



**UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CENTRO TECNOLÓGICO**

DEPARTAMENTO DE INFORMÁTICA E ESTATÍSTICA

Marcos Demétrio

**Coleta de dados do Reddit para análise das linguagens de programação  
mais populares do mercado**

Marcos Demétrio

**Coleta de dados do Reddit para análise das linguagens de programação  
mais populares do mercado**

Trabalho de Conclusão de Curso submetida ao  
Departamento de Informática e Estatística da  
Universidade Federal de Santa Catarina para a  
obtenção do título de Bacharel em Sistemas de  
Informação.  
Orientador: Prof. José Eduardo De Lucca

Dedico este trabalho à minha mãe, Elisa Neli Rehn,  
e ao meu pai Sergio Pereira Demétrio.

## **AGRADECIMENTOS**

Gostaria de agradecer primeiramente aos meus familiares e amigos que sempre me incentivaram e apoiaram em cada etapa da minha trajetória acadêmica. Sem a ajuda deles, certamente não teria chegado até aqui.

Gostaria também de expressar minha gratidão ao professor Eduardo De Lucca, que com sua orientação e contribuições, me ajudou a concluir este trabalho de conclusão do curso.

É com muito orgulho que encerro esta etapa da minha jornada acadêmica como aluno do curso de Sistemas de Informação da Universidade Federal de Santa Catarina. Feliz com o aprendizado adquirido ao longo desses anos, que será base para as próximas etapas da minha vida profissional e pessoal.

*“Tudo evolui; não há  
realidades eternas: tal como  
não há verdades  
absolutas.”  
(Nietzsche)*

## RESUMO

Este trabalho se propôs a desenvolver uma aplicação web que permita a análise das principais linguagens de programação do mercado, utilizando amostras de dados coletados da plataforma Reddit. A aplicação fornece informações sobre engajamento, popularidade e interações dos usuários de cada comunidade. O trabalho também aborda a análise de pesquisas renomadas que fazem o ranqueamento de popularidade das linguagens de programação, como a Pesquisa Anual de desenvolvedores do Stack Overflow e o TIOBE Index.

**Palavras-chave:** Desenvolvedores, Linguagens de programação, *web*, Reddit

## **ABSTRACT**

This work proposed to develop a web application that allows the analysis of the main programming languages in the market, using samples of data collected from the Reddit platform. The application provides information about engagement, popularity and user interactions for each community. The work also addresses the analysis of renowned surveys that rank the popularity of programming languages, such as the Annual Survey of Stack Overflow Developers and the TIOBE Index.

**Keywords:** Developers, Programming languages, web, Reddit.

## LISTA DE FIGURAS

Figura 1 – Top 10 países da pesquisa

Figura 2 – Faixa etária dos desenvolvedores

Figura 3 – Ranking das linguagens da Pesquisa Anual do Stack Overflow

Figura 4 – Ranking de linguagens de programação pelo TIOBE Index

Figura 5 – Plataforma do Reddit

Figura 6 – Subreddit

Figura 7 - Detalhes da postagem

Figura 8 – Plataforma do Twitter

Figura 9 – Notícia no portal Terra

Figura 10 – Diagrama de caso de uso: Usuário.

Figura 11 – Banco de dados NoSQL

Figura 12 – Consultas a API do Reddit

Figura 13 – Tela de análise de comunidades

Figura 14 – Tela de análise de engajamento

Figura 15 – Detalhes da comunidade

Figura 16 – Análise das últimas postagens

Figura 17 – Script de coleta diária dos dados



## LISTA DE ABREVIATURAS E SIGLAS

API	Application Programming Interface
REST	Representational State Transfer
HTTP	HyperText Transfer Protocol
SQL	Structured Query Language
NoSQL	Not Only SQL
JSON	Javascript Object Notation
URL	Uniform Resource Locator
WEB	World Wide Web
UFSC	Universidade Federal de Santa Catarina

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>11</b>
1.1	OBJETIVOS	12
1.1.1	Objetivo Geral	12
1.1.2	Objetivos Específicos	12
1.2	Elaboração do projeto	12
<b>2</b>	<b>TRABALHOS CORRELATOS</b>	<b>13</b>
2.1	Pesquisa correlatas	13
2.1.1	Pesquisa Anual de Desenvolvedores do Stack Overflow	13
2.1.2	TIOBE Index	16
2.2	Redes sociais com concentração de desenvolvedores	17
2.2.1	Reddit	17
2.2.2	Twitter	20
<b>3</b>	<b>MODELAGEM DO PROJETO</b>	<b>23</b>
3.1	Descrição das Etapas	23
3.2	Especificação dos requisitos	23
3.2.1	Requisitos não funcionais	24
3.2.2	Requisitos funcionais	25
3.3	DIAGRAMA DE CASO DE USO	<b>26</b>
<b>4</b>	<b>DESENVOLVIMENTO</b>	<b>28</b>
4.1	Tecnologias utilizadas	28
4.1.1	React	28
4.1.2	Node.js	29
4.1.3	Express.js	29
4.1.4	MongoDB	29
4.1.5	API do React	30
4.1.6	Axios	30
4.2	Banco de dados NoSQL	31
4.3	Desenvolvimento do software	33
4.3.1	Coleta diária dos dados	37
<b>5</b>	<b>CONCLUSÃO</b>	<b>39</b>
5.1	TRABALHOS FUTUROS	<b>39</b>
	<b>REFERÊNCIAS</b>	<b>41</b>
	<b>REPOSITÓRIO DO CÓDIGO</b>	<b>43</b>
	<b>APÊNDICE A</b>	<b>44</b>

## 1 INTRODUÇÃO

Existem milhares de linguagens de programação com diferentes tipos de abordagens e aplicações. A plataforma Stack Overflow<sup>1</sup>, consagrada entre os desenvolvedores, divulga anualmente estatísticas e *rankings* das linguagens de programação mais utilizadas pelos desenvolvedores, por meio de um formulário disponível anualmente aos usuários de todo o mundo. No ano de 2022 foram enviados ao todo 83.439 formulários preenchidos, onde 57,86% desse total é oriundo de 10 países, Estados Unidos, Índia, Alemanha, Reino Unido da Grã-Bretanha e Irlanda do Norte, Canadá, França, Brasil, Polônia, Países Baixos e Itália. A primeira colocação ficou com os Estados Unidos, que contou com 18.33% de todas as respostas enviadas.

