

Trabalho Interdisciplinar: Pesquisa em Engenharia de Software

Professores:

- Felipe Augusto Lima Reis
- José Laerte Pires Xavier Junior

Alunos:

- Guilherme Gabriel S. Pereira
- Henrique P. F. Monteiro
- Lucas Ângelo O. M. Rocha
- Victor Boaventura G. Campos
- Vinícius Marini C. e Oliveira

Estudo de caracterização de aplicações desktop de repositórios do Github por meio de Inteligência Artificial



Pesquisa sobre aplicações desktop de
código aberto em repositórios do Github

Sumário

Apresentação

Glossário

1. Hipótese
2. Partes Interessadas
3. Objetivo
 - 3.1. Objetivo baseado na hipótese
 - 3.2. GQM (Goal, Question, Metric)
4. Perguntas da Pesquisa
 - 4.1. Pergunta 1
 - 4.2. Pergunta 2
 - 4.3. Pergunta 3
5. Metodologia
6. Textos relacionados
7. Proof of concept

Referência

Bibliografia

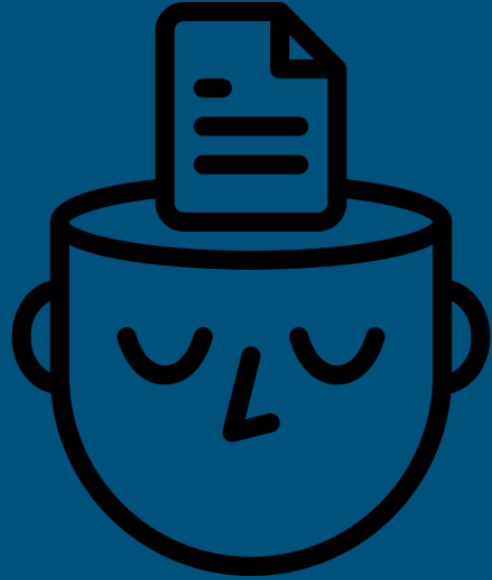


Glossário

1. **Domínio:** refere-se ao propósito do software, a qual ramo ele se destina.
Exemplo: domínio de aplicação de antivírus, editores de vídeos, etc.;
2. **Aplicações desktop:** programa que precisa estar instalado em sua máquina, independente de qual seja sua função.
3. **Repositórios desktop:** quaisquer repositórios que possuam dependências de aplicações desktop na linguagens Java, C# e JavaScript.

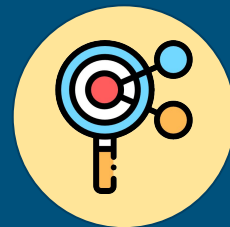


1 - Hipótese



1 - Hipótese

- Ao longo dos anos, os navegadores evoluíram e passaram a ter mais funcionalidades e suportar aplicações mais elaboradas. Nesse contexto, a evolução da internet e a praticidade de acessar os sistemas via web diminuiu a necessidade de se ter uma aplicação desktop. Contudo, alguns domínios de aplicações necessitam de ser desktop, como antivírus;
- Nesse sentido, foi pensado que **ainda existem domínios populares no contexto de aplicações desktop.**



2 - Partes interessadas



2 - Partes interessadas

1. Desenvolvedores analisando tecnologias para projetos;
2. Analistas, clientes e gestores de novos projetos
 - a. que ainda não decidiram qual é o melhor para o seu domínio, web ou desktop.



3 - Objetivo



3.1 - Objetivo baseado na hipótese

- O objetivo principal do projeto é testar a hipótese inicial e registrar o resultado. Dessa forma, por meio de pesquisa e análise de dados, **busca-se descobrir domínios que ainda são populares para aplicações desktop.**



3.2 - GQM (Goal, Question, Metric)

Análise de repositórios de aplicações desktop

com o objetivo de caracterizar aplicações desktop e seus domínios

com relação a quantidade de aplicações e suas respectivas métricas

do ponto de vista de analistas, gerentes e clientes de novos projetos

no contexto dos repositórios do Github que possuem dependências de aplicações desktop das linguagens C# e JavaScript.



4 - Perguntas e métricas



4.1 - Pergunta 1

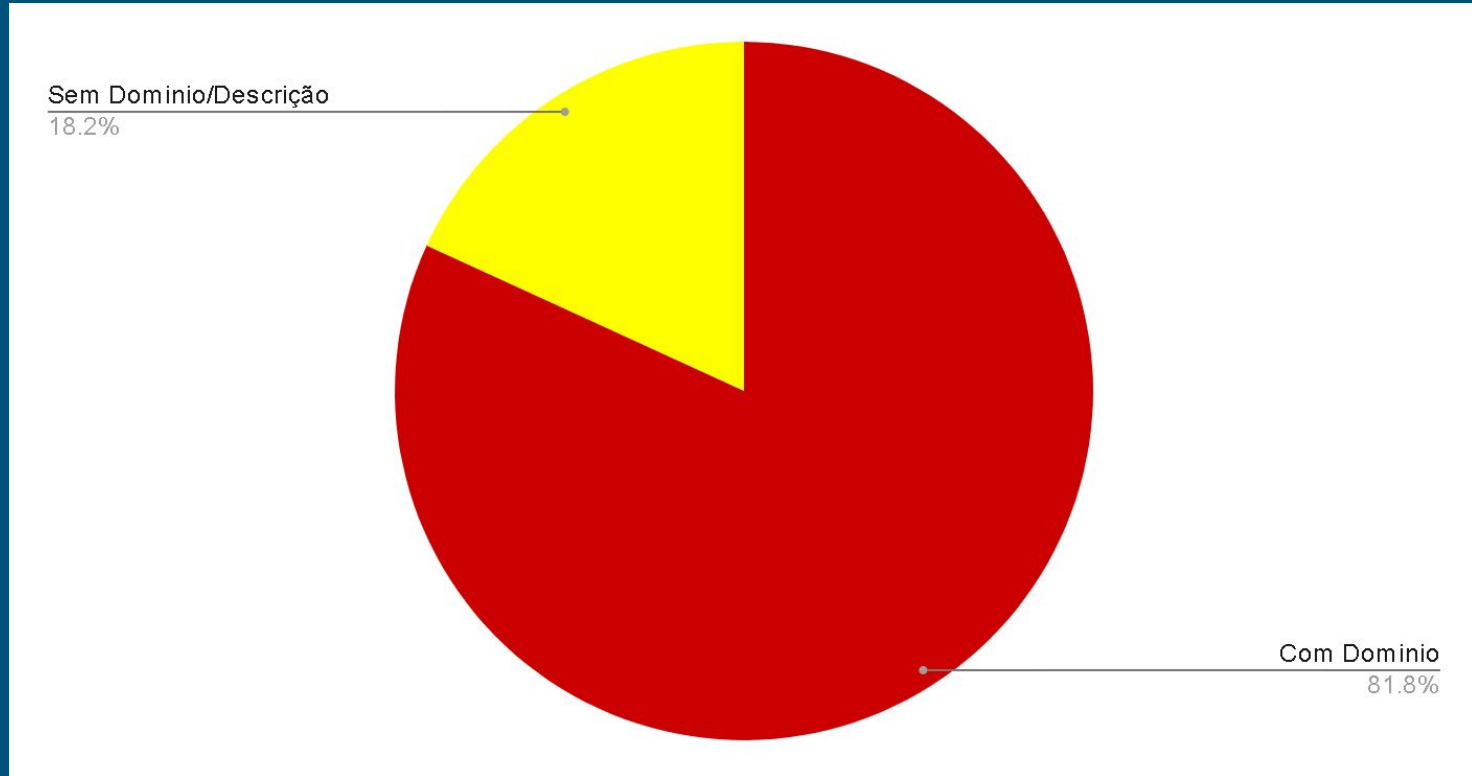
Para as aplicações desktop que ainda são mantidas, qual o domínio que elas se encontram atualmente?

