

Project Summary Report

Tong Zhang 11911831

Yang Yibo 12112222

Abstract

In this final project, we have used Spring Boot, Mysql, to develop a web application that stores, analyzes, and visualizes github repo data we have obtained from Github REST API. The aim of our project is to provide a web application for users who want to know key information about a specific github repo. For instance, total developers number, top 10 active developers info(based on their contribution), number of open issues and closed issues, released version, statistical analysis of commit amounts during each published version as well as analysis of solved issues (presents typical processing of issue resolution time, such as average, extreme difference, variance value, etc.) The presented items mentioned above are all key insights for users to familiarize with a specific github repo.

Set Up

Specifically speaking, the Github REST APIs in our project include:

- 1stDevelopers: <https://api.github.com/repos/mhalbritter/spring-boot/contributors>
- 1stIssues: <https://api.github.com/repos/CMU-Perceptual-Computing-Lab/openpose/issues>
- 1stReleases: <https://api.github.com/repos/CMU-Perceptual-Computing-Lab/openpose/releases>
- 1stCommits: <https://api.github.com/repos/matkob/openpose/commits>
- 2ndDeveloper: <https://api.github.com/repos/ShiqiYu/libfacedetection/contributors>
- 2ndIssues: <https://api.github.com/repos/ShiqiYu/libfacedetection/issues>
- 2ndReleases: <https://api.github.com/repos/ShiqiYu/libfacedetection/releases>
- 2ndCommits: <https://api.github.com/repos/ShiqiYu/libfacedetection/commits>

To make it work, we built a Spring Boot application in java. The structure design is shown below:

CS209-project-22fall [drug] D:\ProProject\JavaProject\ProjDemo\CS209-project

- > .idea
- > .mvn
- ▼ src
 - > DraftJSON
 - ▼ main
 - ▼ java
 - ▼ com.example.springproject
 - ▼ common
 - DateUtil
 - ▼ crawl
 - ClosedIssueInfo
 - ClosedIssueInfo2
 - CommitsInfo
 - CommitsInfo2
 - ContributorInfo
 - ContributorInfo2
 - Druid
 - OpenIssueInfo
 - OpenIssueInfo2
 - ReleaseInfo
 - ReleaseInfo2
 - > domain
 - ▼ service
 - RepoService
 - RepoServiceImpl
 - ▼ web
 - helloTest
 - RepoController
 - SpringprojectApplication
 - ▼ resources
 - ▼ static
 - index.html
 - index2.html
 - Picture.jpg
 - Picture1.jpg
 - application.properties
 - > test
 - > target
 - application.yml
 - LICENSE
 - pom.xml
 - pom1.xml
 - README.md

Crawling

crawl package is for crawling data from Github REST API: The general flow of crawling can be described as follows:

1. Open database using **OpenDB(String database)** (repo)
2. Connect database using **Druid**, which is Alibaba's open source database connection pool project
3. Access url using a self-defined **readByURL** method, since the default amounts for Github REST API is 30 per_page. Hence, we manually set the **page number** and set the **per_page** value to 100 as follows, so that we can access all the items in API:
https://api.github.com/repos/CMU-Perceptual-Computing-Lab/openpose/issues?state=closed&page=1&per_page=100
4. To implement the **readByURL** method, we used **URLConnection**, and **HttpURLConnection**. After building the connection between **localhost** and URL, we used **JSONArray** and **JsonObject** to deal with the desired item (for example, contributions, date, state, name, title, body etc.) we wanted in crawled JSON files.
5. For all the acquired items from JSON files, we added them into SQL **PreparedStatement**. For each record, it is stored in a Batch, and every time we have the full column values for one row, we execute the **Batch**. After executing all the Batches, we finished the crawling process by **committing** the result to database.

Here are the illustrations of the 5 tables we have created:

closedissue

1-500 of 501+

WHERE

number

title

created_at

updated_at

body

1

2158

unity support ver...

2022-12-10

2022-12-11

Hi, I've rtx3070...

2

2158

unity support ver...

2022-12-10

2022-12-11

Hi, I've rtx3070...

3

2158

unity support ver...

2022-12-10

2022-12-11

Hi, I've rtx3070...

4

2147

Improve docs on C...

2022-10-18

2022-10-23

This commit adds ...

5

2145

Link fix in 3d mo...

2022-10-17

2022-10-18

Link broken when ...

6

2143

remote vscode : N...

2022-10-11

2022-10-12

I use vscode to c...

7

2139

Fixed typo in docs

2022-09-28

2022-09-30

<null>

8

2134

Tiyaro demo added...

2022-09-17

2022-09-17

<null>

9

2133

Bump requirements...

2022-09-16

2022-09-17

Openpose is teste...

10

2129

docs: demo, exper...

2022-09-06

2022-09-17

Hello OpenPose Ma...

11

2112

foot_v2.json Unab...

2022-07-18

2022-07-18

I have downloaded...

12

2107

Delete 1/6 evashne...