Segundo a plataforma, quase 60% dos entrevistados dizem ter aprendido a programar a partir de recursos online. Os mais jovens tendem a aprender com cursos online, fóruns e outros recursos na internet. Já os entrevistados mais velhos, aprenderam com meios mais tradicionais, como na escola, universidade e em livros. Do total de entrevistados, aproximadamente 50% estão desenvolvendo *software* há dez anos ou menos, com maior faixa entre 5 e 9 anos, contando com um total de 29.91%. Dos desenvolvedores profissionais que responderam ao formulário, a maior parte, 48%, têm entre 25 e 34 anos. Já das linguagens mais utilizadas pelos desenvolvedores, está na primeira colocação a linguagem JavaScript<sup>2</sup>.

Assim como no Stack Overflow, os desenvolvedores reúnem-se em diversos outros ambientes virtuais para troca de informações. Entre eles o Reddit<sup>3</sup>, uma rede social que também é um agregador de notícias, permitindo que os usuários compartilhem e discutam informações de uma ampla variedade de tópicos. Todas essas comunidades são uma excelente opção para os desenvolvedores que buscam compartilhar conhecimentos e experiências, se mantendo atualizados sobre as últimas tendências na área, além de estabelecer contatos valiosos com outros profissionais do setor.

<sup>1</sup><https://stackoverflow.com/>

<sup>2</sup>JavaScript é uma linguagem de programação sendo uma das principais tecnologias da World Wide Web

<sup>3</sup><https://reddit.com/>

## **1.1 OBJETIVOS**

### **1.1.1 Objetivo Geral**

O objetivo principal deste trabalho é permitir a análise sobre as principais linguagens de programação utilizadas no mercado, com a construção de uma ferramenta que capture dados de alguma plataforma relevante e ajude na análise comparativa das linguagens de programação escolhidas para análise, tornando-se mais uma ferramenta de análise e tomada de decisão para empresas e profissionais da área de tecnologia, como por exemplo, sobre quais linguagens de programação devem investir em seus projetos futuros.

### **1.1.2 Objetivos Específicos**

Para concluir o objetivo do projeto, adotou-se os seguintes objetivos específicos:

- Levantar e analisar as principais pesquisas que indicam a popularidade das linguagens de programação: O objetivo é analisar as fontes mais confiáveis para obter quais as linguagens de programação mais populares do mercado.
- Avaliar e escolher alguma plataforma relevante, que contenham discussões sobre linguagens de programação: Será avaliada a forma de se conectar com a API da plataforma e coletar os dados que mencionem as linguagens de programação selecionadas na pesquisa.
- Comparar a popularidade de cada linguagem, com base na frequência das postagens que façam referência a elas: Realizar uma análise quantitativa das postagens coletadas, para em seguida comparar a popularidade de cada linguagem de programação com base na frequência de menções nas postagens coletadas.
- Construir interface para visualizar os resultados de maneira interativa: Etapa de construção da interface do sistema, utilizando gráficos e outros elementos visuais para apresentar de forma clara e concisa as informações coletadas. Será possível filtrar os resultados por linguagem e data.

## **2 TRABALHOS CORRELATOS**

### **2.1 Pesquisas correlatas**

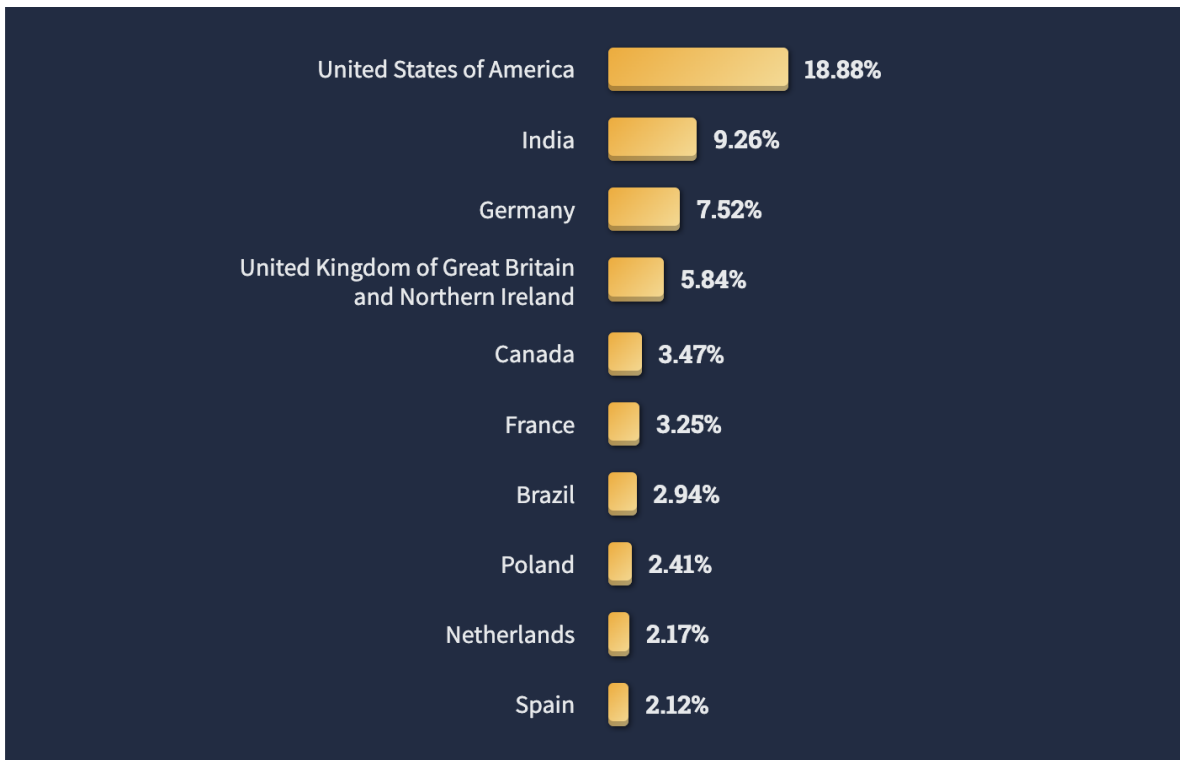
O objetivo deste capítulo é examinar duas das pesquisas mais relevantes em medir a popularidade das linguagens de programação, que são a Pesquisa Anual de Desenvolvedores do Stack Overflow e a pesquisa TIOBE Index. Ambas as pesquisas têm como propósito avaliar a popularidade das linguagens de programação ao redor do mundo, com o intuito de fornecer informações sobre as tendências atuais do mercado de tecnologia. No entanto, essas pesquisas têm maneiras distintas na forma como coletam e analisam os dados. Dessa forma, ao estudar essas pesquisas, permite-se ter uma melhor compreensão sobre as preferências dos desenvolvedores de software, além de obter informações valiosas para tomada de decisões no campo da tecnologia.