Métrica 1: Proporção de repositórios que possuem descrições e domínios versus que não possuem descrições ou domínios;

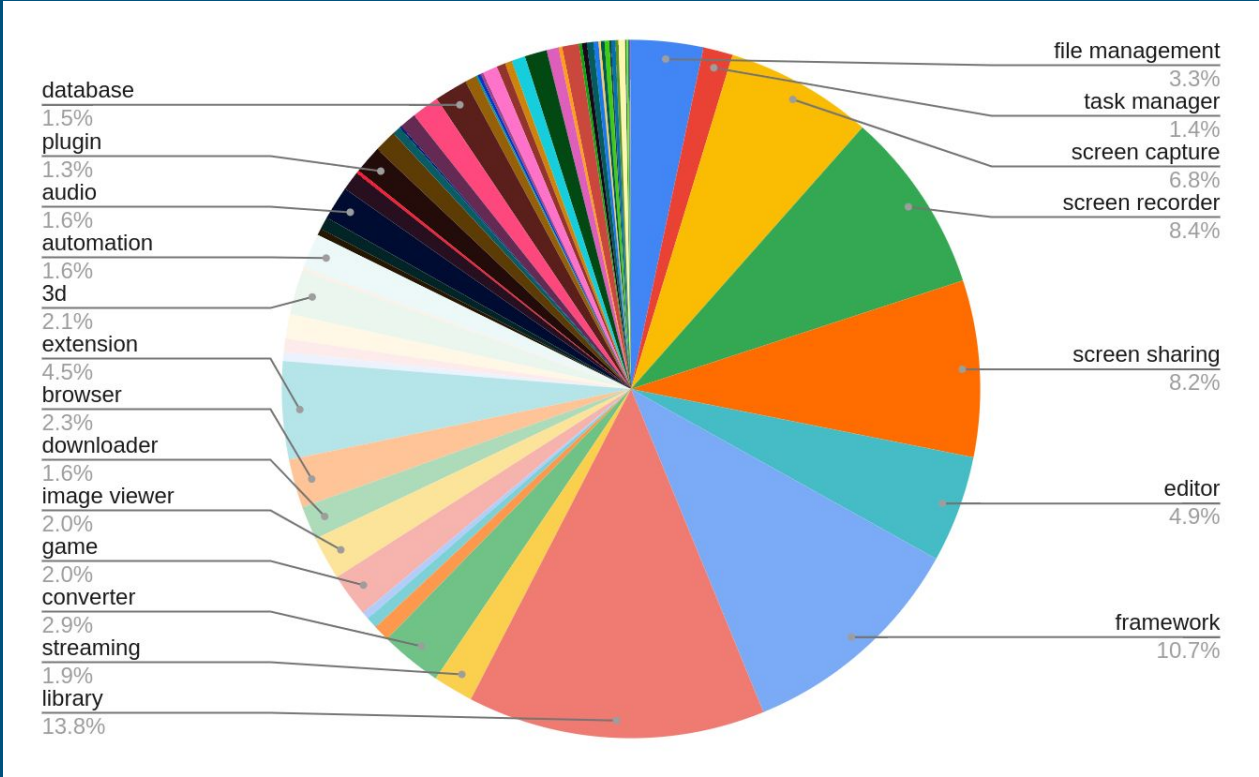
Métrica 2: Percentual da quantidade de repositórios desktop para cada domínio.



4.1.1 - Pergunta 1, métrica 1



4.1.2 - Pergunta 1, métrica 2



4.2 - Pergunta 2

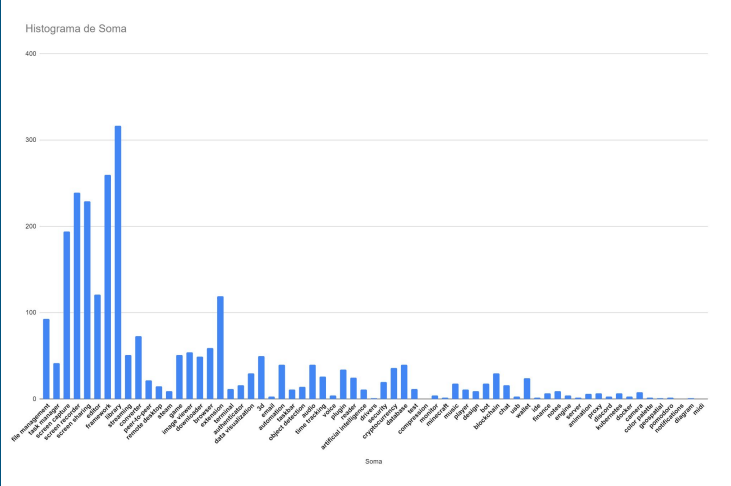
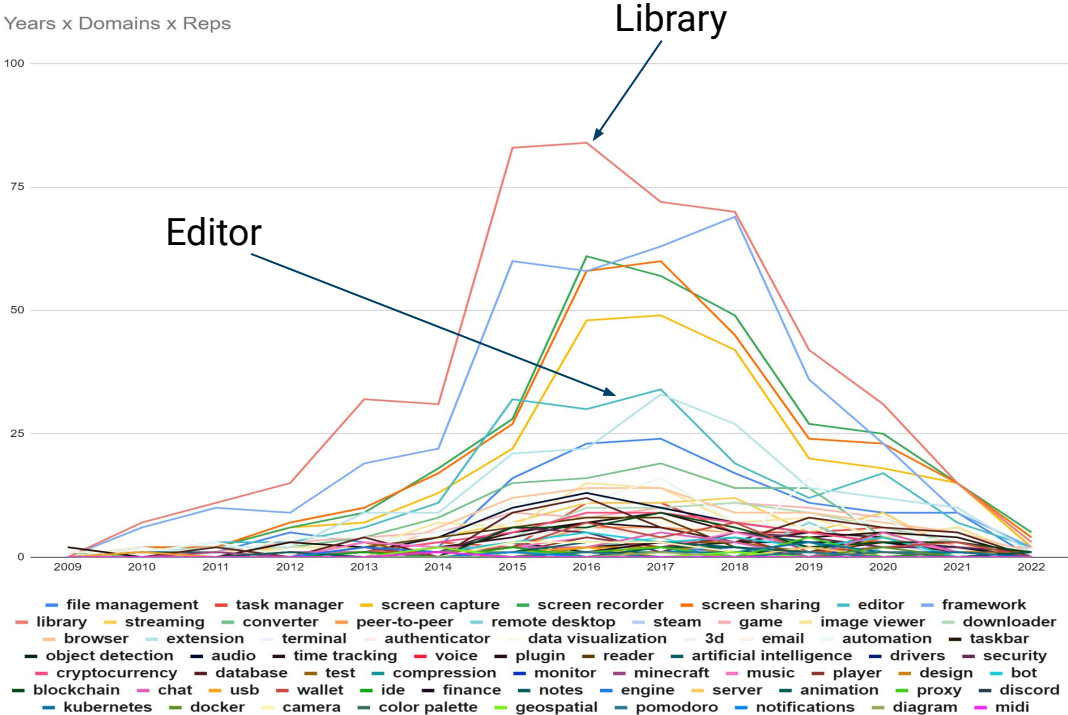
A quantidade de aplicações desktop vem diminuindo ao longo da última década?

