

1322004 - Métricas de Software

Trabalho Prático

Pós-Graduação em Gerência de Projetos de Software na Era de Dados de
Sensores e IA

Guilherme Ramos Coelho (Matrícula 11322d0013)

Juiz de Fora, MG
Outubro de 2025

4) Anexo Técnico e Códigos Utilizados

5.1) Anexos Disponíveis

O projeto demonstrativo foi criado utilizando-se componentes como Git, jscpd e Radon. Os arquivos gerados por estas ferramentas estão disponíveis para conferência no diretório “analysis_data”:

- commits.json
- contributors_last_year.json
- issues_closed.json
- radon_cc.json
- radon_mi.json
- radon_raw.json
- jscpd-report/ jscpd-report.json

Adicionalmente o arquivo “metrics_summary.json” sumariza o resultado dos cálculos com base nos arquivos brutos acima. Os códigos Python de sumarização podem ser conferidos em “Analise Metricas.ipynb”.

5.2) Guia de Execução

5.2.1) Instalar Software Necessário

Os seguintes aplicativos são necessários para executar o código de exemplo:

- Windows 10/11 (64-bit)
- Git for Windows
- Python 3.8+ com pip
- Node.js + npm (para jscpd)
- Conta GitHub e Personal Access Token (<https://github.com/settings/tokens>.)
- Jupyter Notebook: Pode ser instalado via comando PIP (pip install notebook). Mais instruções podem ser encontradas em: <https://jupyter.org/install>. Alternativamente o jupyter pode ser carregado ainda via Docker ou instalado junto do pacote Anaconda (<https://www.anaconda.com/download>)

5.2.2) Baixar o repositório

Antes de mais nada é necessário baixar o repositório para um diretório local a gosto.

| |
|--|
| powershell |
| git clone https://github.com/scrapy/scrapy.git |

5.2.3) Extrair os Dados do Radon

Abra PowerShell, cd até o root do repo (ex: C:\caminho\para\scrapy) e execute estes passos em ordem.

```
powershell
# Complexidade Ciclomática
radon cc -a -j . > analysis_data/radon_cc.json
# Dados brutos para análise dos comentários
radon raw -j . > analysis_data/radon_raw.json
# Dados de Manutenibilidade
radon mi -j . > analysis_data/radon_mi.json
```

5.2.4) Coleta de Dados de Duplicações com o jscpd

```
powershell
npx jscpd . --reporters json --output analysis_data/jscpd-report --min-tokens 50 --ignore
"/scrapy/tests/"
```

5.2.5) Estatísticas de Atividade do Repositório (Git API)

Importante notar que o uso da API requer o Token que pode ser obtido em <https://github.com/settings/tokens>.

```
powershell
# Configuração do Token do GitHub
$env:GITHUB_TOKEN = "seu_token_aqui"

# exporta commits com hash, autor, email, data (ISO)
# tem de ser executada dentro da pasta to Scrapy -> cd Scrapy
$lines = git log --all --pretty=format:'%H|%aN|%aE|%aI'
$objects = @()
foreach ($l in $lines -split "`n") {
    $p = $l -split "|"
    $objects += [PSCustomObject]@{hash=$p[0]; author=$p[1]; email=$p[2]; date=$p[3]}
}
$objects | ConvertTo-Json -Depth 4 > ../analysis_data/commits.json

# Configurações do Nome do Repositório, download issues fechados
$owner = "scrapy"
$repo = "scrapy"
$per_page = 100
$page = 1
$all = @()
$headers = @{User-Agent='PS'}
if ($env:GITHUB_TOKEN) { $headers['Authorization'] = "token $env:GITHUB_TOKEN" }
```

```

do {
    $url = "https://api.github.com/repos/$owner/$repo/issues?state=closed&per_page=$per_page&page=$page"
    $resp = Invoke-RestMethod -Uri $url -Headers $headers
    $all += $resp
    $page++
} while ($resp.Count -eq $per_page)

# filtra só issues (remove PRs)
$issues = $all | Where-Object { -not $_.pull_request }
$issues | ConvertTo-Json -Depth 10 > ../analysis_data/issues_closed.json

# Contribuições no último ano
$since = (Get-Date).AddYears(-1).ToString('yyyy-MM-dd')
$short = git shortlog -sne --since="$since"
$objs = @()
foreach ($line in $short -split "`n") {
    if ($line -match '^s*(\d+)\s+(.*)\s+<(.*)>$') {
        $objs += [PSCustomObject]@{commits=[int]$matches[1]; author=$matches[2]; email=$matches[3]}
    }
}
$objs | ConvertTo-Json -Depth 4 > ../analysis_data/contributors_last_year.json

