Etapas e Resultados

Influência de boas práticas de

- (i) Arquitetura de software
- (ii) Padrões e Estilos Arquiteturais
- (iii) Padrões de Projeto na Produtividade

Membros da Equipe:

- 1. Letícia Cavalcanti
- 2. Maria Diniz

Disciplina:

COMP0439 - Engenharia de Software II (2024.2 - T03)

1. Apresentação do Projeto Escolhido

1.1. Nome do Projeto e URL de Acesso

Projeto: Flutter

URL: https://github.com/flutter/flutter

1.2. Finalidade do Projeto

Flutter é um framework de código aberto desenvolvido pelo Google para a criação de aplicativos que funcionam em diversas plataformas com um único código. Com ele, é possível desenvolver aplicativos para Android, iOS, Web, Desktop e até sistemas embarcados. O Flutter se destaca por sua estrutura baseada em widgets, onde cada elemento da interface (como botões, textos e imagens) faz parte de uma árvore de componentes (widget tree). Isso torna a criação de telas mais flexível e organizada.

Outro ponto forte do Flutter é o uso do mecanismo gráfico Skia, que permite a renderização rápida e suave dos elementos visuais, garantindo alto desempenho e animações fluidas. A linguagem utilizada no Flutter é o Dart, criada pelo próprio Google. Ela foi projetada para ser simples, eficiente e otimizada para interfaces gráficas, facilitando o desenvolvimento de aplicativos com uma aparência moderna e responsiva.

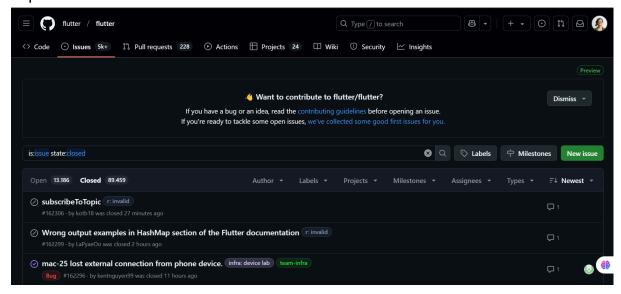
1.3. Características Relevantes do Projeto

- Grande comunidade de desenvolvedores: Possui mais de 1.481 contribuidores no GitHub.
- Elevado número de issues: O projeto apresenta mais de 89.503
- issues fechadas
- Quantidade expressiva de pull requests: Conta com mais de 58.125 pull requests fechados registrados.
- Arquitetura modular e flexível: Utiliza conceitos de boas práticas de arquitetura de software, facilitando a extensibilidade e manutenção.
- Uso de padrões e estilos arquiteturais: O Flutter adota padrões como MVU (Model-View-Update) e arquiteturas como Redux e Bloc para gerenciamento de estado.
- Adoção de padrões de projeto: Implementa padrões como Singleton, Factory e Observer em diversas partes do código.

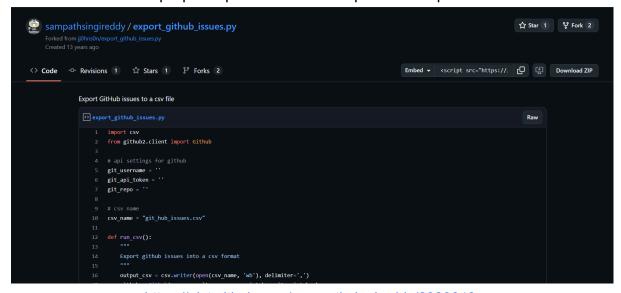
A partir dessas observações, daremos sequência às próximas etapas do estudo, realizando a coleta, classificação e análise das issues do Flutter relacionadas a boas práticas de arquitetura de software, padrões e estilos arquiteturais e padrões de projeto.

2. Etapas

2.1. Optamos pelo Flutter após uma busca por projetos de código aberto que atendessem aos critérios exigidos na atividade ou estivessem próximos desses requisitos.