Métrica 1: Média de repositórios com dependências de aplicações desktop criados por ano para cada domínio;

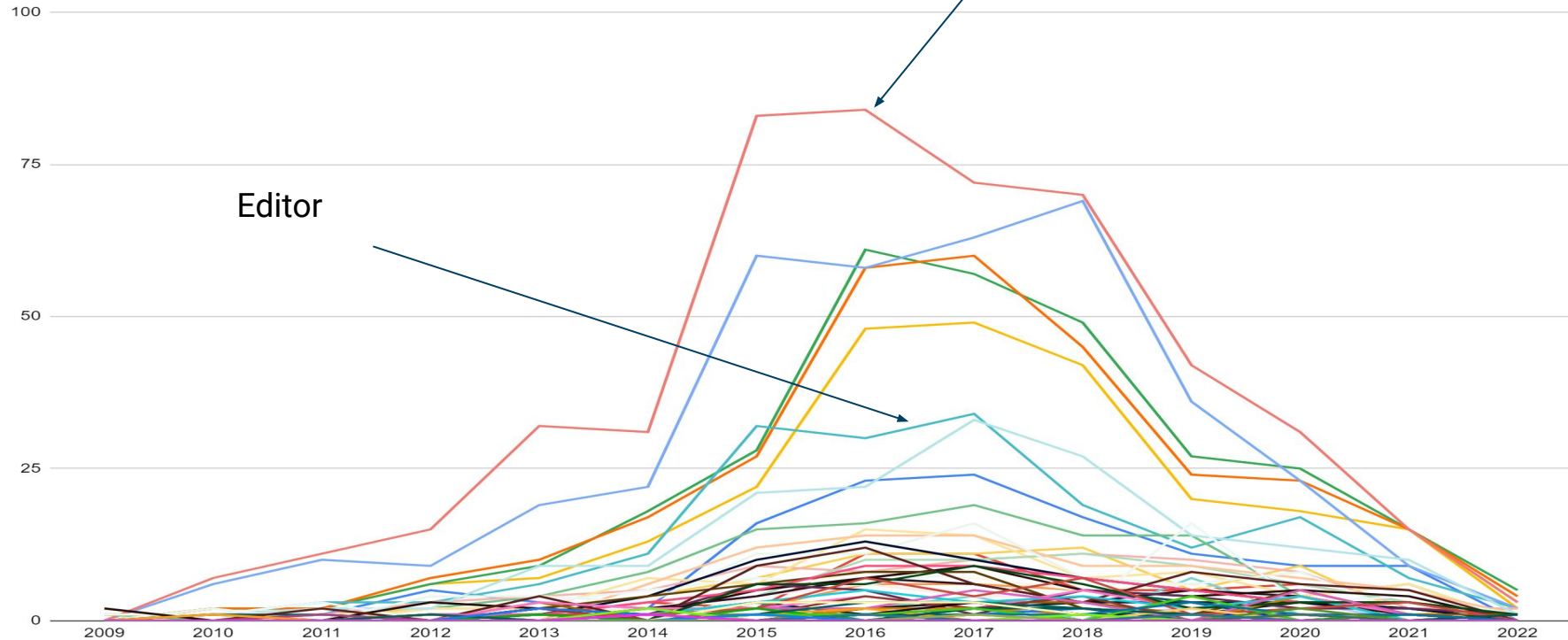
Métrica 2: Média de repositórios com dependências de aplicações desktop criados por ano.



4.2.1 - Pergunta 2, métrica 1

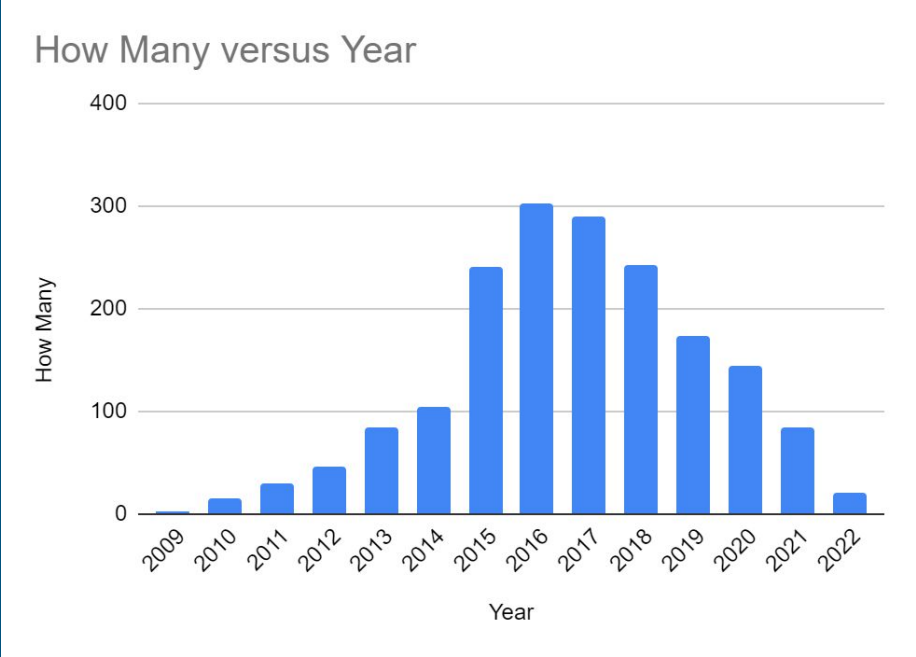
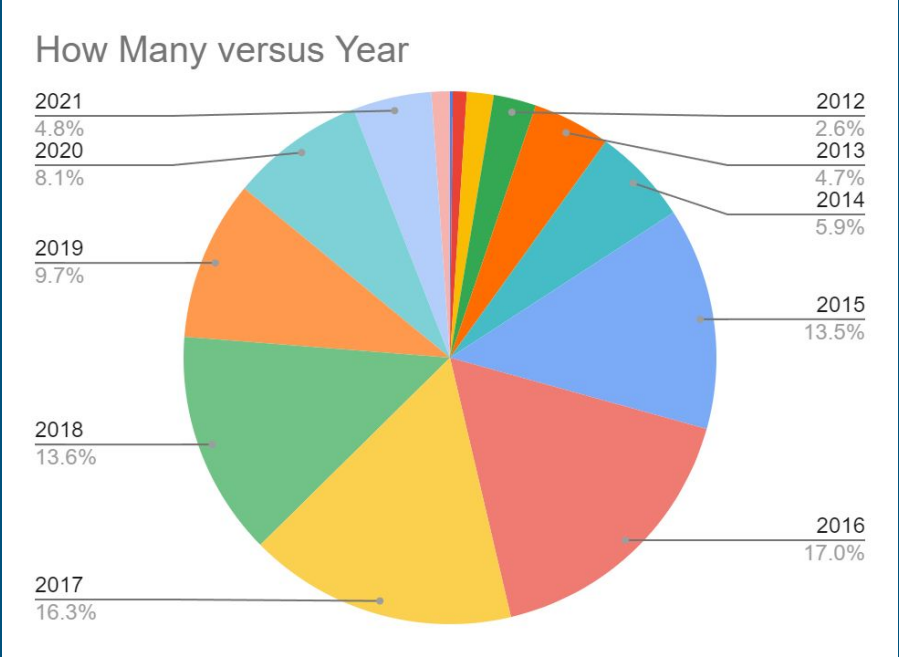


