CS 3050: Software Engineering
Professor Hibbeler
Samuel Clear, Michael Piscione, Max Pursche, Kyrill Serdyuk
09/08/2023

GitHub Repo Database Description

We chose our database to be highly-rated OOP GitHub Repos to take a look into what keywords and topics attract users to these repositories! Below, each of the entities within each document of the JSON file are defined:

id: Unique identifying String variable representing the entity in GitHub
repo_name: A String variable for the name of the repository in GitHub
description: String variable representing the given description of the project in GitHub
created: String variable representing the date of creation
language: Optional String variable representing the programming language used in the repository
type: String variable representing the type of user submitting the repository to GitHub
stars: Integer variable counting the total number of stars that each user has given the repository in total
forks: Integer variable counting the total number of forks that each user has performed with the given repository in total
subscribers: Integer variable counting the total number of subscribers to the repository
topics: Optional List variable containing a collection of topics chosen by the owner of the repository to describe the GitHub Repository

The optional subcollection which is within each Repository entity is called "topics". This subcollection is a collection of topics each consisting of a parent group and the specific topic name. Each of these topics are either created uniquely by the user, or they are chosen from a list of popular topics given to other projects in GitHub.