#### **2.1.1 Pesquisa Anual de Desenvolvedores do Stack Overflow**

A pesquisa vem sendo aplicada anualmente desde 2011, onde são analisados vários aspectos das preferências dos desenvolvedores, como as tecnologias favoritas e o controle de versão favorito. A pesquisa é aplicável tanto para aqueles que estão aprendendo a codificar, como para desenvolvedores profissionais. A plataforma Stack Overflow ([stackoverflow.com.br](https://stackoverflow.com.br)) recebe a visita de 100 milhões de usuários todos os meses, sendo um dos sites mais acessados do mundo. Trata-se de uma plataforma de perguntas e respostas na área de tecnologia e desenvolvimento de software.

No ano de 2022, a Pesquisa Anual de Desenvolvedores do Stack Overflow recebeu mais de 70 mil respostas de mais de 180 países, pesquisa esta que é realizada todos os anos desde 2011, onde se examinam diversos aspectos da experiência dos desenvolvedores de *software* através de questionário com perguntas como quais suas tecnologias favoritas e experiências profissionais.

Figura 1 – Top 10 países da pesquisa

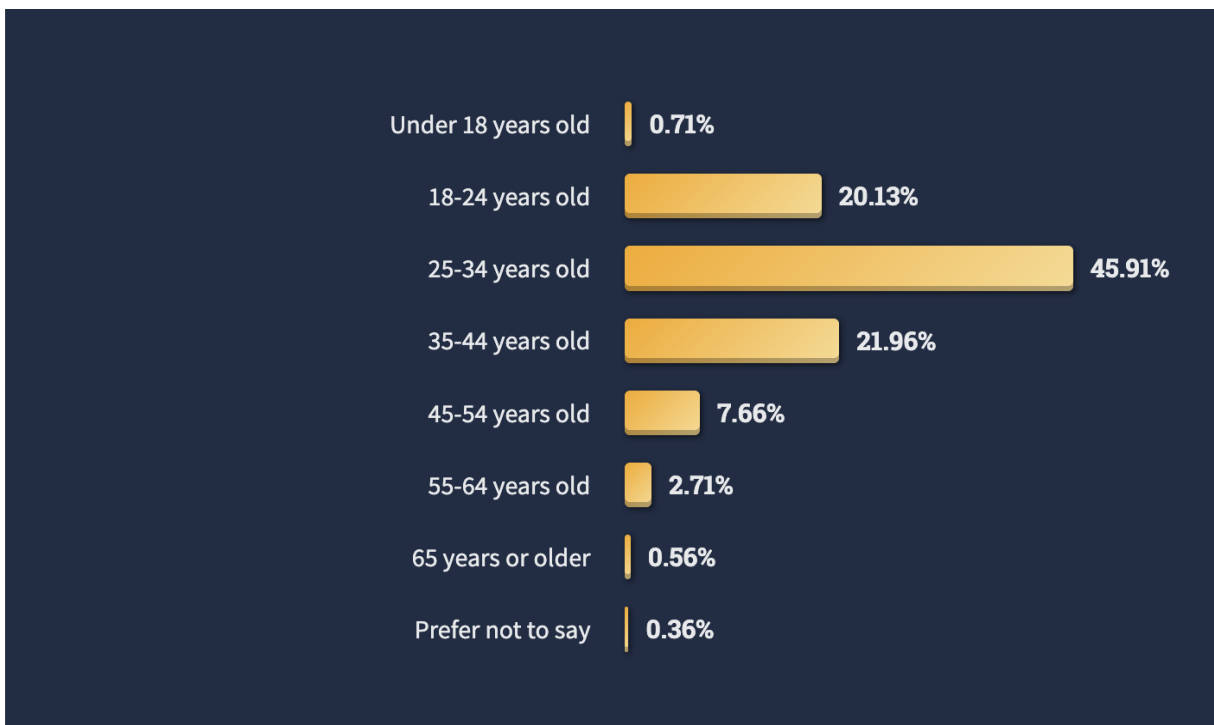


Fonte: Printscreen da página <https://survey.stackoverflow.co/2022/#key-territories-country>, feito em 17/10/2022.

O gráfico apresentado na figura 1, mostra os 10 países com mais desenvolvedores que participaram da pesquisa Anual de Desenvolvedores do Stack Overflow do ano de 2022, contando com 49,116 respostas no total. A partir dele, é possível verificar que os Estados Unidos lideram a lista com mais de 9.000 desenvolvedores participantes, seguido pela Índia, em segundo lugar com mais de 4.500 participantes. O Brasil aparece em sétimo lugar, com mais de 1400 participantes.

Na pesquisa Anual de Desenvolvedores do Stack Overflow do ano de 2022, uma das informações mais interessantes apresentadas foi a faixa etária dos desenvolvedores que atuam profissionalmente no mercado de desenvolvimento de software, revelando que o cenário é composto na sua maioria por jovens, onde mais de 66% dos entrevistados possuem entre 18 e 34 anos.