Years x Domains x Reps

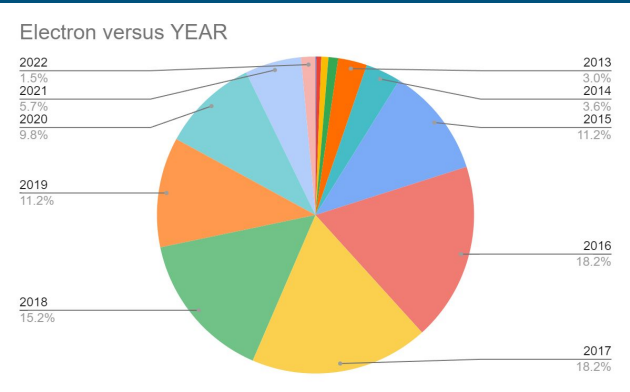
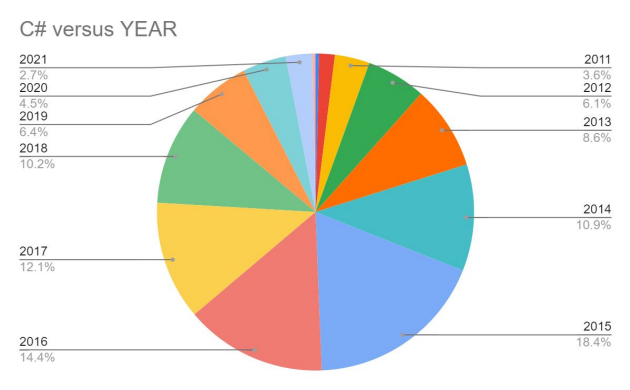
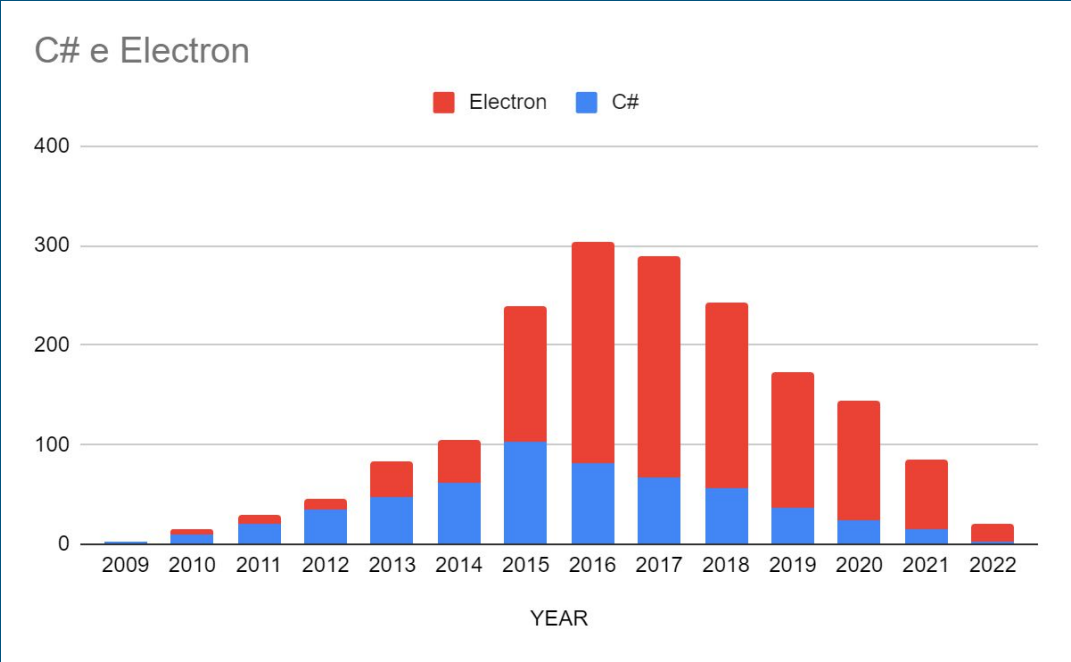


file management task manager screen capture screen recorder screen sharing editor framework
library streaming converter peer-to-peer remote desktop steam game image viewer downloader
browser extension terminal authenticator data visualization 3d email automation taskbar
object detection audio time tracking voice plugin reader artificial intelligence drivers security
cryptocurrency database test compression monitor minecraft music player design bot
blockchain chat usb wallet ide finance notes engine server animation proxy discord
kubernetes docker camera color palette geospatial pomodoro notifications diagram midi

4.2.1 - Pergunta 2, métrica 2



4.2.1 - Pergunta 2, métrica 2



4.3 - Pergunta 3

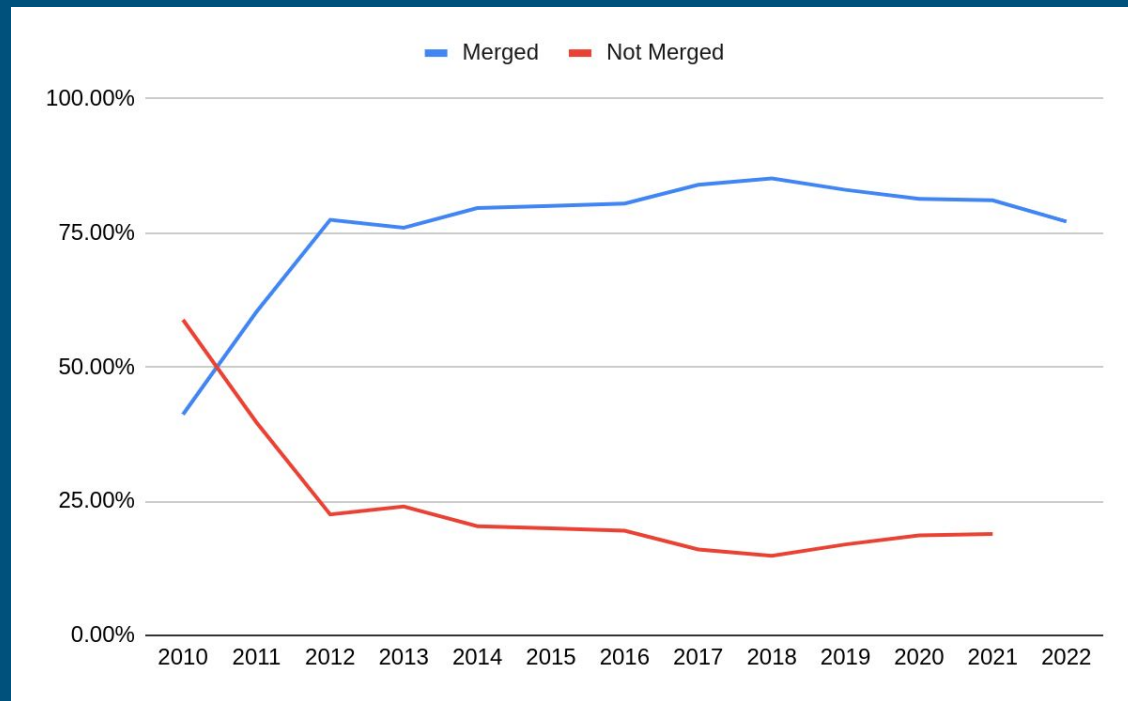
Aplicações desktop tem engajamento da comunidade?

Métrica 1: Percentual de pull requests *merged* em relação aos não *merged* dos repositórios desktop por ano;

Métrica 2: Percentual de issues fechadas em relação a não fechadas em repositórios desktop por ano.