```

5.2.6) Resultados - Complexidade por Função

| Resultados | | | | |
|--|--|----------|------------|------------|
| Funções analisadas: 11119 Complexidade média: 2.82 Distribuição por rank: complexity A 9976 B 852 C 244 D 33 E 9 F 5 Name: count, dtype: int64 | | | | |
| | file | name | complexity | linen o |
| 10190 | venv\Lib\site-packages\pip_vendor\rich\table.py | _render | 49 | 747 |
| 10108 | venv\Lib\site-packages\pip_vendor\rich\style.py | __init__ | 49 | 122 |

| | | | | |
|--------------|--|-----------------|----|-----|
| 8029 | venv\Lib\site-packages\pip_vendor\msgpack\ fal... | _pack | 49 | 676 |
| 6267 | venv\Lib\site-packages\pip_internal\locations ... | get_scheme | 47 | 216 |
| 5874 | venv\Lib\site-packages\pip_internal\cli\autoc. .. | autocomplete | 43 | 17 |
| 6066 | venv\Lib\site-packages\pip_internal\command s\... | run | 39 | 281 |
| 5577 | scrapy\tests\test_settings__init_ .py | test_get | 36 | 247 |
| 10109 | venv\Lib\site-packages\pip_vendor\rich\style. py | __str__ | 35 | 285 |
| 10570 | venv\Lib\site-packages\pip_vendor\urllib3\co n... | urlopen | 35 | 535 |
| 10293 | venv\Lib\site-packages\pip_vendor\rich\trace b... | extract | 34 | 413 |
| 9536 | venv\Lib\site-packages\pip_vendor\rich\cons ol... | __init__ | 32 | 625 |
| 8030 | venv\Lib\site-packages\pip_vendor\msgpack\ fal... | _unpack | 31 | 488 |
| 6866 | venv\Lib\site-packages\pip_internal\req\req_ u... | from_dist | 31 | 427 |
| 7269 | venv\Lib\site-packages\pip_internal\utils\sub. .. | call_subprocess | 31 | 59 |
| 7456 | venv\Lib\site-packages\pip_vendor\cachecon tro... | cache_response | 30 | 324 |
| 9444 | venv\Lib\site-packages\pip_vendor\resolvelib \... | resolve | 30 | 413 |

| | | | | |
|--------------|---|--------------------------|----|-----|
| 10191 | venv\Lib\site-packages\pip_vendor\rich\table.py | _calculate_column_widths | 29 | 523 |
| 8031 | venv\Lib\site-packages\pip_vendor\msgpack\fa... | _read_header | 28 | 392 |
| 10319 | venv\Lib\site-packages\pip_vendor\rich_insp... | _render | 28 | 121 |
| 7688 | venv\Lib\site-packages\pip_vendor\distlib\uti... | parse_requirement | 28 | 144 |

5.2.7) Resultados - NCLOC e Densidade de Comentários

| Resultados |
|---|
| Total LLOC: 119, Total SLOC: 155, Comentários: 28, % Comentários = 18.06% |

5.2.8) Resultados - Índice de Manutenibilidade

| Resultados | | | |
|--|---|----------|------|
| Média geral do Índice de Manutenibilidade: 66.60 | | | |
| Arquivos com pior manutenibilidade: | | | |
| | file | mi | rank |
| 239 | scrapy\tests\test_http_response.py | 0.000000 | C |
| 391 | scrapy\tests\test_settings__init__.py | 0.000000 | C |
| 620 | venv\Lib\site-packages\pip_vendor\pkg_resourc... | 0.000000 | C |
| 234 | scrapy\tests\test_feedexport.py | 0.000000 | C |
| 587 | venv\Lib\site-packages\pip_vendor\distlib\uti... | 0.000000 | C |
| 692 | venv\Lib\site-packages\pip_vendor\rich\consol... | 0.000000 | C |
| 199 | scrapy\tests\test_crawl.py | 0.000000 | C |
| 238 | scrapy\tests\test_http_request.py | 0.000000 | C |
| 200 | scrapy\tests\test_crawler.py | 0.000000 | C |
| 225 | scrapy\tests\test_engine.py | 3.727745 | C |

