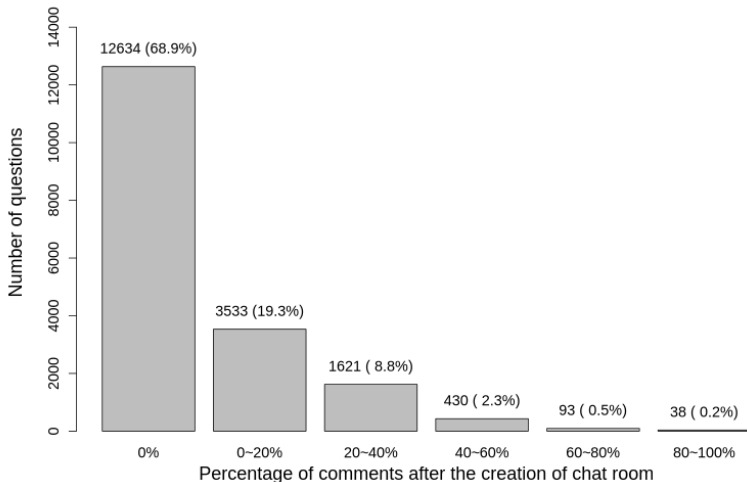
Figure 7: Distribution of the comment percentage after the creation of the chat room

Discussions in chat rooms are mainly between 2 users, and they have 3.6 times the amount of messages.

In 90.0% (i.e., 17,616) of chat rooms, two users participates in the chat room.
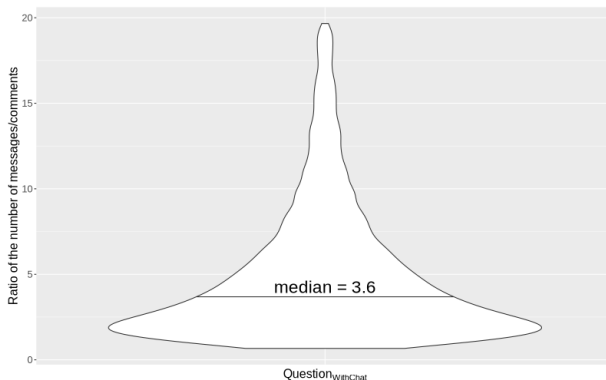
Figure 8: Distribution of the ratio: # Messages over # comments per question

More than half (i.e., 55.2%) of chat rooms are not created immediately after the option is available.
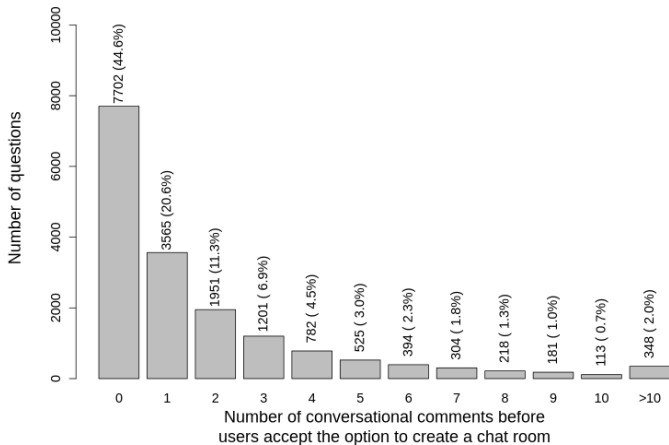
Figure 9: Distribution of additional conversational comments before chat room creation

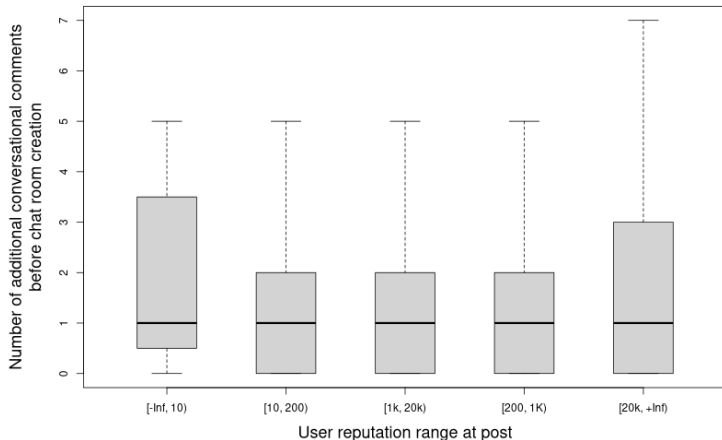# Reputation does not affect users' behavior related to chat room creation

Figure 10: Reputation and additional conversational comments before chat room creation

How does chatting affect question answering activities?

After the trigger event/creation of chat room questions leveraging chat rooms are more likely to receive accepted answers.

44.5% (i.e., 7,634) of $Question_{WithChat}$ receives accepted answers after the creation of a chat room.

14.8% (i.e., 35,466) of $Question_{IgnoredChat}$ receives accepted answers after triggering the option to create a chat room.

Chatting allow a more responsive environment for discussions and is dense in information

Users on average (median) respond to each other 5.2 times faster in messages compared to comments in the same question.

Messages are 4.7 times the density of comments on average (median) for the same question.

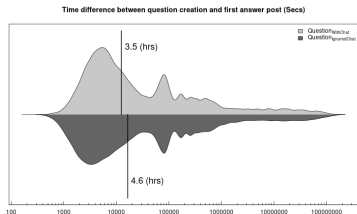# Creating chat rooms is associated with faster answers

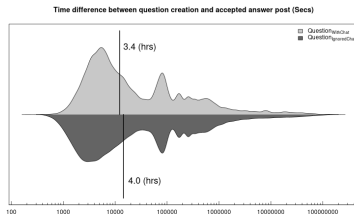Figure 11: Distribution of time difference from question post to first answer

Figure 12: Distribution of time difference from question post to accepted answer
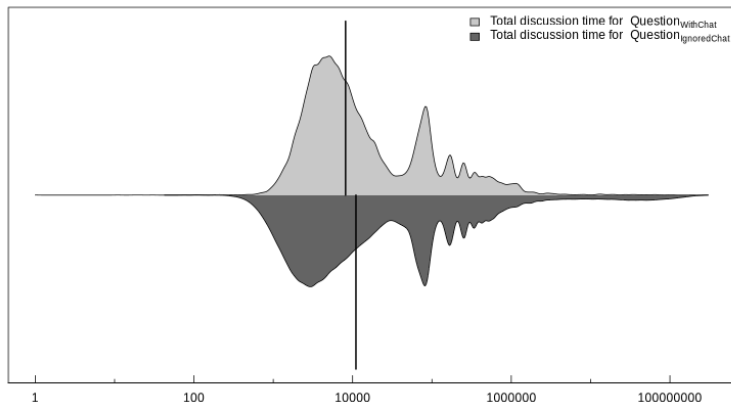
Leveraging chat rooms saves the total discussion time

Figure 13: Distribution of total discussion time for $Question_{WithChat}$ and $Question_{IgnoredChat}$

Implications and discussions

Importance of discussions

- ► Answers do not come from nowhere, many of them are created following discussions
- ► While commenting is the dominate communication channel for discussions, users also have the choice of chat rooms

Chat rooms on Stack Overflow

- ▶ Users are not aware of the feature
- ▶ Leveraging chat rooms will raise the possibility of receiving accepted answers
- ▶ Contains many and important information but often hidden and requires the access to separate webpage

Thank you!