Constructing
Effective
Stack
Overflow
Questions

Jane Hsieh

Constructing Effective Stack Overflow Questions

Jane Hsieh

Fall 2019 Honors Presentation

Advisor: Cynthia Taylor

Constructing Effective Stack Overflow Questions

Jane Hsieh

 \blacksquare Substantial portion of barriers that impede Stack Overflow users are significantly problematic for females (every 5 of 14, or $\approx 35\%)$

Constructing
Effective
Stack
Overflow
Questions

- \blacksquare Substantial portion of barriers that impede Stack Overflow users are significantly problematic for females (every 5 of 14, or $\approx 35\%)$
- Previous research focuses primarily on improving answer quality

Constructing
Effective
Stack
Overflow
Questions

- I Substantial portion of barriers that impede Stack Overflow users are significantly problematic for females (every 5 of 14, or $\approx 35\%$)
- Previous research focuses primarily on improving answer quality
 - Mining Successful Answers in Stack Overflow
 - Perceptions of Answer Quality in an Online Technical Question and Answer Forum
 - Eliciting Answers on StackOverflow
 - Moving to Stack Overflow: Best-Answer Prediction in Legacy Developer Forums
 - Towards predicting the best answers in community-based question-answering services
 - Predicting best answerers for new questions: An approach leveraging topic modeling and collaborative voting

Constructing
Effective
Stack
Overflow
Questions

Jane Hsieh

- Substantial portion of barriers that impede Stack Overflow users are significantly problematic for females (every 5 of 14, or $\approx 35\%$)
- Previous research focuses primarily on improving answer quality
 - Mining Successful Answers in Stack Overflow
 - Perceptions of Answer Quality in an Online Technical Question and Answer Forum
 - Eliciting Answers on StackOverflow
 - Moving to Stack Overflow: Best-Answer Prediction in Legacy Developer Forums
 - Towards predicting the best answers in community-based question-answering services
 - Predicting best answerers for new questions: An approach leveraging topic modeling and collaborative voting
- 3 Many questions (3 out of every 10) remain unanswered

As of today, according to stackexchange.com/sites

Constructing Effective Stack Overflow Questions

Jane Hsieh

- Substantial portion of barriers that impede Stack Overflow users are significantly problematic for females (every 5 of 14, or $\approx 35\%$)
- Previous research focuses primarily on improving answer quality
 - Mining Successful Answers in Stack Overflow
 - Perceptions of Answer Quality in an Online Technical Question and Answer Forum
 - Eliciting Answers on StackOverflow
 - Moving to Stack Overflow: Best-Answer Prediction in Legacy Developer Forums
 - Towards predicting the best answers in community-based question-answering services
 - Predicting best answerers for new questions: An approach leveraging topic modeling and collaborative voting
- Many questions (3 out of every 10) remain unanswered

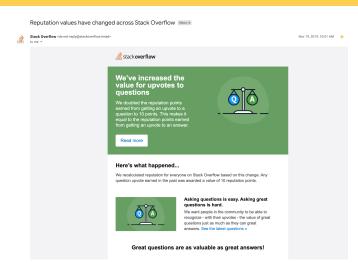
As of today, according to stackexchange.com/sites

4 More effective questions are needed

Call for questions

Constructing Effective Stack Overflow Questions

Jane Hsieh



Last month, the Stack Overflow forum itself to started rewarding upvoted questions with more reputation points.



Literature Review - Barriers

Constructing
Effective
Stack
Overflow
Questions

- Barriers on SO vary by gender, usage & experience
- Muddy lens perspective
 - lack of experience w/ site
 - unsure about the existence of unanswered questions
 - fear of being judged for "slacking" at work
 - fear of not ever receiving a quality answer or adding a duplicate post
- Impersonal interactions
 - fear of being judged
 - fear of relying on strangers rather than friends
- On-ramp roadblocks
 - difficulties in making the questions free of proprietary information / making sure there doesn't already exist a duplicate post
 - struggles with time constraints and self-confidence with expertise in the topic
 - effort required to learn proper etiquette of the community

Other Factors Common to Unanswered Questions

Constructing
Effective
Stack
Overflow
Questions

Jane Hsieh

- Too short, vague, or hard to follow
- Too hard, specific, or time-consuming
- Too specific and without context
- Duplicate question
- Proprietary technology
- Impatient, irregular, inconsiderate answerers / participants
- Lack of answer
- Fails to attract an expert member (22%)
- Works for me!
- hw questions

Asaduzzaman, Muhammad, et al. "Answering questions about unanswered questions of stack overflow."

Proceedings of the 10th Working Conference on Mining Software Repositories. IEEE Press, 2013.

Literature Review - success factors

Constructing
Effective
Stack
Overflow
Questions

Jane Hsieh

Factors that predict the success of an answer include

- answerer's reputation
- presentation quality (code snippets, URLs)
- not so much affect

Calefato, Fabio, et al. "Mining successful answers in stack overflow." Proceedings of the 12th Working Conference on Mining Software Repositories. IEEE Press, 2015.

Factors of questions that elicit good answers

- question type
- technology in question
- user identity
- time and day of asking
- code snippet presence
- length

Treude, Christoph, Ohad Barzilay, and Margaret-Anne Storey. "How do programmers ask and answer questions on the web?: Nier track." 2011 33rd International Conference on Software Engineering (ICSE).

Measuring Successful Questions

Constructing
Effective
Stack
Overflow
Questions

Jane Hsieh

- Metric for Questions:
 - $= m = \frac{\text{score}}{\text{view count}}$
 - score = upvotes downvotes
 - Successful questions are those with m > 0.13
 - Questions needing improvement have m < -0.2
- 2 Verification: Two-proportions z-test used to compare samples due to:
 - 1 binary data
 - 2 simple random sampling
 - 3 independent samples

Result: determines whether difference between two proportions is significant

Results

Constructing
Effective
Stack
Overflow
Questions

Jane Hsieh

	Code	Length	Attempt	Link
Confidence	99.92%	96.2%	82.2%	52.2%

■ Successful Questions Sample Size: 1057

■ Improvement-needed Questions Sample Size: 1568

■ Sampling size: 52

■ Number of trials: 1000

■ Attempt-signifying words: attempt, try, tried, tries

■ Medium length of post: 1350 characters

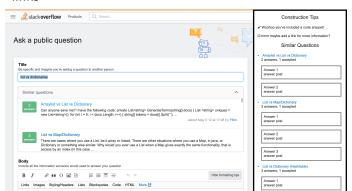
Preliminary Implementations

Constructing
Effective
Stack
Overflow
Questions

Jane Hsieh

Chrome Sidebar Features

- Previews relevant posts and answers based on current question title
- Assesses current state of question formation and suggest actionable improvements such as adding code snippets or links





Future steps

Constructing
Effective
Stack
Overflow
Questions

- Externally query Stack Overflow to construct a preview of similar questions' answers (WT)
- Dynamically assess question quality as users formulates their inquiries (WT & Spring)
- Using current set of successful & low-ratio questions, apply classification/regression techniques to provide feedback by giving a probable score for each question under construction (WT & Spring)