5.2.9) Resultados - Duplicação

| Resultados |
|--|
| Duplicated lines (sum blocks) = 1342, % = 865.81% (usando total_sloc do radon) |

5.2.10) Resultados - Commits

| Resultados | | |
|---------------------------|---------|-----|
| month | commits | |
| 196 | 2024-10 | 33 |
| 197 | 2024-11 | 37 |
| 198 | 2024-12 | 29 |
| 199 | 2025-01 | 43 |
| 200 | 2025-02 | 27 |
| 201 | 2025-03 | 158 |
| 202 | 2025-04 | 14 |
| 203 | 2025-05 | 60 |
| 204 | 2025-06 | 58 |
| 205 | 2025-07 | 19 |
| 206 | 2025-08 | 12 |
| 207 | 2025-09 | 8 |
| Média commits/mês: 57.875 | | |

5.2.11) Resultados - Tempo Médio para Fechamento de Issues

| Resultados |
|--|
| Estatísticas de tempo para fechar issues (dias): count 2721.000000 mean 243.102835 std 582.007651 min 0.000081 25% 0.451516 50% 7.204641 75% 138.591366 max 4199.540324 Name: days_to_close, dtype: float64 Média dias para fechar: 243.10283521853347 |

5.2.12) Resultados - Contribuidores (últimos 12 meses)

| Results | | | |
|--|-------------|-----------------------|---|
| Número de contribuidores ativos (último ano): 38 | | | |
| | commit s | author | email |
| 0 | 253 | Andrey Rakhmatullin | wrar@wrar.name |
| 1 | 43 | Adrián Chaves | adrian@chaves.io |
| 2 | 7 | Laerte Pereira | 5853172+Laerte@users.noreply.github.com |
| 3 | 5 | Mikhail Korobov | kmike84@gmail.com |
| 4 | 4 | Rotzbua | Rotzbua@users.noreply.github.com |
| 5 | 4 | Thalison Fernandes | thalisondev@gmail.com |
| 6 | 4 | anubhav | protokoul@users.noreply.github.com |
| 7 | 3 | Laerte Pereira | laertefbk@gmail.com |
| 8 | 3 | M Ikram Ullah Khan | 44160462+IkramKhanNiazi@users.noreply.github.com |
| 9 | 3 | Mehraz Hossain Rumman | 59512321+MehrazRumman@users.noreply.github.com |
| 13 | 2 | Rohitkr117 | 145501871+Rohitkr117@users.noreply.github.com |
| 12 | 2 | Rohit Kumar Singh | 145501871+Rohitkr117@users.noreply.github.com |
| 11 | 2 | Keval Sakhiya | 37344767+kevalsakhiya@users.noreply.github.com |
| 10 | 2 | Felipe Benevolo | 77981110+fbenevolo@users.noreply.github.com |
| 26 | 1 | NavpreetSSidhu | 116211004+NavpreetSSidhu@users.noreply.github.com |
| 36 | 1 | mmoriniere | maxime.moriniere@hotmail.fr |
| 35 | 1 | guillermo-bondonno | 95530227+guillermo-bondonno@users.noreply.gith... |

| | | | |
|----------------|---|----------------|---|
| 3 4 | 1 | Vsevolod Breus | vsevolodbreus1@gmail.com |
| 3 3 | 1 | ThunderMind | 46158218+ThunderMind2019@users.noreply.github.com |
| 3 2 | 1 | Swayam Gupta | 78016781+swayam0322@users.noreply.github.com |

5.2.13) Resultados – Comparação com as Metas

| Questão | Métrica | Valor obtido | Meta | Atende? |
|----------------|----------------------------------|---------------------|-------------|----------------|
| 0 | Complexidade média | 2.821567 | ≤ 10 | True |
| 1 | Densidade comentários (%) | 18.064516 | $\geq 20\%$ | False |
| 2 | Duplicação (%) | 865.806452 | $\leq 5\%$ | False |
| 3 | Índice Manutenibilidade (média) | 66.603809 | ≥ 70 | False |
| 4 | Commits / mês (média) | 57.875000 | ≥ 10 | True |
| 5 | Tempo médio fechar issues (dias) | 243.102835 | ≤ 14 | False |
| 6 | Contribuidores (últ. 1 ano) | 38.000000 | ≥ 10 | True |

Por simplificação para o relatório final os dados aqui foram arredondados para a unidade.