

# Signals Matter: Understanding Popularity and Impact of Users on Stack Overflow

## Summary

Stack Overflow is a QnA portal for programmers that awards its users with some points and badges upon completing some tasks. This is actually a form of using game design elements in the non-game context which encourages an individual to contribute more to the community. This paper focuses on two important questions, i.e., what social qualities (if any) do reputation scores and badges intend to signal, and to what extent do these game elements actually signal or indicate the qualities that users expect them to?

At first, researchers conducted an online survey of Stack Overflow users. The results from the survey showed the increasing trust of programmers towards Stack Overflow. Since this research primarily focuses on two factors triggered by these game elements: popularity and impact, thus, from the results, it was obtained that people having low reputation scores often consider reputation scores to better indicators of popularity than badges, and for people with high reputation scores, it's vice-versa (humbleness? maybe). A majority of them agreed that badges are actually important, and reputation scores are impactful.

Researchers also conducted experiments on a publicly available dataset of 3831147 users and the large complete time-stamped history of their actions on Stack Overflow spanning a decade. The dataset comprised the following attributes: Reputation Scores, Badges, and Active Users. Based on the survey, researchers made some hypotheses and tested them.

Researchers worked on three non-linear regression models for predicting the popularity and impact of users based on reputation scores and the top five most important badges. In the results which they got, they found that reputation points are good predictors of both impact and popularity. Yet, Badge Model performed better in predicting signals of impact than predicting signals of popularity. Also, in each of the cases, all of the five important badges were Answer badges. The reputation model was faster than the badges model, yet the badges model gave more accurate results because it used decision trees and had more features.

They also differentiated popular(P) and impactful(I) users based on these game elements. They found out that a majority of users belonging to LPHI, HPLI, and HPHI (L-Low, H-High) groups have more badges for well-received answers as compared to questions. The results showed that answers derive popularity whereas questions offer more influence.

To conclude, the results that the researchers got are satisfactory, and it proves the fact that game symbols actually helps to filter out the quality community on a platform. But there are some limitations to this research. First, the metrics are biased towards the number of user views and posts. Second, it focuses only on reward-based features and does not incorporate content-based features. And as per my opinion, if they would have used more badges as a metric, the results would have been more accurate. But this research paper is not the end. In the future, one can examine the linguistic attributes of posts and can further explore the role of game elements as symbols of social status in socio-technical systems.

**Shlok Pandey**  
**2020121008**

---