# StackExchange Data Analytics

Project by :- Ashwin Bhide, Chaitanya Bapat
Nidhi Menon, Sneha Venkat, Vaibhav Tendulkar

# Current Situation

- Questions go unanswered
- People answer for “points” rather than to provide the correct answer
- Absence of question routing to experts

# Objectives

- Given a question, find expert users to answer
- User engagement driven ML solution
  1. Visualizing StackExchange as a Graph
  2. Applying ML algorithms

# Literature Survey

- Understanding StackOverflow using Quantitative, Qualitative analysis

  -“Design Lessons from the Fastest Q&A Site in the West”

- Recommends questions to users based on their novel algorithm

  -“Question Recommendation for Collaborative Question Answering Systems with RankSLDA”

# Literature Survey (contd.)

- Feature-based recommendation approach using Random Forest classifier

  -“Who will Answer my Question on Stack Overflow?”

- A generative modelling approach combining topic-level information with term-level similarity for expert recommendation

  -“Tapping on the Potential of Q&A Community by Recommending Answer Providers”

# Our Approach

- Feature extraction to identify experts
- Analyze user engagement using visualization software
- Reason for Success :
  - Increase in answer accuracy due to intelligent question routing

# Our Approach (contd)

- Measure of Success - Ratio of predicted users to groundtruth
- Risks and Payoffs - Accuracy and Time
- Cost - None (open source)
- Time - 45 days

# Timeline

- Milestone 1
  - Data Cleaning, Feature Extraction
  - Preliminary Data Analysis and Visualization
- Milestone 2
  - Fine Tuning Data Analysis
  - Final Visualizations

# Thank You!

-Messiahs