2022-06-20

2022-06-24

<null>

ORDER BY

Database

repo@localhost 1 of 8

repo

tables 5

closedissue

columns 5

number mediumtext

title varchar(1000)

created_at date

updated_at date

body varchar(10000)

commits

contributor

openissue

releases

Server Objects

commits

1-500 of 501+

WHERE

message

post_date

1

Doc link typos fix (#2145)

2022-10-17

2

Fixed typo in docs (#2139)

2022-09-30

3

Tiyaro demo added [as asked by...

2022-09-17

4

Bump requirements for Ubuntu 2...

2022-09-17

5

Update openpose_python.cpp (#2...

2022-06-19

6

Update README.md

2022-06-10

7

Update README.md

2022-06-10

8

Fix typos. (#2086)

2022-05-17

9

Typo fixed in 02_output.md (#2...

2022-03-24

ORDER BY

Database

repo@localhost 1 of 8

repo

tables 5

closedissue

commits

columns 2

message varchar(100)

post_date date

contributor

openissue

releases

Server Objects

contributor					
id	name	contributions	avatar_url	html_url	
1	7797094	513	https://avatars.g...	https://github.co...	
2	3193196	26	https://avatars.g...	https://github.co...	
3	4561285	11	https://avatars.g...	https://github.co...	
4	5785922	9	https://avatars.g...	https://github.co...	
5	18014833	6	https://avatars.g...	https://github.co...	
6	8867900	6	https://avatars.g...	https://github.co...	
7	18168769	5	https://avatars.g...	https://github.co...	
8	3905501	4	https://avatars.g...	https://github.co...	
9	12238661	3	https://avatars.g...	https://github.co...	
10	4690241	2	https://avatars.g...	https://github.co...	
11	22472380	2	https://avatars.g...	https://github.co...	

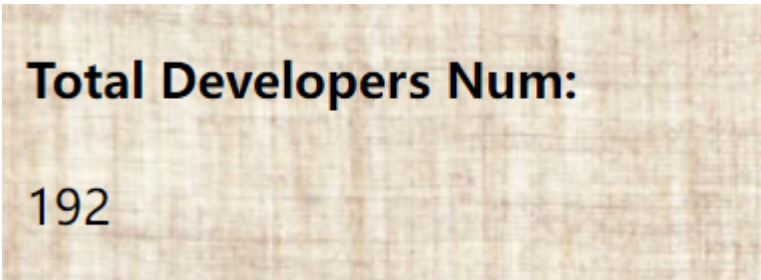
openissue					
number	title	created_at	updated_at	body	
1	2157	Facing cuDNN version error in ...	2022-12-07	2022-12-07	### Posting rules=1. **No dupl...
2	2156	Python API	2022-11-29	2022-12-16	Hi, I am trying to run some py...
3	2155	Hand Standalone Detector	2022-11-25	2022-11-25	Hi, I am trying to run the sta...
4	2154	Windows: Inference speed depen...	2022-11-24	2022-11-24	### Issue Summary=OpenPose inf...
5	2153	how to get it run in online st...	2022-11-16	2022-11-17	Installation guide provides de...
6	2152	You need to compile OpenPose w...	2022-11-10	2022-11-10	Have you ever encountered this...
7	2151	Unable to recognize the invert...	2022-11-07	2022-11-07	Hello, I tried some non-uprigh...
8	2150	How to obtain 25-keypoint body...	2022-11-06	2022-11-06	hello, anybody could help me w...
9	2149	(OSX) do not install szip	2022-11-04	2022-11-28	szip is removed, see https://g...
10	2148	EntryPointNotFoundException: ...	2022-11-02	2022-11-02	### Issue Summary=EntryPointNo...
11	2146	how to draw body_25	2022-10-18	2022-10-18	how to draw body_25=i have coo...

releases				
assets_url	created_at	published_at	name	
1	https://api.githu...	2020-11-17	2020-11-17	OpenPose v1.7.0
2	https://api.githu...	2020-04-27	2020-04-27	OpenPose v1.6.0
3	https://api.githu...	2019-09-04	2019-09-04	OpenPose v1.5.1
4	https://api.githu...	2019-05-16	2019-05-16	OpenPose v1.5.0
5	https://api.githu...	2018-09-02	2018-09-02	OpenPose v1.4.0
6	https://api.githu...	2018-03-24	2018-03-24	OpenPose v1.3.0
7	https://api.githu...	2018-01-09	2018-01-09	OpenPose v1.2.1
8	https://api.githu...	2017-11-03	2017-11-03	OpenPose v1.2.0
9	https://api.githu...	2017-09-20	2017-10-31	OpenPose v1.1.0
10	https://api.githu...	2017-09-03	2017-10-31	OpenPose v1.0.2

Front End

Our front end shows the following information:











1. The total number of developers.



2. The information of the top 10 active developers info(based on their contribution).

Top10 Developers:

>

Id	Name	Developers	avatar	avatar_url	Personal Page
7797094	gineshidalgo99	513		https://avatars.githubusercontent.com/u/7797094?v=4	https://github.com/gineshidalgo99
3193196	soulslicer	26		https://avatars.githubusercontent.com/u/3193196?v=4	https://github.com/soulslicer
4561285	bikz05	11		https://avatars.githubusercontent.com/u/4561285?v=4	https://github.com/bikz05
5785922	xiangyann	9		https://avatars.githubusercontent.com/u/5785922?v=4	https://github.com/xiangyann
18014833	MatthijsBurgh	6		https://avatars.githubusercontent.com/u/18014833?v=4	https://github.com/MatthijsBurgh
8867900	ZheC	6		https://avatars.githubusercontent.com/u/8867900?v=4	https://github.com/ZheC
18168769	ZiuTinyat	5		https://avatars.githubusercontent.com/u/18168769?v=4	https://github.com/ZiuTinyat
3905501	bryant1410	4		https://avatars.githubusercontent.com/u/3905501?v=4	https://github.com/bryant1410
12238661	subail	3		https://avatars.githubusercontent.com/u/12238661?v=4	https://github.com/subail
4690241	cngzhnp	2		https://avatars.githubusercontent.com/u/4690241?v=4	https://github.com/cngzhnp

3. The number of the open issues, the closed issues. The mean value of the solution time of issues, the extreme value of solution time, and the variance of solution time.

Issues :

Open Issues Num:	414
Closed Issues Num:	1865
Mean of Solution Time(Hours):	3873.1367
Extreme Value of Solution Time(Days):	1959
Variance of Solution Time(Days):	97601.8457908846

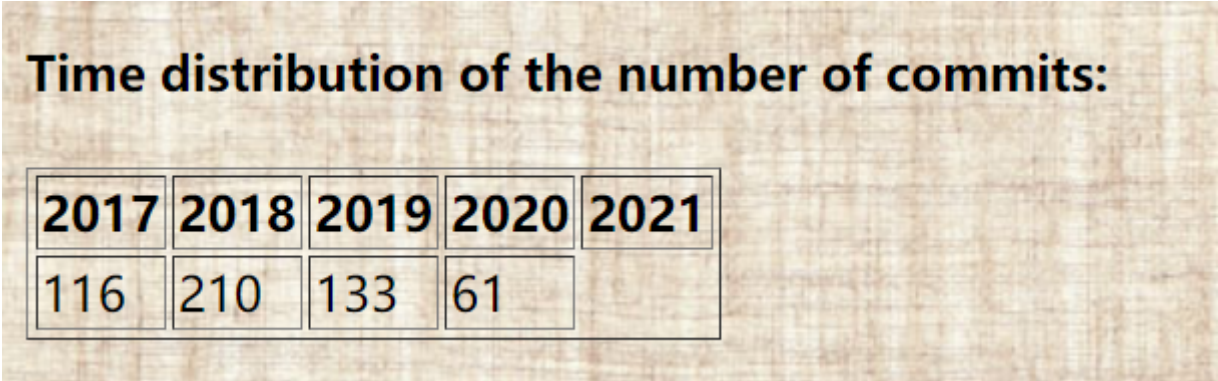
4. The number of the releases and the number of commits.

Releases Num:
15
Commits Num
566

5. The commit numbers between each two releases.

Assets url	Created at	Published at	Name
https://api.github.com/repos/CMU-Perceptual-Computing-Lab/openpose/releases/34039502/assets	2020-11-17	2020-11-17	OpenPose v1.7.0
commits num between last release and next release			32
https://api.github.com/repos/CMU-Perceptual-Computing-Lab/openpose/releases/25905045/assets	2020-04-27	2020-04-27	OpenPose v1.6.0
commits num between last release and next release			50
https://api.github.com/repos/CMU-Perceptual-Computing-Lab/openpose/releases/19734438/assets	2019-09-04	2019-09-04	OpenPose v1.5.1
commits num between last release and next release			22
https://api.github.com/repos/CMU-Perceptual-Computing-Lab/openpose/releases/17395810/assets	2019-05-16	2019-05-16	OpenPose v1.5.0
commits num between last release and next release			144
https://api.github.com/repos/CMU-Perceptual-Computing-Lab/openpose/releases/12689688/assets	2018-09-02	2018-09-02	OpenPose v1.4.0
commits num between last release and next release			87
https://api.github.com/repos/CMU-Perceptual-Computing-Lab/openpose/releases/10243963/assets	2018-03-24	2018-03-24	OpenPose v1.3.0
commits num between last release and next release			65
https://api.github.com/repos/CMU-Perceptual-Computing-Lab/openpose/releases/9169064/assets	2018-01-09	2018-01-09	OpenPose v1.2.1
commits num between last release and next release			27

6. Time distribution of the number of commits.



7.This web page can jump to another page, which contains the information of another repository that follows the same logic as the first page we described above.



Configuration

We add our dependencies in *pom.xml*, including 'org.springframework.boot', 'com.javaclimb', 'org.apache.httpcomponents', 'org.jsoup', 'us.codecraft', 'com.google.code.gson', 'org.projectlombok', 'mysql', 'com.alibaba', 'org.apache.shiro', 'junit', 'com.github.theborakompanion'. Besides pom.xml, we also added a 'json-simple-1.1.1.jar' as a *local library* so that we could analyze the JSON files we obtained from crawling.