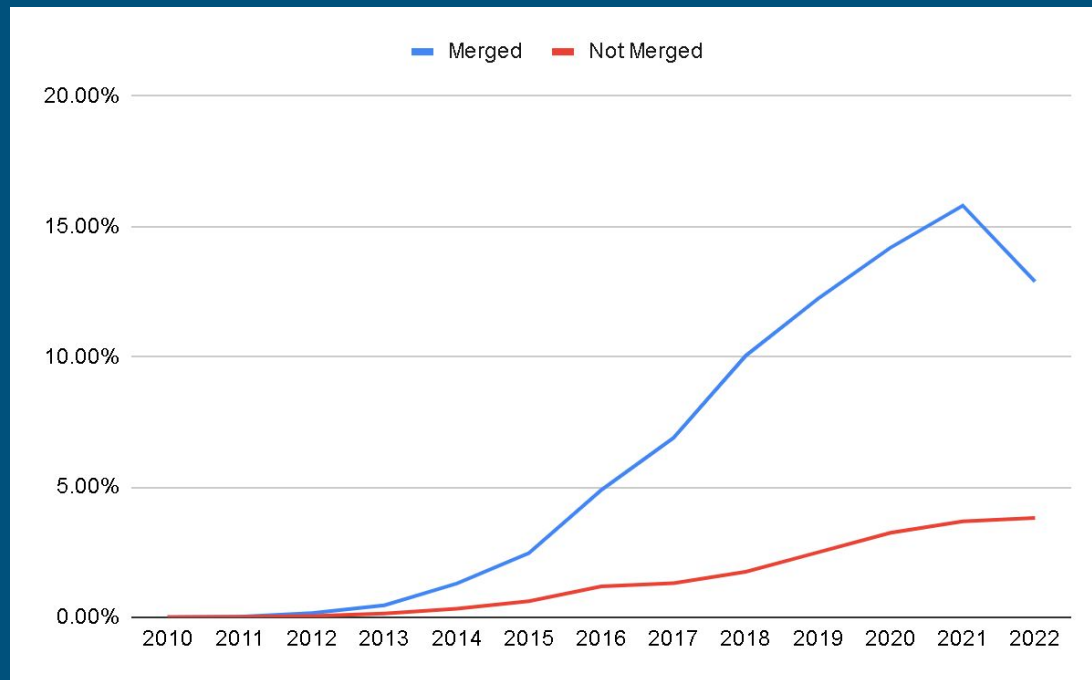


4.3.1 - Pergunta 3, métrica 1



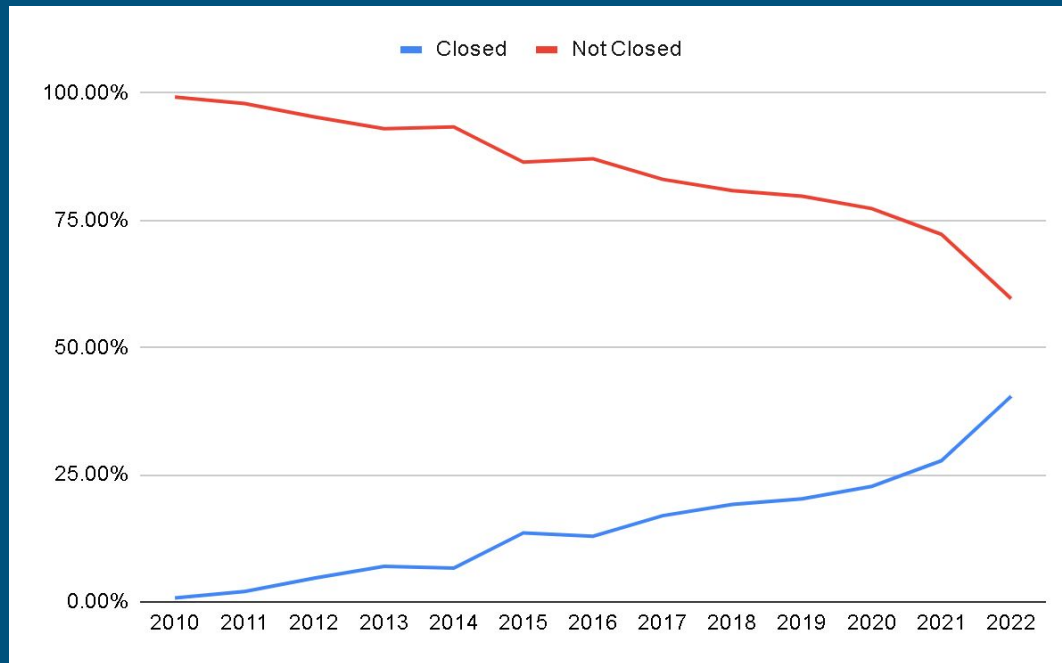
Percentual de Pull Requests mergeados versus não mergeados em média para cada ano

4.3.1 - Pergunta 3, métrica 1



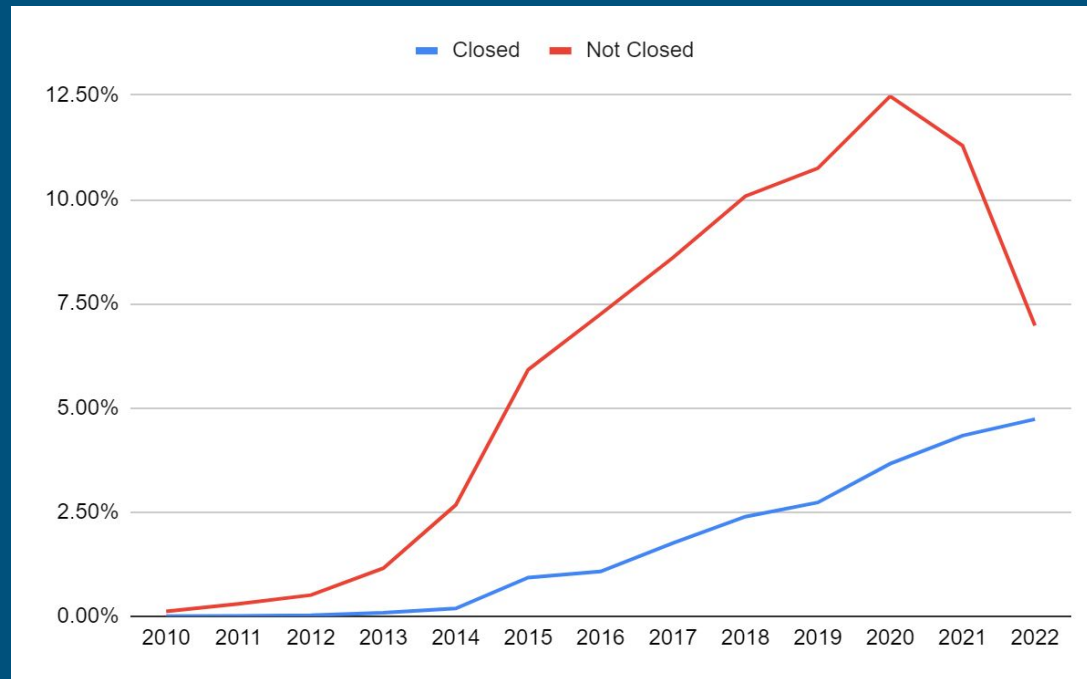
Percentual de Pull Requests *merged* versus não *merged* em média para todos os anos

4.3.2 - Pergunta 3, métrica 2



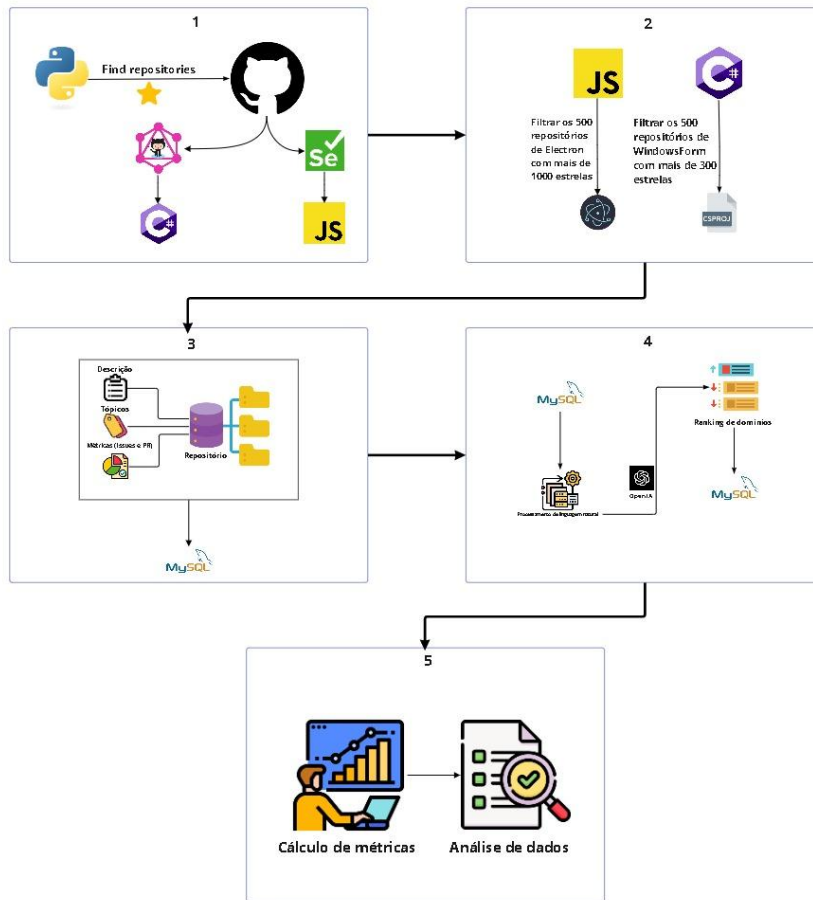
Percentual de Issues fechadas versus não fechadas em média para cada ano

