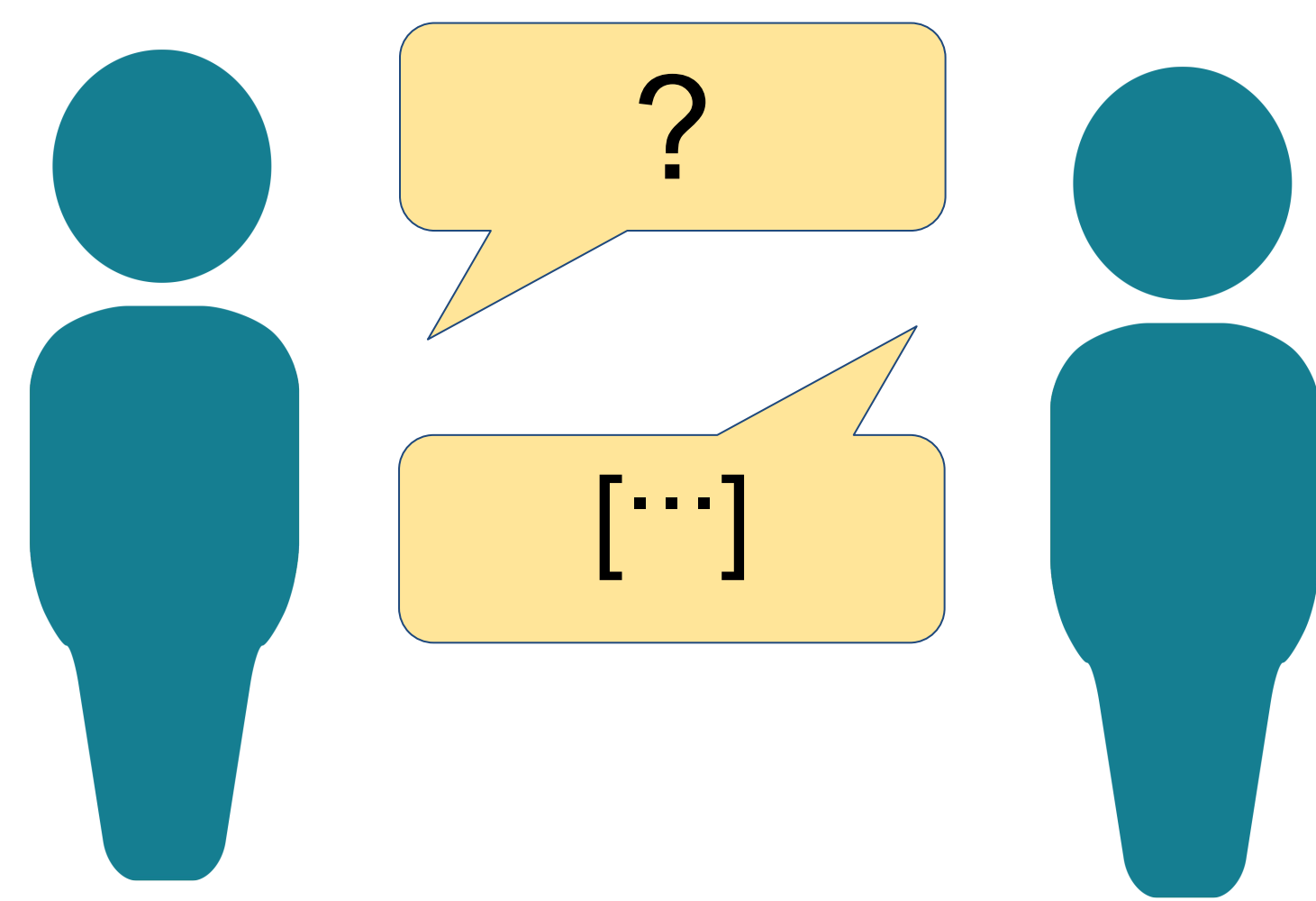


**Samantha Lee**  
Advisor: Dr. Chirag Shah

## ABSTRACT

In recent years, online Q&A services have grown in popularity. By posing and answering questions, users have created whole communities based around information exchange. Researchers seek to understand these communities and must consider the problems users face, as well as the factors involved in a successful search for information. A literature search was conducted to assess the different approaches to improve community Q&A platforms. Three publications were selected for the review. They examine datasets from popular sites to study various factors. The methodologies implemented in the studies primarily feature predictive frameworks and mixed-method analysis. All three studies achieved conclusive results and thus greater understandings of community Q&A services and how to enhance them.