2.2. Para facilitar a extração das 300 issues fechadas (closed) do repositório do Flutter no GitHub, buscamos um código aberto que automatizasse esse processo. Encontramos um script que exportava as issues para um arquivo CSV.



https://gist.github.com/sampathsingireddy/2993646

2.3. Modificamos o código original para que, em vez de exportar os dados para um arquivo CSV, as issues fossem diretamente inseridas em um banco de dados PostgreSQL. Além disso, adicionamos os atributos extras exigidos na atividade,

como data de abertura, data de conclusão, tempo de resolução, prioridade e milestone.

Utilizamos o ChatPT, através do modelo 4o-mini, ajustar para baixar as issues direto no Postgres com o seguinte prompt principal:

"\\${código} + como ao inves de converter para csv eu jogar direto no banco postgres?"

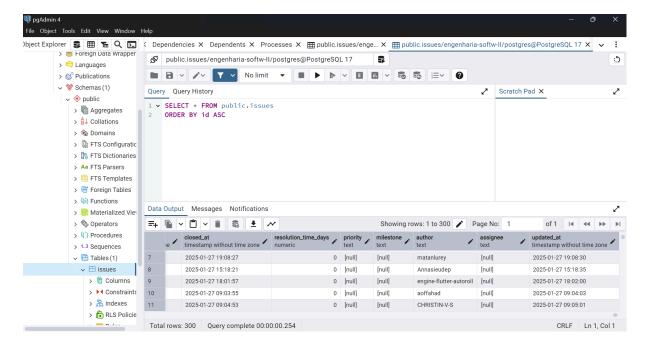
E o resultado foi obtido de acordo com nossas máquinas e banco:

```
import requests
import psycopg2
from datetime import datetime
   conn = psycopg2.connect(
       password="1234",
       host="localhost",
       port="5432"
except Exception as e:
   print(f"Erro ao conectar ao banco de dados: {e}")
   exit()
   cursor.execute("""
```

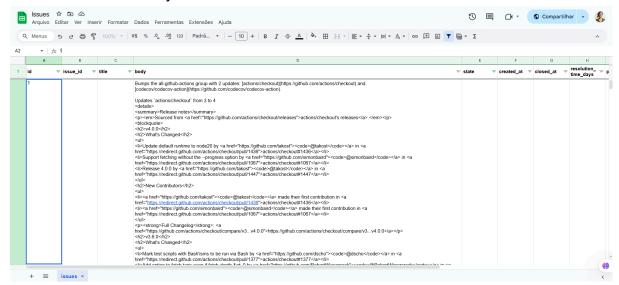
```
conn.commit()
   print("Tabela 'issues' criada/verificada com sucesso.")
except Exception as e:
   print(f"Erro ao criar/verificar a tabela: {e}")
   conn.close()
   exit()
url = "https://api.github.com/repos/flutter/flutter/issues"
headers = {"Authorization": "TOKEN: DADO PROTEGIDO"}  # Substitua pelo seu
per page = 100 # Máximo permitido por página
total issues = 300  # Quantidade total desejada
issues closed = 0 # Contador de issues já buscadas
page = 1  # Página inicial
try:
   while issues closed < total issues:</pre>
           per page = 1
           per_page = per_page
       paginated url = f"{url}?state=closed&per page={per page}&page={page}"
       response = requests.get(paginated url, headers=headers)
       response.raise for status()
        issues = response.json()
        if not issues: # Se não houver mais issues, sair do loop
       print(f"Página {page}: {len(issues)} issues fechadas encontradas.")
                issue id = issue.get('id')
                title = issue.get('title', 'Sem título')
                body = issue.get('body') if 'body' in issue else ''
                state = issue.get('state', 'unknown')
```

```
created at = issue.get('created at')
                closed at = issue.get('closed at')
                updated at = issue.get('updated at')
                if created at and closed at:
                    resolution time days = (
                        datetime.fromisoformat(created at[:-1])
                    ).days
                    resolution time days = None
                labels = issue.get('labels', [])
                for label in labels:
                    label name = label.get('name', '').lower()
                    if "high" in label name:
                        priority = "High"
                    elif "medium" in label name:
                        priority = "Medium"
                    elif "low" in label name:
                milestone data = issue.get('milestone')
                milestone = milestone data.get('title') if milestone data
else None
                user data = issue.get('user')
                author = user data.get('login') if user data else None
                assignee data = issue.get('assignee')
                assignee = assignee data.get('login') if assignee data else
None
                cursor.execute("""
created at, closed at, resolution time days, priority, milestone, author,
assignee, updated at)
                """, (issue id, title, body, state, created at, closed at,
resolution time days, priority, milestone, author, assignee, updated at))
```

2.4. Após realizar ajustes no código, executamos uma série de testes para garantir que a exportação estava funcionando corretamente. Após validações bem-sucedidas, conseguimos estabelecer a conexão com o banco de dados e inserir os registros. Confirmamos que os dados estavam devidamente armazenados no banco de dados e realizamos a exportação para um arquivo CSV para facilitar a análise posterior.