4.3.2 - Pergunta 3, métrica 2



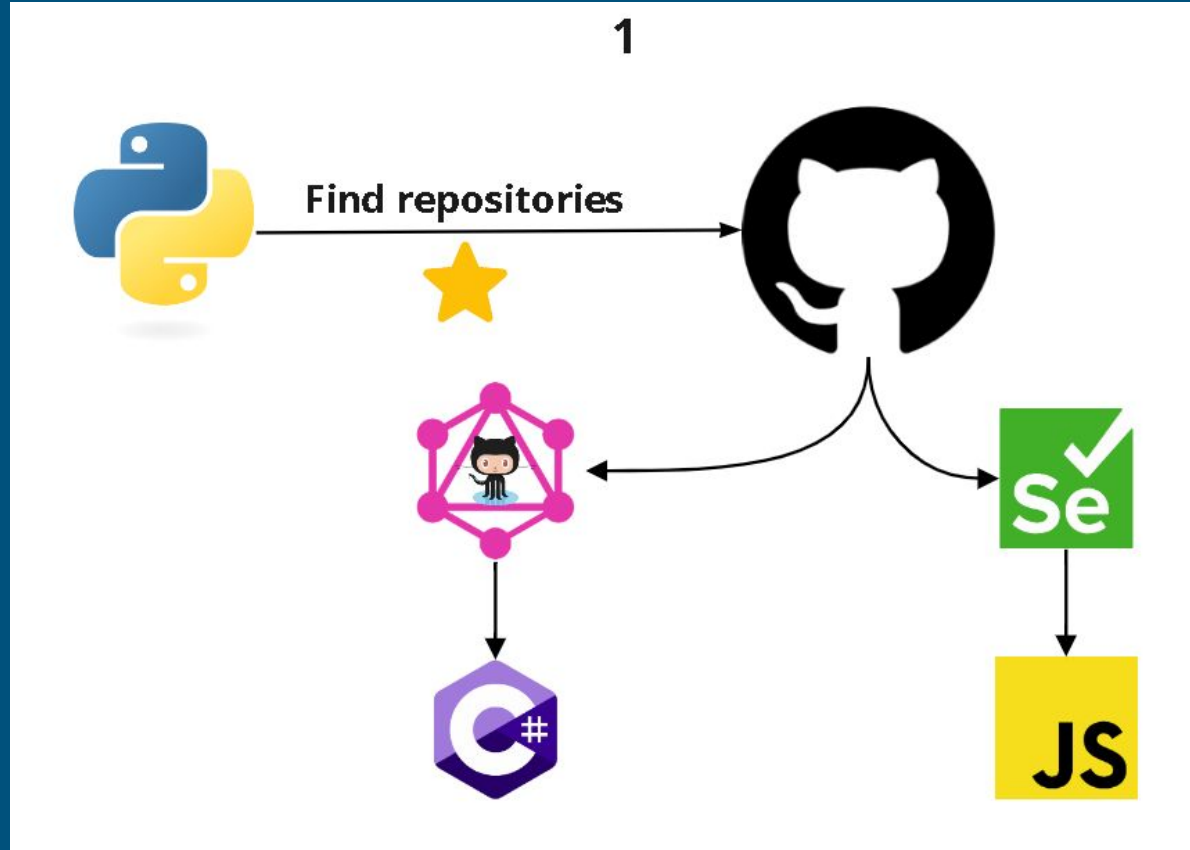
Percentual de issues fechadas em relação a não fechadas em repositórios desktop por ano.

5 - Metodologia



5.1 - Busca de repositórios

O passo 1 desta metodologia, foi buscar repositórios populares por meio de scripts desenvolvidos em Python, que acessam o site do Github por meio da API (Interface de Programação de Aplicação, do inglês Application Programming Interface) de GraphQL e o Selenium. Com a API, foi buscado os repositórios das linguagens C# que possuem WindowsForms, ou seja, uma aplicação desktop. Por fim, para detectar os repositórios de JavaScript, foi utilizado o Selenium que detecta os repositórios dependentes da dependência de aplicação desktop Electron.

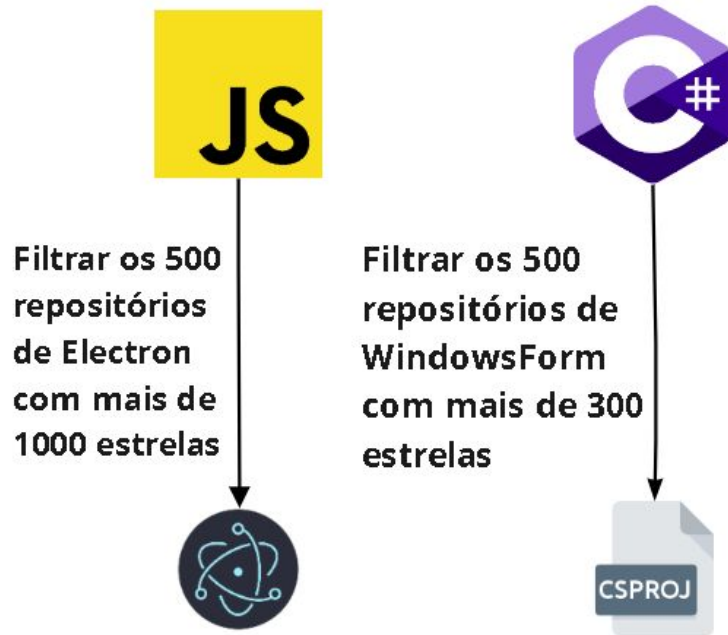


5.2 - Filtrando repositórios

O passo 2, que se baseia em processar todos os repositórios detectados, capturando as informações das dependências, textos e as métricas do repositório, que inclui dados de quantitativos como quantidade de estrelas, issues e pull requests em diferentes estados, isto para cada repositório.

Dados e Métricas para: P2M1.

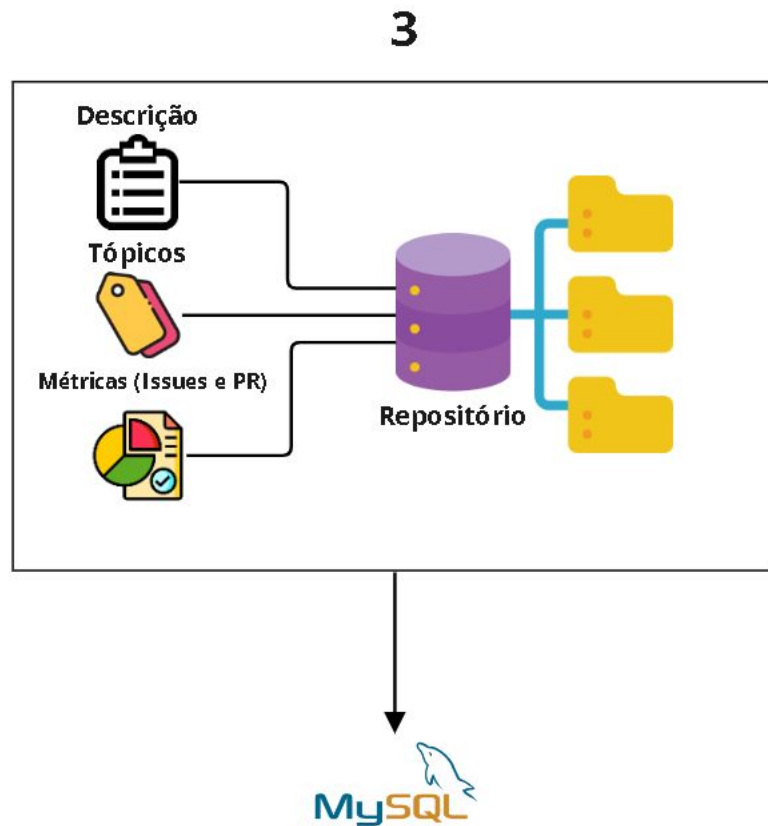
2



5.3 - Capturando dados

O passo 3, salva as informações de métricas, descrição e tópicos (tags do Github) dos repositórios no MySQL. Tais informações serão essenciais para o objetivo do projeto nos próximos passos.