Figura 2 – Faixa etária dos desenvolvedores

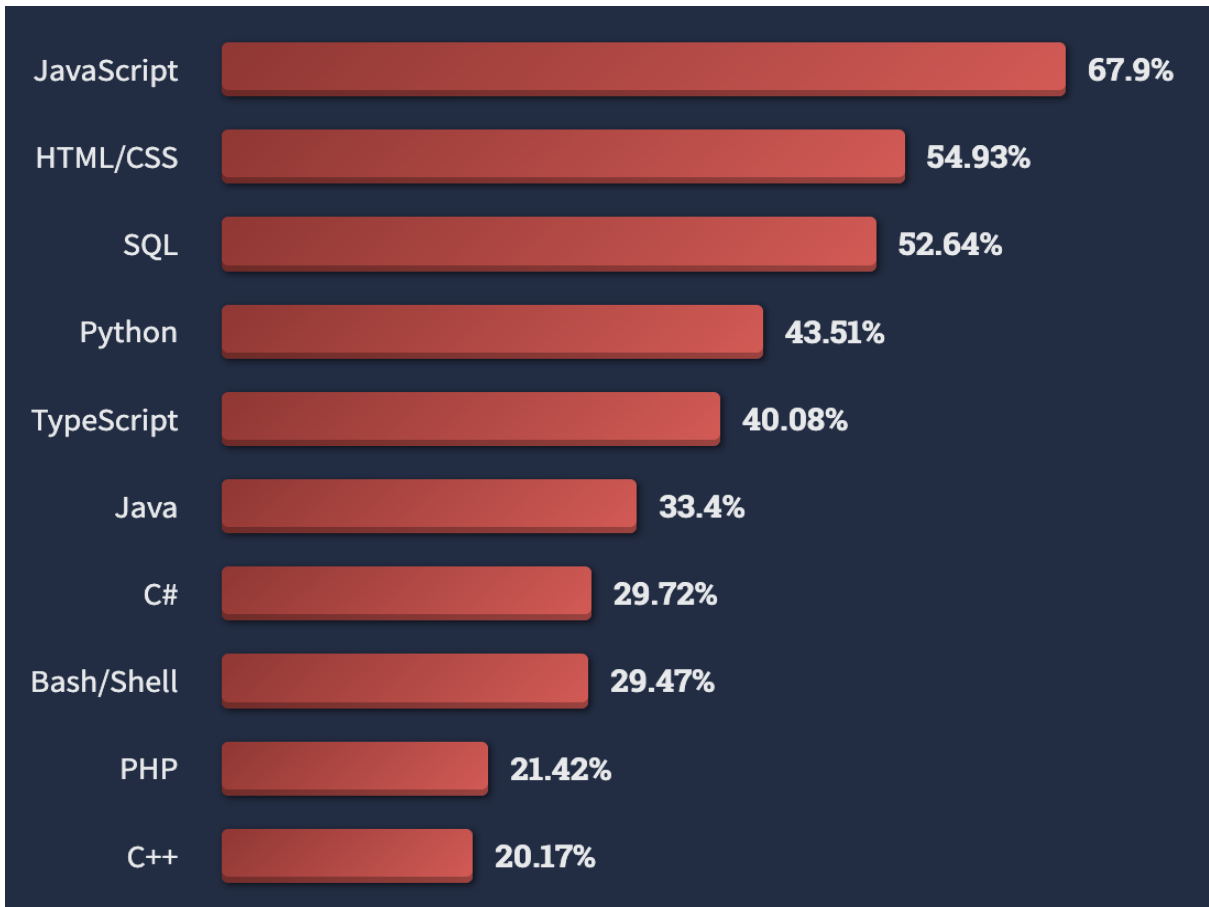


Fonte: Printscreen da página <https://survey.stackoverflow.co/2022/#key-territories-country>, feito em 17/10/2022.

O gráfico da figura 2 apresenta as porcentagem da análise de distribuição de idade dos desenvolvedores participantes da pesquisa, que totalizaram 53.041 de respostas no total. Apesar da concentração de profissionais de sua maioria jovens, a pesquisa deste ano também evidencia a presença de desenvolvedores com mais experiências no mercado de trabalho. As faixas etárias acima de 45 anos, juntas, representam mais de 11% dos profissionais que responderam a pesquisa.

A figura 3 indica o ranking das principais linguagens de programação de 2022, utilizadas pelos desenvolvedores profissionais que responderam à Pesquisa Anual de Desenvolvedores do Stack Overflow. O Javascript ocupou a primeira colocação, a qual vem ocupando há 10 anos em sequência na pesquisa, sendo um reflexo de sua ampla utilidade para diferentes tipos de projetos de desenvolvimento de software, bem como seu engajamento e popularidade entre uma vasta comunidade de desenvolvedores.

Figura 3 – Ranking das linguagens da Pesquisa Anual do Stack Overflow



Fonte: Printscreen da página <https://survey.stackoverflow.co/2022/#key-territories-country>, feito em 17/10/2022.

### 2.1.2 TIOBE Index

O site TIOBE mede a popularidade das linguagens de programação através das buscas realizadas na *web*. Mensalmente é mensurado o total de buscas por linguagens de programação nos motores de busca como Google, Bing, Yahoo!, Wikipedia, Amazon, YouTube, entre outros 19 buscadores. O índice conta com centenas de linguagens de programação, e dá uma visão geral das tendências atuais e futuras da indústria de tecnologia, pois acompanhando a evolução da popularidade das linguagens de programação, é possível ter uma ideia das tecnologias que estão emergindo, ajudando empresas de tecnologia a planejar seus investimentos.



Figura 4 – Ranking de linguagens de programação pelo TIOBE Index

Programming Language	2022	2017	2012	2007	2002	1997	1992	1987
Python	1	5	8	7	13	28	17	-
C	2	2	1	2	2	1	1	1
Java	3	1	2	1	1	16	-	-
C++	4	3	3	3	3	2	2	5
C#	5	4	5	8	12	-	-	-
Visual Basic	6	16	-	-	-	-	-	-
JavaScript	7	7	10	9	8	21	-	-
Assembly language	8	12	-	-	-	-	-	-
SQL	9	-	-	-	7	-	-	-
PHP	10	8	6	5	6	-	-	-

Fonte: Printscreen da página <https://www.tiobe.com/tiobe-index/>, feito em 02/12/2022.