- 2.5. Importamos o arquivo CSV para o Google Planilhas e adicionamos a coluna "tema relacionado". Em seguida, iniciamos a classificação das issues de acordo com os três temas da atividade:
 - Arquitetura de Software
 - Padrões e Estilos Arquiteturais0
 - Padrões de Projeto



2.6. Ao analisar o tempo necessário para classificar manualmente todas as 300 issues, percebemos que esse processo seria demorado e pouco eficiente. Para otimizar a tarefa, decidimos desenvolver uma ferramenta baseada em Inteligência Artificial que pudesse ler e classificar automaticamente cada issue com base nas colunas "title" e "body" do banco de dados.

Para isso, utilizamos o ChatGPT, através do modelo 3.5 Turbo, para auxiliar na geração do código. Fizemos a seguinte solicitação:

"Quer gerar um código usando OpenAl e LangChain para ler os itens das colunas 'title' e 'body' no banco de dados e me retornar a classificação delas numa coluna ou num arquivo TXT?"

Após validar a abordagem inicial, fizemos um novo pedido para armazenar diretamente os resultados no banco de dados:

"Tem como salvar as classificações direto no banco numa coluna nova chamada 'tema relacionado'?"

Com isso, conseguimos automatizar o processo, reduzindo significativamente o esforço manual necessário para a categorização das issues.

```
import psycopg2
from langchain.chat models import ChatOpenAI
from langchain.schema import HumanMessage
load_dotenv()
DB NAME = os.getenv("DB NAME")
   USER = os.getenv("DB USER")
   PASSWORD = os.getenv("DB PASSWORD")
   HOST = os.getenv("DB_HOST")
DB_PORT = os.getenv("DB_PORT")
OPENAI API KEY = os.getenv("OPENAI API KEY")
        conn = psycopg2.connect(
           dbname=DB NAME,
            user=DB USER,
            password=DB_PASSWORD,
        print(f"Erro ao conectar ao banco de dados: {e}")
        exit()
def classify texts with gpt35(body):
        chat = ChatOpenAI(model="gpt-3.5-turbo", temperature=0.3,
openai api key=OPENAI API KEY)
        message = HumanMessage(
            content=(
        response = chat([message])
        classification = response.content.strip()
```

```
classification = "Padrões e Estilos Arquiteturais"
            classification = "Padrões de Projeto"
alteração será feita.")
       return classification
       print(f"Erro ao classificar o texto: {e}")
def main():
   conn = connect to db()
                    SELECT 1
                END IF;
        conn.commit()
       print("Coluna 'tema relacionado' verificada/adicionada com sucesso.")
       cursor.execute("SELECT id, body FROM issues WHERE tema relacionado IS
       rows = cursor.fetchall()
       print("Classificando os textos...")
            issue id, body = row
            classification = classify texts with gpt35(body)
                    UPDATE issues
```

```
print(f"Issue {issue_id} classificada como:
{classification}")
    else:
        print(f"Issue {issue_id} não classificada.")

    conn.commit()
    print("Classificações salvas no banco de dados com sucesso.")

except Exception as e:
    print(f"Erro ao processar os dados: {e}")

finally:
    cursor.close()
    conn.close()
    print("Conexão com o banco encerrada.")

if __name__ == "__main__":
    main()
```

- 2.7. Para garantir a precisão do modelo comparamos as classificações geradas pela IA com nossa interpretação humana. Esse processo permitiu avaliar a assertividade da ferramenta e ajustar eventuais inconsistências.
- 2.8. Para facilitar a análise de métricas, modificamos uma coluna do banco de dados para armazenar o tempo de resolução de cada issue em horas. Dessa forma, conseguimos obter um indicador mais preciso do tempo médio necessário para a conclusão das tarefas.