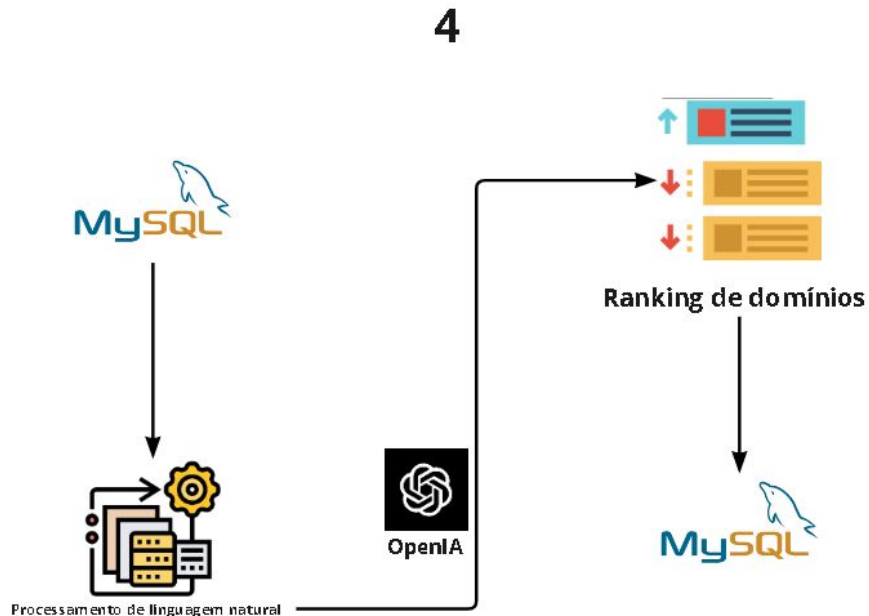
Dados e Métricas para: P1M1, P2M2, P3M1 e P3M2.



5.4 - Processando domínios

No passo 4, inicia-se um estágio muito complexo, no qual se baseia em utilizar processamento de linguagem natural da OpenIA, criado a partir da GPT-3, para detectar as palavras chaves com os melhores níveis de precisão. Sendo possível assim, determinar automaticamente qual é o mais provável domínio de cada repositório a partir dos dados salvos no MySQL. Ao finalizar, os prováveis domínios e seus percentuais são salvos no MySQL.

Dados e Métricas para: P1M2.



5.5 - Análise de resultados

Analisa-se os resultados e calculadas as métricas para auxiliar a responder as perguntas levantadas e alcançar o objetivo definido nas fases iniciais do trabalho