Na figura 4 está o ranking de linguagens de programação de acordo com o TIOBE Index, pesquisa que é realizada desde 2001, sendo que a companhia foi fundada em outubro do ano 2000. A pesquisa é feita através do total de buscas realizadas em um total de 25 motores de busca e atualizada mensalmente. Por conta disso, a companhia consegue divulgar os resultados pegando anos anteriores a sua fundação, coletando resultados que vão até 1987, como mostra a figura.

De acordo com a pesquisa, Python aparece pela primeira vez como líder no ano de 2022. Desde então apenas as linguagens C e Java haviam conquistado a primeira colocação do TIOBE Index.

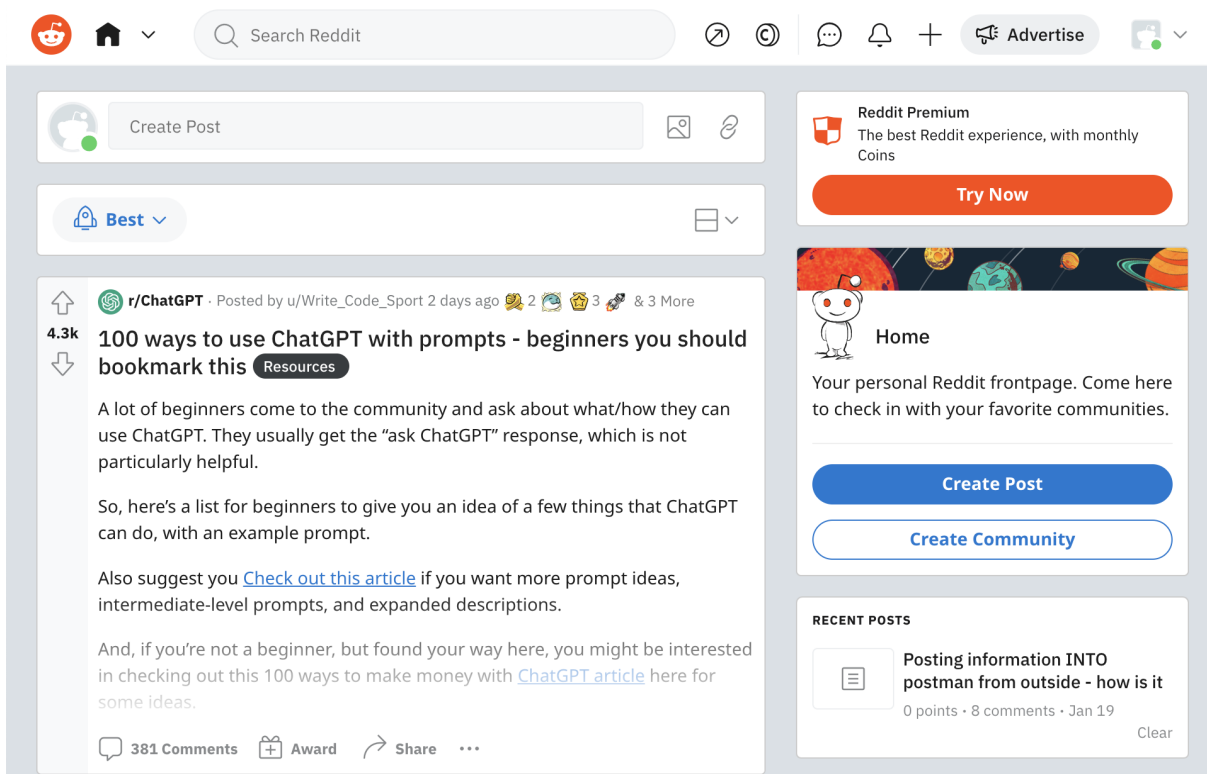
## 2.2 Redes sociais com concentração de desenvolvedores

### 2.2.1 Reddit

O Reddit ([www.reddit.com](http://www.reddit.com)) é uma rede social considerada um dos maiores agregadores de notícias e fóruns de discussões da internet, contendo uma ampla variedade de tópicos onde os usuários da plataforma compartilham conteúdo e participam das discussões. Existem tópicos sobre uma gama grande de temas como tecnologia, notícias, política, ciência e criptomoedas, por exemplo, onde os usuários podem avaliar as postagens dos outros usuários, positivamente ou negativamente, através da funcionalidade de votação que faz com que os usuários aumentem a relevância de um post na comunidade dentro da plataforma. Isso permite que as

postagens mais relevantes sejam destacadas e apareçam para mais usuários dentro da plataforma.