Utilizamos o seguinte prompt no ChatGPT, através do modelo 4-o mini, para gerar o código necessário:

"Gere um código Python que lê as colunas 'created_at' e 'closed_at' e devolve o tempo de fechamento em horas. Considere que meus arquivos vão estar no env."

A seguir, implementamos o código responsável pela conversão do tempo de resolução das issues em horas:

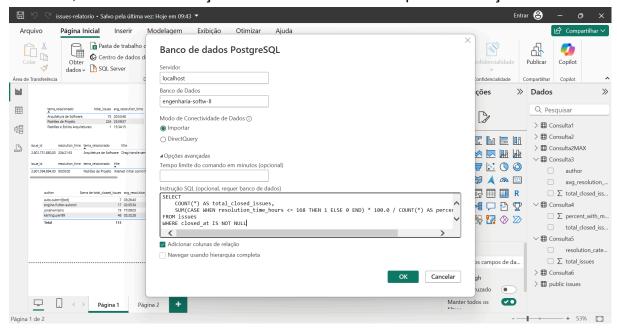
```
import os
import psycopg2
from dotenv import load_dotenv
from datetime import datetime

# Carregar variáveis de ambiente do arquivo .env
load_dotenv()

# Configurações do banco de dados
DB_NAME = os.getenv("DB_NAME")
DB_USER = os.getenv("DB_USER")
DB_PASSWORD = os.getenv("DB_PASSWORD")
DB_HOST = os.getenv("DB_HOST")
DB_PORT = os.getenv("DB_PORT")
```

```
def connect to db():
       conn = psycopg2.connect(
           dbname=DB NAME,
           user=DB USER,
            password=DB PASSWORD,
           port=DB PORT,
       print(f"Erro ao conectar ao banco de dados: {e}")
       exit()
def calcular_tempo_fechamento():
   cursor = conn.cursor()
       issues = cursor.fetchall()
       for issue id, created at, closed at in issues:
            created at dt = datetime.fromisoformat(str(created at))
            closed at dt = datetime.fromisoformat(str(closed at))
            resolution time hours = (closed at dt -
created at dt).total seconds() / 3600
               UPDATE issues
            """, (resolution time hours, issue id))
        conn.commit()
       print(f"{len(issues)} issues atualizadas com tempo de fechamento em
horas.")
       print(f"Erro ao calcular tempo de fechamento: {e}")
       cursor.close()
       conn.close()
   calcular tempo fechamento()
```

2.9. Para analisar os dados extraídos, realizamos consultas SQL diretamente no Power BI, extraindo informações relevantes sobre o tempo de resolução das issues.



Consultando as seguintes informações:

2. Tempo máximo e mínimo de resolução de issues

```
-- Issue mais demorada

WITH max_issue AS (

SELECT

issue_id,

tema_relacionado,

title,

TO_CHAR(

INTERVAL '1 second' * ROUND(resolution_time_hours * 3600),

'HH24:MI:SS'

) AS resolution_time
```

```
FROM issues
        WHERE closed at IS NOT NULL
        ORDER BY resolution_time_hours DESC
        LIMIT 1
      )
      SELECT * FROM max issue;
-- Issue mais rápida
      WITH min issue AS (
        SELECT
          issue_id,
          tema_relacionado,
          title,
          TO CHAR(
            INTERVAL '1 second' * ROUND(resolution_time_hours * 3600),
            'HH24:MI:SS'
          ) AS resolution time
        FROM issues
        WHERE closed at IS NOT NULL
        ORDER BY resolution time hours ASC
        LIMIT 1
SELECT * FROM min_issue;
3. Top 5 autores que mais fecharam issues e tempo médio de resolução por
autor
      SELECT
        author,
        COUNT(*) AS total_closed_issues,
        TO CHAR(
          INTERVAL '1 second' * ROUND(AVG(resolution time hours) *
      3600),
          'HH24:MI:SS'
        ) AS avg resolution time
      FROM issues
      WHERE closed at IS NOT NULL AND resolution time hours IS NOT
      NULL
      GROUP BY author
      HAVING COUNT(*) > 5
      ORDER BY total_closed_issues DESC
      LIMIT 5;
4. Impacto das issues com base no tempo de fechamento
      SELECT
            issue id,
```

```
title AS titulo,
milestone AS marco,
CASE
WHEN resolution_time_hours <= 24 THEN 'Curto Prazo'
WHEN resolution_time_hours BETWEEN 24 AND 168 THEN
'Médio Prazo'
ELSE 'Longo Prazo'
END AS impacto_tempo
FROM issues
WHERE state = 'closed';
```