### RESEARCH QUESTIONS

1. How can one derive the quality of an answer in a Q&A environment?
2. How can a system seek out potential answerers?
3. How does a user's expectations affect his or her information-seeking behaviors?

## METHODS

The publications selected for this literature satisfy the following criteria:

- published by InfoSeeking
- published between 2016 and 2018
- focus on community-based Q&A (CQA)

## METHODS

**Table 1: List of features classified into four groups.**

Feature	Examples
Personal	Number of answers given, number of questions asked, ranking of users
Community Based	Number of thanks that user received, number of warnings that user received, number of spam reports that user received, number of friends in community
Textual	The length of answer, the readability of answer, the format of answer
Contextual	The grade level of question, the grade difference between answerer & question, The rank difference between answerer & asker

**Table 2: List of features classified into four groups.**

Features
Similarity between question content and user profile
Similarity between question topics and user expertise topics
Similarity between asker and answerer in the information network
User's activity level

**Table 3: List of common user expectations when seeking information.**

Expectation-Based Factors
Looking for quick response
Looking for additional or alternative information
Looking for accurate or complete information
Looking for social or emotional support
Looking for verification for own belief of knowledge
Looking for trustworthy sources

### Evaluating Answer Quality - dataset from Brainly

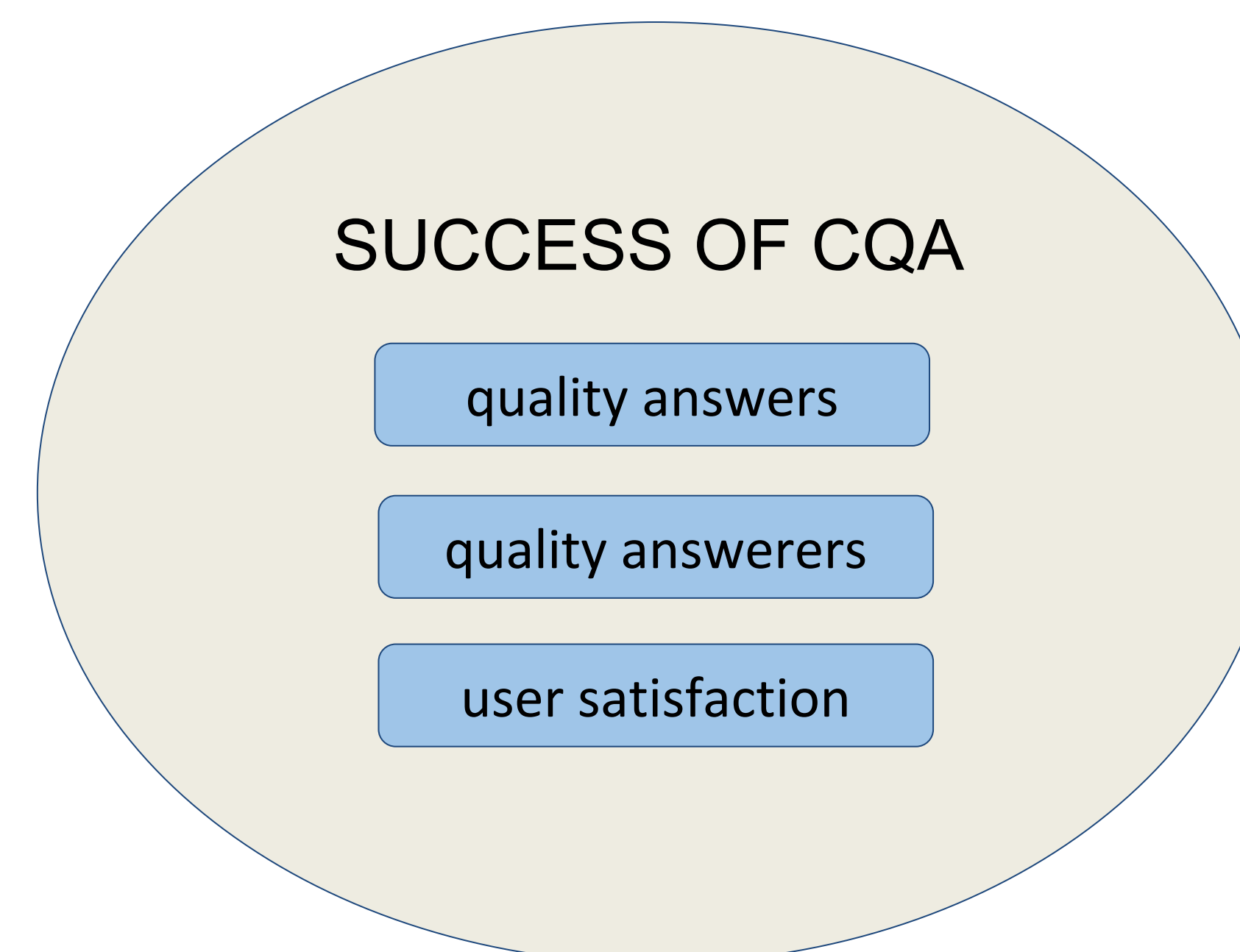
- Classify the quality of answers by features
- Measure accuracy (percentage of answers classified correctly)
- Compare different algorithms using framework