Figura 5 – Plataforma do Reddit

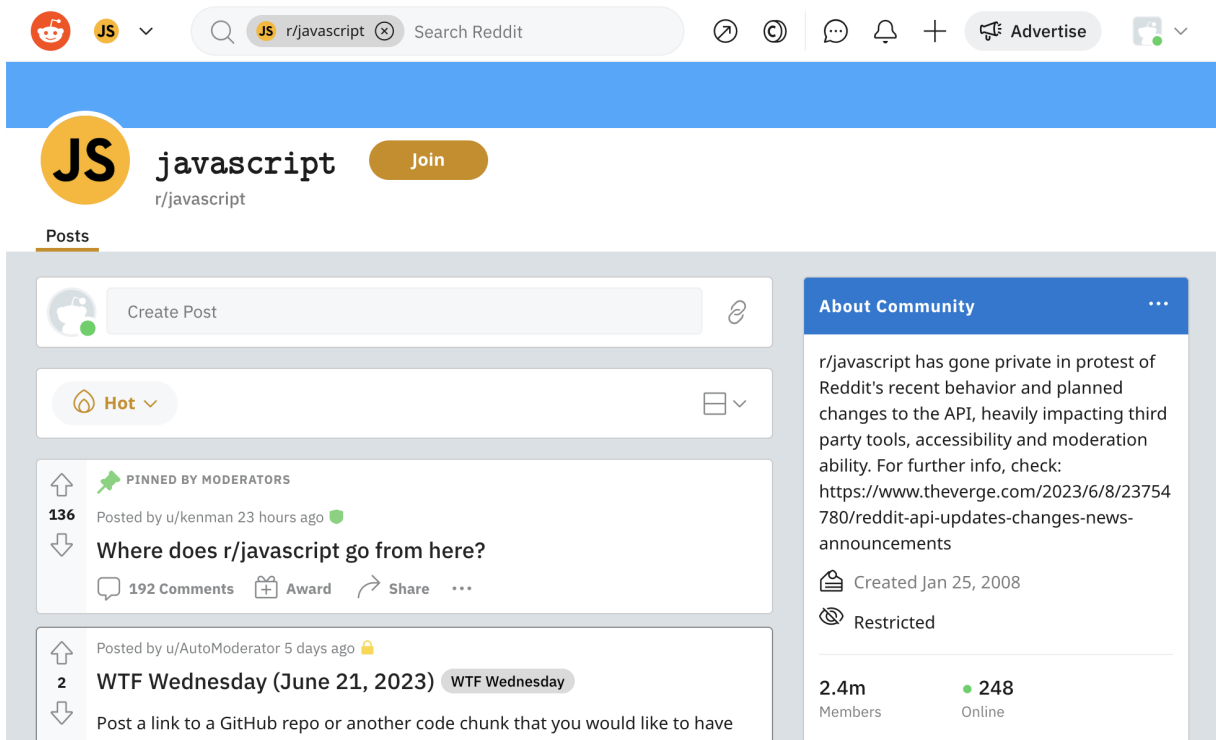


Fonte: Printscreen da página <https://reddit.com>, feito em 07/04/2023.

As postagens na plataforma são categorizadas através de subreddits, que são fóruns temáticos criados pelos próprios usuários. Por exemplo, existem subreddits dedicados à tecnologia, ciência, política, entre muitos outros temas. Dentro de cada subreddit, os usuários podem adicionar tags em suas postagens para ajudar na organização e na busca por conteúdo relevante. As tags são chamadas de "*flairs*" e podem ser definidas pelos moderadores do subreddit ou pelos próprios usuários, dependendo das configurações definidas para o subreddit. Além disso, os subreddits podem permitir que os usuários personalizem suas *flairs* com texto ou imagens.

Qualquer usuário pode criar subreddits, o que ocasiona em uma ampla variedade de subreddits existentes na plataforma. É importante frisar que a plataforma mantém moderadores que são os responsáveis por garantir que os subreddits sigam as regras da plataforma, removendo conteúdo inadequado e mantendo a discussão de acordo com as regras da rede social.

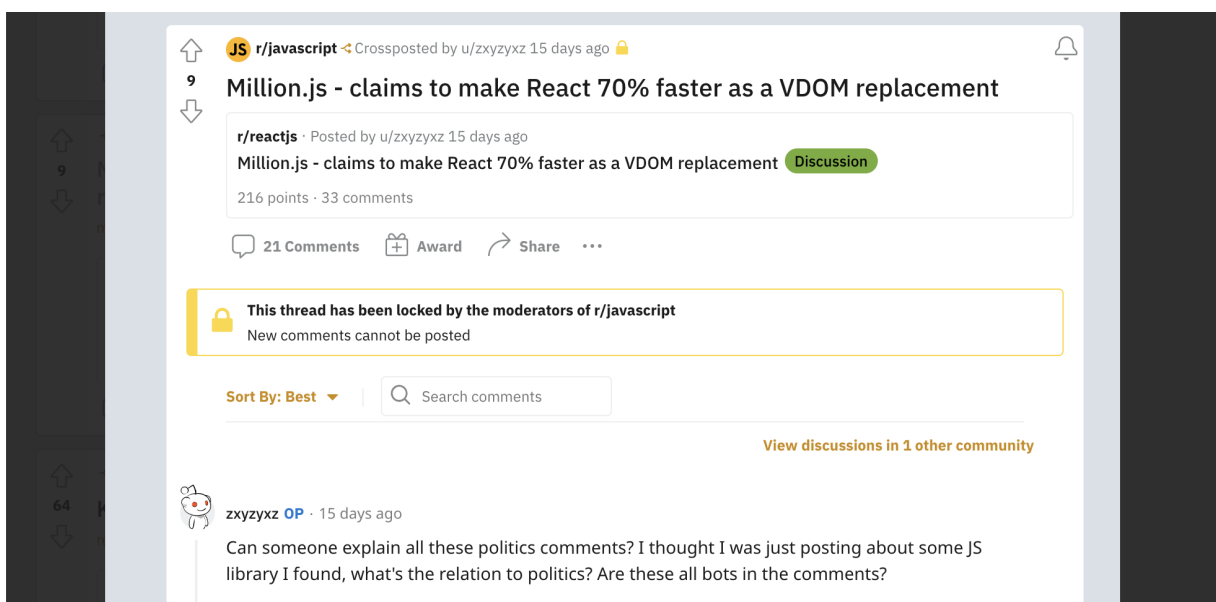
Figura 6 – Subreddit



Fonte: Printscreen da página <https://reddit.com>, feito em 25/05/2023.

A figura 6 mostra o subreddit Javascript, com um total de 2 milhões e 400 mil membros, onde é possível visualizar a descrição da comunidade, quantidade de membros, usuários online e as últimas postagens.

Figura 7 – Detalhes da postagem



Fonte: Printscreen da página <https://reddit.com>, feito em 25/05/2023.

A figura 7 mostra as informações presentes nos detalhes de uma postagem, contendo o título da postagem, o conteúdo da postagem, os votos dos usuários, os *flairs* em caso de o usuário ter utilizado, os comentários, entre outros.

Para que os desenvolvedores de software acessem os dados da plataforma, como postagens, comentários, usuários, entre outros, o Reddit disponibiliza uma API (*Application Programming Interface*), para que dessa forma os desenvolvedores possam integrá-los a suas aplicações. A API da plataforma é baseada no modelo arquitetural REST (*Representational State Transfer*), sendo totalmente gratuita e aberta para uso. Também é possível interagir com a plataforma através da API, postando novos conteúdos e gerenciando contas de usuário, onde é possível automatizar processos de gerenciamento de conteúdo, por exemplo.

### **2.2.1 Twitter**

O Twitter ([www.twitter.com](http://www.twitter.com)) é uma das maiores redes sociais do mundo, sendo utilizado por milhões de pessoas para compartilhar informações, notícias e opiniões sobre assuntos diversos. Na plataforma, os usuários podem seguir outras pessoas e empresas para verem suas postagens, chamadas de *tweets*. Assim como no Reddit, existe uma ampla variedade de tópicos que são discutidos no Twitter, incluindo política, tecnologia, esportes, entre outros.