5. Issues fechadas dentro de 24 horas vs. acima disso

Para entender quantas issues são resolvidas rapidamente:

SELECT

CASE WHEN resolution_time_hours <= 24 THEN 'Menos de 1 dia' ELSE 'Mais de 1 dia' END AS resolution_category, COUNT(*) AS total_issues

FROM issues

WHERE closed_at IS NOT NULL GROUP BY resolution_category;

6. Percentual de issues fechadas dentro de um prazo razoável (exemplo: 7 dias)

SELECT
COUNT(*) AS total_closed_issues,
SUM(CASE WHEN resolution_time_hours <= 168 THEN 1 ELSE 0
END) * 100.0 / COUNT(*) AS percent_within_7_days
FROM issues
WHERE closed_at IS NOT NULL;

2.10. Desenvolvemos um dashboard no PowerBI, utilizando gráficos gerados a partir de consultas SQL executadas diretamente na plataforma.

3. Resultados

Resultados: Análise de 300 issues do Flutter



3.1. Nº de issues e tempo médio de resolução por tema

Análise:

- O tempo médio de resolução das issues mostra uma variação significativa entre os temas.
- Padrões de Projeto tem o maior tempo médio, o que pode indicar maior complexidade ou prioridade menor.
- Arquitetura de Software tem um tempo médio próximo a 21 horas, possivelmente porque envolve decisões técnicas que demandam mais avaliação.

3.2. Tempo máximo e mínimo de resolução de issues

Análise:

- Menor tempo: 5 minutos e 24 segundos
 - Esse tempo sugere que as issues mais rápidas de resolver são aquelas de baixa complexidade, como pequenos ajustes ou correções no código, revisões de documentação ou alterações visuais menores.
 - Pode também ser um reflexo de situações onde a solução já estava preparada antes da abertura da issue ou onde ferramentas automatizadas facilitaram o trabalho.
- Maior tempo: 204 horas (8,5 dias)
 - Esse tempo reforça a ideia de que algumas issues exigem análise aprofundada, múltiplas etapas de desenvolvimento e testes rigorosos.

- Problemas críticos ou de alta complexidade, como reestruturações de arquitetura, otimizações de desempenho ou decisões estratégicas, frequentemente demandam esse nível de esforço.
- Outro fator pode ser a dependência de terceiros ou o envolvimento de diversas equipes no processo.

3.3. Top 5 autores que mais fecharam issues e tempo médio de resolução por autor

Análise:

- A presença de um bot do Flutter, o "engine-flutter-autoroll", como o maior responsável pelo fechamento de issues sugere automação no processo de fechamento. Por este motivo, foi um dos selecionados no ranking.
 - O engine-flutter-autoroll é um bot automatizado do GitHub responsável por sincronizar atualizações entre o repositório da engine do Flutter (flutter/engine) e o repositório principal do framework (flutter/flutter).
 - Principais funções:
 - Monitoramento de atualizações: Identifica novas mudanças na engine do Flutter.
 - Criação de Pull Requests automáticos: Gera PRs para integrar as atualizações ao repositório principal do Flutter.
 - Execução de testes: Dispara testes automatizados para validar a estabilidade das atualizações.
 - Mesclagem automatizada: Caso os testes sejam bem-sucedidos, o bot pode mesclar as alterações automaticamente ou aguardar aprovação manual.
 - o Importância do engine-flutter-autoroll:
 - A utilização desse bot garante que a engine do Flutter esteja sempre sincronizada com as últimas mudanças, reduzindo a necessidade de intervenções manuais e minimizando riscos de incompatibilidades. Além disso, o processo automatizado contribui para a estabilidade e confiabilidade do framework, assegurando que apenas atualizações testadas sejam incorporadas.
- Autores humanos têm tempos médios muito variados, o que pode indicar que alguns lidam com issues mais complexas do que outros.
- jonahwilliams apresenta o maior tempo médio de resolução, possivelmente porque trata de issues mais complexas.
- 3.4. Impacto das issues com base no tempo de fechamento