5



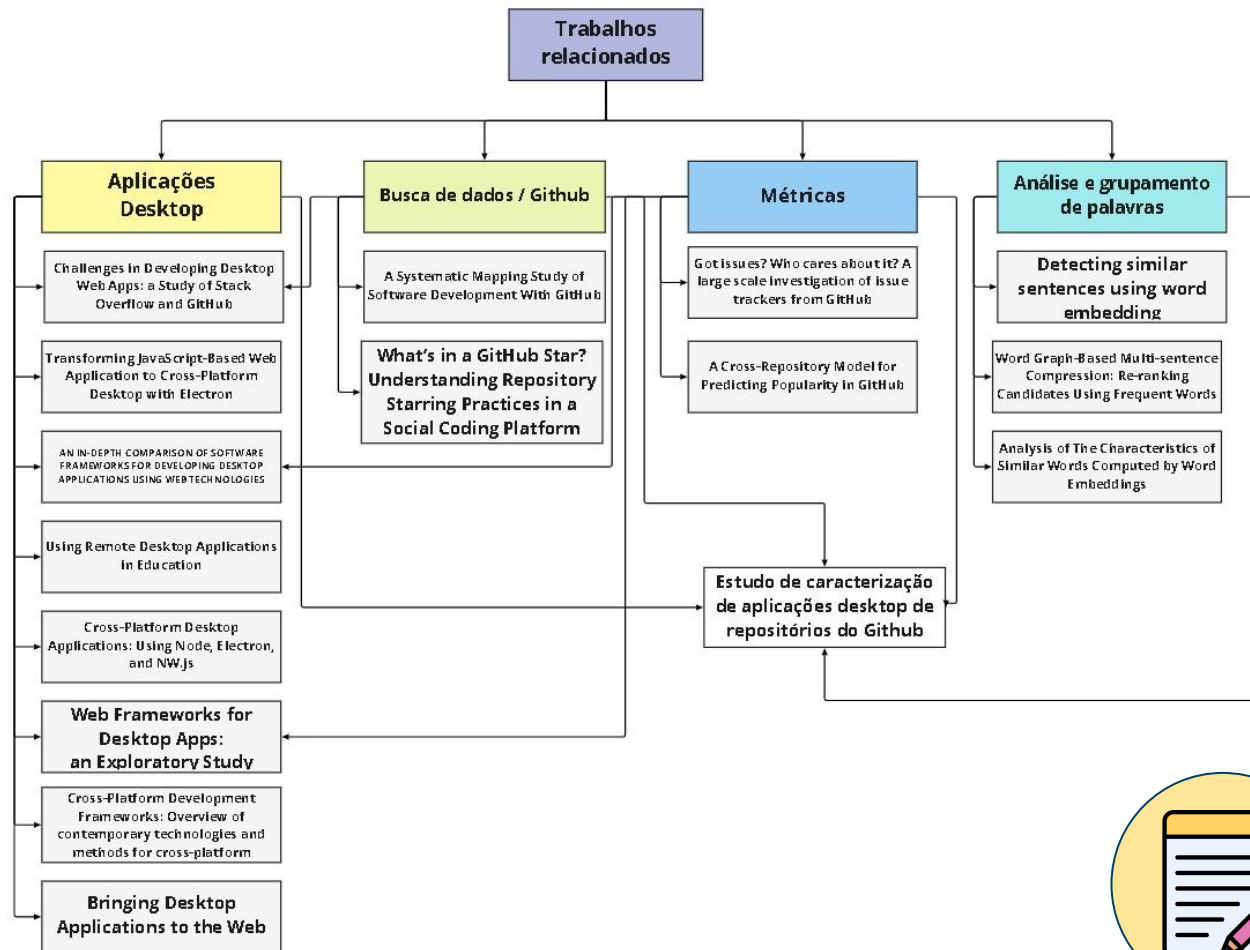
Cálculo de métricas



Análise de dados



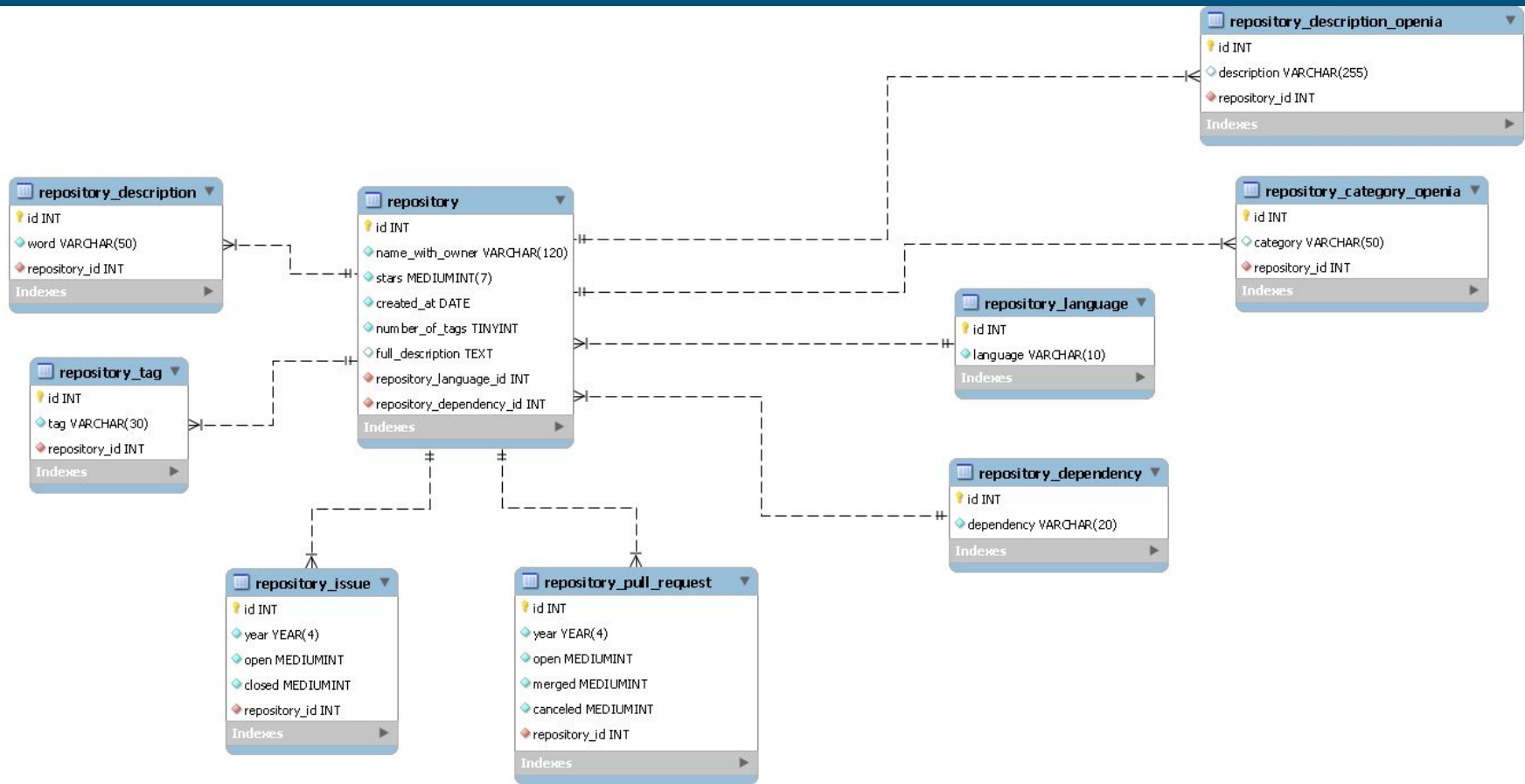
6 - Textos relacionados



7 - Proof of concept



7 - Proof of concept (Modelo do banco de dados)



7 - Proof of concept (Planilha da POC)

	A	B	C	D	E	F	G	H	I	J	K	L	M
1		LOWER(category)	Contagem		Primeiro crivo		Segundo crivo	Final					
2	<input checked="" type="checkbox"/>	analysis	3	<input type="checkbox"/>	analysis	<input type="checkbox"/>							
3	<input checked="" type="checkbox"/>	animation	3	<input checked="" type="checkbox"/>	animation	<input checked="" type="checkbox"/>	animation	animation,					
4	<input checked="" type="checkbox"/>	artificial intelligence	2	<input checked="" type="checkbox"/>	artificial intelligence	<input checked="" type="checkbox"/>	artificial intelligence	artificial intelligence,					
5	<input checked="" type="checkbox"/>	authenticator	2	<input checked="" type="checkbox"/>	authenticator	<input checked="" type="checkbox"/>	authenticator	authenticator,					
6	<input checked="" type="checkbox"/>	automation	6	<input checked="" type="checkbox"/>	automation	<input checked="" type="checkbox"/>	automation	automation,					
7	<input checked="" type="checkbox"/>	backup	4	<input type="checkbox"/>	backup	<input type="checkbox"/>							
8	<input checked="" type="checkbox"/>	browser	14	<input checked="" type="checkbox"/>	browser	<input checked="" type="checkbox"/>	browser	browser,					
9	<input checked="" type="checkbox"/>	browser extension	3	<input type="checkbox"/>	browser extension	<input type="checkbox"/>							
10	<input type="checkbox"/>	browsers	4	<input type="checkbox"/>		<input type="checkbox"/>							
11	<input checked="" type="checkbox"/>	converter	2	<input checked="" type="checkbox"/>	converter	<input checked="" type="checkbox"/>	converter	converter,					
12	<input checked="" type="checkbox"/>	cryptocurrency	13	<input checked="" type="checkbox"/>	cryptocurrency	<input checked="" type="checkbox"/>	cryptocurrency	cryptocurrency,					
13	<input type="checkbox"/>	cryptography	4	<input type="checkbox"/>		<input type="checkbox"/>							
14	<input type="checkbox"/>	data analysis	4	<input type="checkbox"/>		<input type="checkbox"/>							
15	<input checked="" type="checkbox"/>	data visualization	8	<input checked="" type="checkbox"/>	data visualization	<input checked="" type="checkbox"/>	data visualization	data visualization,					
16	<input checked="" type="checkbox"/>	database	6	<input checked="" type="checkbox"/>	database	<input checked="" type="checkbox"/>	database	database,					
17	<input type="checkbox"/>	databases	7	<input type="checkbox"/>		<input type="checkbox"/>							
18	<input type="checkbox"/>	deep learning	5	<input type="checkbox"/>		<input type="checkbox"/>							
19	<input type="checkbox"/>	desktop automation	2	<input type="checkbox"/>		<input type="checkbox"/>							
20	<input checked="" type="checkbox"/>	downloader	4	<input checked="" type="checkbox"/>	downloader	<input checked="" type="checkbox"/>	downloader	downloader,					
21	<input type="checkbox"/>	downloading	2	<input type="checkbox"/>		<input type="checkbox"/>							
22	<input checked="" type="checkbox"/>	editor	20	<input checked="" type="checkbox"/>	editor	<input checked="" type="checkbox"/>	editor	editor,					
23	<input checked="" type="checkbox"/>	encryption	7	<input type="checkbox"/>	encryption	<input type="checkbox"/>							
24	<input checked="" type="checkbox"/>	extension	10	<input checked="" type="checkbox"/>	extension	<input checked="" type="checkbox"/>	extension	extension,					
25	<input type="checkbox"/>	extensions	3	<input type="checkbox"/>		<input type="checkbox"/>							
26	<input type="checkbox"/>	file conversion	2	<input type="checkbox"/>		<input type="checkbox"/>							
27	<input checked="" type="checkbox"/>	file management	4	<input checked="" type="checkbox"/>	file management	<input checked="" type="checkbox"/>	file management	file management,					
28	<input checked="" type="checkbox"/>	file sharing	3	<input checked="" type="checkbox"/>	file sharing	<input type="checkbox"/>	file sharing						
29	<input checked="" type="checkbox"/>	firewall	3	<input type="checkbox"/>	firewall	<input type="checkbox"/>							
30	<input checked="" type="checkbox"/>	framework	20	<input checked="" type="checkbox"/>	framework	<input checked="" type="checkbox"/>	framework	framework,					
31	<input checked="" type="checkbox"/>	image processing	3	<input type="checkbox"/>	image processing	<input type="checkbox"/>							

marcar um x se fizer sentido a palavra (as primeiras tao marcada só de exemplo). apos ultimo crivo elas vem aqui pra lista abaixo

LISTA DAS PALAVRAS

animation,artificial intelligence,authenticator,automation,browser,converter,cryptocurrency,data visualization,database,downloader,editor,extension,file management,framework,image viewer,library,object detection,peer-to-peer,reader,remote desktop,screen capture,screen recorder,screen sharing,steam,streaming,task manager,taskbar,time tracking,usb,3d,audio,blockchain,bot,camera,chat,color palette,compression,crm,design,diagram,discord,docker,drawing,drivers,email,engine,finance,game,geospatial,ide,iot,keyboard,kubernetes,midi,minecraft,monitor,music,notes,notifications,player,plugin,pomodoro,proxy,security,server,terminal,test,voice,wallet,

7 - Proof of concept (Domínios em planilha)

Domínios de aplicações descobertas pela OpenAI (utilizando GPT-3), de aplicações Electron e WindowsForms:



https://docs.google.com/spreadsheets/d/18AUkPrT_lcXHZNaq_RolO4Y97ceDX6BHAixA5vG210s



Referência

- <https://xgen.com.br/blog/aplicacao-web-ou-desktop-qual-a-melhor-solucao>
- <https://wktechnology.com.br/aplicacoes-desktop-o-que-sao-e-como-funcionam/>
- <https://www.quora.com/Is-it-still-worth-to-build-desktop-applications>
- <https://www.devmedia.com.br/forum/aplicacoes-desktop/599445>
- <https://skymail.com.br/2022/07/05/e-possivel-manter-aplicacoes-desktop-em-nuvem/>
- <https://beta.openai.com/>



Bibliografia

- PRESSMAN, Roger S.; MAXIM, Bruce R.. Engenharia de software: uma abordagem profissional. 8. ed. Porto Alegre: AMGH, 2016. ISBN 978-8580555349. E-book (Livro Eletrônico).
- SOMMERVILLE, Ian. Engenharia de software. 10. ed. São Paulo: Pearson, 2019. ISBN 978-8543024974. E-book (Livro Eletrônico).
- Basili, Victor; Gianluigi Caldiera; H. Dieter Rombach (1994). The Goal Question Metric Approach.



FIM