### Determining Potential Answerers - datasets from Stack Overflow and Yahoo! Answers

- Develop algorithm (QRec) to compute scores between a given question and all user profiles
- Measure correctness (likelihood that at least one potential answerer will answer)
- Compare against different algorithms

### Assessing User Expectations - datasets from Yahoo! Answers and WikiAnswers

- Collect behavioral data via open-ended prompts, self-assessments, questions about online behavior, phone interviews, diary entries
- Measure expectation-based factors (identified in previous literature) on a 5-point Likert scale



## RESULTS

### Evaluating Answer Quality

- Personal and community features best predict the quality of an answer
- Random Forest algorithm most successful

### Determining Potential Answerers

- QRec which considers the content similarity, topic similarity, and topics and the user similarity yields highest correctness

### Assessing User Expectations

- Users sought information to find (in order of frequency) quick responses, alternative information, and complete information

## DISCUSSION

- There are many ways to improve Q&A communities and thus enhance the platforms as a whole
- While there is a breadth of factors to focus on, several stand out as notable influencers when it comes to answer quality and determining answerer qualities
- Users tend to focus on themselves when conducting a search

## FUTURE DIRECTIONS

- Focus on
  - a user's unique situation in relation to expectations
  - a user's interests
  - struggling users in the CQA communities
- Coagmento (data collecting browser extension) development



## REFERENCES

1. Choi, E., and Shah, C. (2017). Asking for more than an answer: What do askers expect in online Q&A services? *Journal of Information Science*, 43(3), 424-435. <https://doi.org/10.1177/0165551516645530>
2. Le, L. T. and Shah, C. (2018), Retrieving people: Identifying potential answerers in Community Question-Answering. *Journal of the Association for Information Science and Technology*, 69: 1246-1258. doi:10.1002/asi.24042
3. Le, L. T., Shah, C., and Choi, E. (2016). Evaluating the Quality of Educational Answers in Community Question-Answering. *JCDL '16 Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries*, 129-138. <https://doi.org/10.1145/2910896.2910900>

## ACKNOWLEDGEMENTS

I would like to thank Dr. Chirag Shah and Dr. Matthew Mitsui for their mentorship and support, as well as to the Douglass Project for providing guidance through Project SUPER, the Introduction to Scientific Research class, and the peer mentor program.