1. Knowledge Graphing Git Repositories: A Preliminary Study

Yanjie Zhao (1); Haoyu Wang (1); Lei Ma (2); Yuxin Liu (3); Li Li (3); Grundy, J. (3)
Source: *2019 IEEE 26th International Conference on Software Analysis, Evolution and Reengineering (SANER). Proceedings*, p 599-603, 2019; **ISBN-13:** 978-1-7281-0591-8; **DOI:** 10.1109/SANER.2019.8668034; **Conference:** 2019 IEEE 26th International Conference on Software Analysis, Evolution and Reengineering (SANER), 24-27 Feb. 2019, Hangzhou, China; **Publisher:** IEEE, Piscataway, NJ, USA
Author affiliation: (1) Beijing University of Posts and Telecommunications, China (2) Harbin Institute of Technology, China (3) Monash University, Melbourne, VIC, Australia
Abstract: Knowledge Graph, being able to connect information from a variety of sources, has become very famous in recent years since its creation in 2012 by Google. Researchers in our community have leveraged Knowledge Graph to achieve various purposes such as improving API caveats accessibilities, generating answers to developer questions, and reasoning common software weaknesses, etc. In this work, we would like to leverage the knowledge graph concept for helping developers and project managers to comprehend software repositories. To this end, we design and implement a prototype tool called GitGraph, which takes as input a Git repository and constructs automatically a knowledge graph associated with the repository. Our preliminary experimental results show that GitGraph can correctly generate knowledge graphs for Git projects and the generated graphs are also useful for users to comprehend the projects. More specifically, the knowledge graph, on one hand, provides a graphic interface that users can interactively explore the integrated artefacts such as commits and changed methods, while on the other hand, provides a convenient means for users to search for advanced relations between the different artefacts. (0 refs)
Inspec controlled terms: application program interfaces - graph theory - knowledge management - software engineering - software tools
Uncontrolled terms: knowledge graph concept - knowledge graphing git repositories - API caveats accessibilities - GitGraph - Git projects
Classification Code: C6110B Software engineering techniques - C1160 Combinatorial mathematics - C6150J Operating systems
IPC Code: G06F9/44 - G06F9/46
Treatment: Practical (PRA) - Theoretical or Mathematical (THR)
Database: Inspec
Data Provider: Engineering Village