

A Comparison Study of Open Source License Crawlers

Abstract

Open-source licenses allow source code, blueprint or design to be used, modified and/or shared under defined terms and conditions. However, there is no standard where to place the license text. Different tools such as Fossology or [Go License Detector](#), [license classifier by Google](#), [license checker \(ls\)](#), and others attempt to detect licenses within source code - using different approaches and results. In this thesis, we evaluated existing licenses crawlers with respect to design, precision/recall, and computational performance on a test sample of 1,000 top GitHub projects. {RESULTS OF THE BENCHMARK}