No twitter é possível usar hashtags, que são palavras ou frases precedidas do símbolo "#", servindo para encontrar tweets relacionados a um determinado assunto em específico. Quando algum usuário adiciona uma hashtag em sua postagem, ela é automaticamente indexada a uma lista de tweets que utilizem a mesma hashtag.

Para que os desenvolvedores de software possam acessar os dados do Twitter, a plataforma disponibiliza exclusivamente aos desenvolvedores uma plataforma para acesso a sua API (<https://developer.twitter.com/en/docs/twitter-api>), permitindo a integração de aplicações com a plataforma do Twitter. A API do Twitter também é baseada no modelo arquitetural REST. Com a API, os desenvolvedores podem automatizar processos de gerenciamento de conteúdo, postagens, respostas automáticas, entre outros.

Figura 8 – Plataforma do Twitter



Fonte: Printscreen da página <https://twitter.com>, feito em 23/04/2023.

Apesar de o Twitter ser uma plataforma muito interessante para fazer a coleta de informações, por ser uma das maiores plataformas do planeta, com um enorme número de usuários ativos, a sua gestão vem gerando dúvidas sobre a continuação das APIs de acesso gratuito.

Por tais fatores, para este projeto optou-se pela plataforma Reddit, já que sua API é totalmente gratuita e aberta para uso, permitindo que desenvolvedores de software possam integrá-la às suas aplicações com facilidade e sem custo. Além disso, o Reddit é uma plataforma muito popular e ativa, possuindo uma grande quantidade de usuários ativos e dados disponíveis para análise.

Figura 9 – Notícia no portal Terra: "Twitter encerra a API gratuita e afeta aplicativos"



**Twitter encerra a API gratuita e afeta aplicativos**

O Twitter finalmente encerrou o acesso à sua API grátis, mas sem implementar efetivamente a versão paga, o que deixou muitos desenvolvedores e apps no escuro

 Alveni Lisboa

7 abr 2023 - 15h30 (atualizado às 22h06) [Compartilhar](#) [Ver comentários](#)

[Ouvir texto](#) 0:00 

O [Twitter](#) finalmente encerrou o suporte oficial à sua API após muitos adiamentos e polêmicas. A rede social interrompeu o uso da integração nesta quinta-feira (6), o que deixou vários aplicativos fora do ar.

Fonte:

<https://www.terra.com.br/byte/twitter-encerra-a-api-gratuita-e-afeta-aplicativos,b8f6343df2e78136bcd40319cdf5a9a6big3243o.html>

### **3 MODELAGEM DO PROJETO**

Adotou-se as etapas a seguir para a conclusão do projeto:

#### **3.1 Descrição das Etapas**

- Levantamentos de requisitos da aplicação: Identificação de todos os requisitos funcionais e requisitos não funcionais necessários para construção do projeto.
- Diagrama de caso de uso: Criação do diagrama de caso de uso para visualizar as funcionalidades mais importantes do sistema, no contexto de interação com o usuário do sistema.
- Prototipação das telas: Criação dos protótipos das telas da aplicação com o design da aplicação. Esses protótipos serão o guia para construção do layout no frontend do sistema.
- Modelagem do banco de dados: Realização da modelagem do banco de dados NoSQL, definindo as coleções e seus campos necessários para armazenar os dados da plataforma.
- Etapa de implementação: Desenvolvimento das funcionalidades da aplicação, integrando o backend em Node.js, com o frontend em ReactJS e o banco de dados MongoDB, seguindo os requisitos e protótipos das telas previamente definidos. Serão realizados testes de unidade e integração para garantir a qualidade do código e a integridade do sistema.
- Etapa de Implantação: Implantação da aplicação em ambiente de produção na hospedagem da plataforma em nuvem Heroku.

#### **3.2 Especificação dos requisitos**

Os requisitos foram identificados considerando a avaliação das informações obtidas por meio das pesquisas, Tiobe Index e a Stack Overflow Developer Survey, além de levar em consideração as informações documentadas na API do Reddit disponíveis em: <https://www.reddit.com/dev/api/>.

### 3.2.1 Requisitos não funcionais

#### **[RNF001] Programação**

Prioridade:  Essencial  Importante  Desejável

O sistema deve ser desenvolvido em linguagem Javascript usando Node.js junto ao framework Express.js no backend e utilizando a biblioteca React para interface web.

#### **[RNF002] Banco de dados**

Prioridade:  Essencial  Importante  Desejável

O sistema deve ser desenvolvido utilizando o banco de dados MongoDB.

#### **[RNF003] Compatibilidade**

Prioridade:  Essencial  Importante  Desejável

O sistema deve ser compatível com os navegadores Google Chrome e Mozilla Firefox.

#### **[RNF004] Responsividade**

Prioridade:  Essencial  Importante  Desejável

O sistema deve ser responsivo tanto para dispositivos móveis quanto para computadores, devendo ser adotado um design responsivo, fazendo com que a interface se adapte automaticamente ao tamanho de tela de celulares e computadores.

#### **[RNF005] Desempenho**

Prioridade:  Essencial  Importante  Desejável



O sistema deve processar e apresentar as informações coletadas de forma fluida, de forma que proporcione uma boa experiência de uso aos usuários.

### 3.2.2 Requisitos funcionais

#### [RF001] Comparação entre comunidades

Prioridade:  Essencial  Importante  Desejável

O sistema deverá permitir ao usuário realizar comparações entre diferentes comunidades de programação, com base em fatores como o número de seguidores, curtidas, comentários, o tamanho dos threads de discussão, entre outros.

#### [RF002] Filtrar as postagens por linguagem

Prioridade:  Essencial  Importante  Desejável

O usuário deve ter a opção de visualizar as postagens individualmente por linguagem de programação.

#### [RF003] Análise da evolução temporal dos dados

Prioridade:  Essencial  Importante  Desejável

O sistema deve apresentar uma análise da evolução temporal dos dados, permitindo ao usuário visualizar como a popularidade das linguagens de programação mudou ao longo do tempo.

