

1. AndroZooOpen: collecting large-scale open source android apps for the research community

Pei Liu (1); Li Li (1); Yanjie Zhao (1); Xiaoyu Sun (1); Grundy, J. (1)

Source: MSR '20: *Proceedings of the 17th International Conference on Mining Software Repositories*, p 548-52, 29 June 2020; **ISBN-13:** 978-1-4503-7517-7; **DOI:** 10.1145/3379597.3387503; **Conference:** MSR '20: 17th International Conference on Mining Software Repositories, 29-30 June 2020, Seoul, South Korea; **Sponsor:** SIGSOFT; **Publisher:** ACM, New York, NY, USA

Author affiliation: (1) Monash University, Melbourne, VIC, Australia

Abstract: It is critical for research to have an open, well-curated, representative set of apps for analysis. We present a collection of open-source Android apps collected from several sources, including Github. Our dataset, AndroZooOpen, currently contains over 45,000 app artefacts, a representative picture of Github-hosted Android apps. For apps released on Google Play, metadata including categories, ratings and user reviews, are also stored. We share this new dataset as part of our ongoing research to better support and enable new research topics involving Android app artefact analysis, and as a supplement dataset for AndroZoo, a well-known app collection of close-sourced Android apps. (0 refs)

Inspec controlled terms: Android (operating system) - meta data - mobile computing - public domain software

Uncontrolled terms: AndroZooOpen - open source Android apps - Github-hosted Android apps - close sourced Android apps - Google Play - metadata

Classification Code: C6190V Mobile, ubiquitous and pervasive computing - C6150J Operating systems

IPC Code: G06F9/44 - G06F9/46

Treatment: Practical (PRA)

Database: Inspec

Data Provider: Engineering Village

Copyright 2020, The Institution of Engineering and Technology