Análise:

- O tempo foi classificado como curto para questões resolvidas em até 24 horas, médio para aquelas resolvidas entre 24 horas e 7 dias, e longo para questões que levaram mais de 7 dias para serem resolvidas.
- A distribuição mostra uma eficiência na resolução da maioria das issues, com quase 76% resolvidas rapidamente.
- Poucas issues permaneceram abertas por um longo período, o que indica uma boa capacidade de resposta e priorização das issues mais críticas.

3.5. Issues fechadas dentro de 24 horas vs. acima disso

Análise:

- Apesar da maioria das issues levarem mais de 24h para serem resolvidas, cerca de 1/4 das issues são resolvidas em menos de um dia, o que é um indicador positivo de eficiência.
- Poderia ser interessante analisar quais tipos de issues são resolvidas rapidamente e quais demoram mais para identificar padrões e possíveis otimizações.
- 3.6. Percentual de issues fechadas dentro de um prazo razoável (7 dias)

Análise:

- Esse é um indicador altamente positivo. Mostra que a grande maioria das issues não fica aberta por longos períodos, sugerindo que o time tem uma política eficiente de fechamento.
- O 1,33% restante pode representar casos mais críticos ou menos prioritários.

4. Conclusões Gerais

Pontos Positivos

- Alto índice de resolução rápida: 75% das issues são resolvidas rapidamente.
- Baixo percentual de issues de longo prazo: Apenas 1,33% permanecem abertas por um tempo prolongado.
- Automação eficiente: O bot desempenha um papel relevante no fechamento de issues.
- Ótimo prazo médio de resolução: 98,67% das issues fechadas em menos de 7 dias.

Pontos de Melhoria

- Grande variação no tempo médio de resolução por autor: Alguns autores têm tempos de resolução significativamente maiores que outros.
- Padrões de Projeto demora mais para ser resolvido: Poderia ser interessante investigar se esse tempo maior afeta a produtividade.

Sugestões

- Investigar os tipos de issues que demoram mais tempo para serem resolvidas e avaliar formas de otimização.
- Analisar os autores que têm tempos de resolução mais altos e verificar se há gargalos no processo.

Análise das Issues no Processo de Arquitetura de Software

- Issues de Arquitetura de Software:
 - As questões relacionadas à arquitetura de software apresentam um tempo de resolução moderado, indicando que elas frequentemente exigem uma análise técnica mais aprofundada e criteriosa antes de serem implementadas. Isso sugere que, embora sejam importantes, as decisões podem demandar mais tempo devido à necessidade de um planejamento detalhado, que leve em consideração a escalabilidade, a manutenção e a eficiência do sistema a longo prazo.
- Issues de Padrões e Estilos Arquiteturais:
 As questões relacionadas a padrões e estilos
 - As questões relacionadas a padrões e estilos arquiteturais seguem um ritmo de resolução similar ao das questões de arquitetura de software, o que indica que as decisões estruturais dentro do projeto são feitas de maneira eficiente e organizada. A adoção de padrões bem estabelecidos contribui para uma solução mais coesa e alinhada com as melhores práticas do setor, facilitando a comunicação e a implementação do projeto.
- Issues de Padrões de Projeto:
 - As questões ligadas a padrões de projeto tendem a ser mais complexas e demoram mais para serem resolvidas. Isso pode ser um reflexo de uma maior complexidade técnica ou de uma priorização mais baixa no fluxo de trabalho. Quando essas questões não são resolvidas rapidamente, podem afetar diretamente a produtividade do time de desenvolvimento, retardando a evolução do projeto.

Resumo Final

O relatório sugere que o time do Flutter tem um bom desempenho na resolução de issues, com alta eficiência e baixa taxa de atraso. Há oportunidades de melhoria na variação do tempo de resolução entre os autores, mas no geral, os dados indicam um processo bem otimizado.