

# A Manual Categorization of Android App Development Issues Using Stack Overflow Posts

Stefanie Beyer, Martin Pinzger

University of Klagenfurt

The discussion of issues related to the development of mobile applications (apps) has gained more and more popularity on Q&A-platforms such as Stack Overflow.<sup>1</sup> Barua *et al.* [1] stated that Android is among the topics with the largest increase in the number of posts on Stack Overflow. The success of a mobile application depends on the quality of the application. Lineares-Vasquez *et al.* [2] found that there is a dependency between the quality of an API and the success of the mobile app that uses this API. The APIs used by successful apps use less fault-prone APIs than those used by unsuccessful apps. Although there are many tutorials, a lot of documentation and several examples on how to develop mobile applications, developers often have problems and questions concerning their implementation. The goal of this research is to give recommendations to developers to improve the quality of mobile applications and to minimize the amount of problems during development. As a first step in this research, we need to find out what the main problems and topics of Android app developers are.

We make the following contributions:

- A qualitative evaluation of Android development issues concerning the main problem and question categories.
- An evaluation of the dependencies between the problems and question categories of posts.
- A manually created benchmark for Android-related post classification.
- An initial evaluation of Apache Lucene’s kNN algorithm to automate the classification.

We manually investigated 450 Android related posts of Stack Overflow to get information about the main issues of Android development. In particular, we used *Card Sorting* to categorize posts concerning the issues stated in the post and the kind of question that is asked. We refined the categories iteratively and evaluated them with three experienced Android app developers. Fleiss’ Kappa is calculated to measure the inter-rater agreement. Furthermore, we used the classified posts to train a model with Apache Lucene’s implementation of kNN (k-nearest-neighbours)-algorithm in order to perform this classification automatically.

Analyzing the questions and problems we found that developers mainly have problems with the usage of API components, such as *User Interface* and *Core Elements*. Developers also ask if the realization of their ideas is possible with

---

<sup>1</sup> <http://stackoverflow.com/>

*Libs/APIs* and *User Interface* components. Errors are often mentioned in questions related to *Network*, *Database*, and *Fragments*. Furthermore, version changes cause problems in the components *Webkit* and *ActionBar*. Based on our categorization we then investigated to which extent the classification of posts can be automated with Apache Lucene’s kNN algorithm. While the classification obtained with Lucene significantly outperformed the baseline, it needs to be improved.

The three approaches closest to this study are from Linares-Vasquez *et al.* [3], Barua *et al.* [1], and Treude *et al.* [4]. We differ from these approaches by using card sorting for the categorization of the posts instead of automated approaches, such as LDA. We furthermore differ from related studies in investigating the dependencies between question and problem categories to get better insights into the issues of Android app development.

In future work we first will expand this study by an investigation of more Android related posts on Stack Overflow. We also will compare our manual categorization with categorizations obtained through IR-techniques, such as LDA and LSI (Latent Semantic Indexing). We will use these techniques to also improve the automated classification of posts. Based on the categorization, we then plan to investigate the evolution of questions and problems discussed on Stack Overflow.

## References

1. Anton Barua, Stephen W. Thomas, and Ahmed E. Hassan. What are developers talking about? an analysis of topics and trends in stack overflow. *Empirical Software Engineering*, 19(3):1–36, 2012.
2. Mario Linares-Vásquez, Gabriele Bavota, Carlos Bernal-Cárdenas, Massimiliano Di Penta, Rocco Oliveto, and Denys Poshyvanyk. Api change and fault prone-ness: A threat to the success of android apps. In *Proceedings of the Joint Meeting on Foundations of Software Engineering*, pages 477–487. ACM, 2013.
3. Mario Linares-Vásquez, Bogdan Dit, and Denys Poshyvanyk. An exploratory analysis of mobile development issues using stack overflow. In *Proceedings of the 10th Working Conference on Mining Software Repositories*, pages 93–96. IEEE Press, 2013.
4. Christoph Treude, Ohad Barzilay, and Margaret-Anne Storey. How do programmers ask and answer questions on the web? (nier track). In *Proceedings of the International Conference on Software Engineering*, pages 804–807. ACM, 2011.