#### [RF004] Análise de flairs

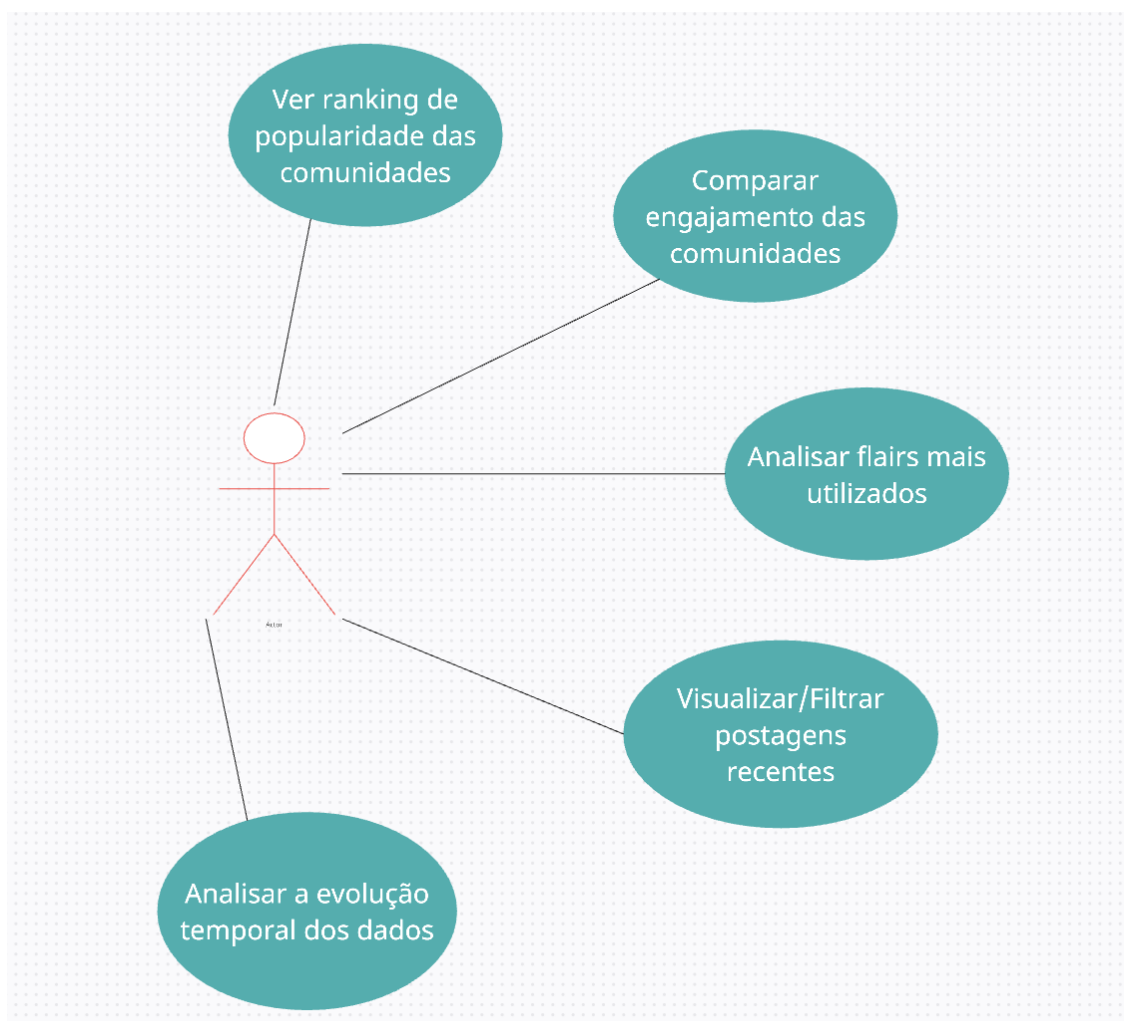
Prioridade:  Essencial  Importante  Desejável

O sistema deve permitir realizar a análise dos *flairs* utilizados nas postagens coletadas e permitir filtrar as postagens pelo *flair* selecionado.

### 3.3 DIAGRAMA DE CASO DE USO

O diagrama de caso de uso mostra as funcionalidades do sistema tendo como referência as interações do usuário do sistema, mostrando de forma clara e simples quais são as funcionalidades que o sistema oferece para atender as necessidades do usuário. Neste caso, um usuário com interesse em conhecer quais as linguagens mais populares do mercado.

Figura 10 – Diagrama de caso de uso: Usuário.



Fonte: Autor

O diagrama de caso de uso mostrado na figura 10 descreve as interações do usuário com o sistema. Na página principal, é exibido um ranking com algumas das comunidades de linguagens de programação mais engajadas do Reddit. O usuário pode então fazer as análises em cima do engajamento das comunidades, como o total de inscritos, usuários online, visualizar comentários

recentes, principais *flairs* utilizados, além de poder fazer a análise temporal dos dados.

## 4 DESENVOLVIMENTO

No desenvolvimento da aplicação proposta para este projeto de TCC, optou-se por utilizar uma combinação de tecnologias robustas, eficientes e de fácil integração, além de dar preferência por tecnologias as quais o autor já teve experiências profissionais

### 4.1 Tecnologias utilizadas

Foram escolhidos ReactJS<sup>1</sup> para compor o frontend da aplicação, o NodeJS<sup>2</sup> para compor o backend e MongoDB<sup>3</sup> para o banco de dados, fazendo utilização do Javascript tanto no frontend quanto no backend da aplicação.

O ReactJS será utilizado para criar a interface de usuário, permitindo visualizar os dados coletados da plataforma do Reddit. O NodeJS será utilizado para criar um servidor que se conectará à API do Reddit e também à aplicação ReactJS. Já o MongoDB será responsável por armazenar os dados coletados, permitindo que sejam consultados posteriormente pelo backend em NodeJS, que por fim irá entregar as informações ao frontend em ReactJS, permitindo a análise do usuário sobre a popularidade das linguagens de programação da plataforma do Reddit.

#### 4.1.1 React

React ([www.react.dev](http://www.react.dev)) é uma biblioteca JavaScript de código aberto desenvolvida pelo Facebook, utilizada para criar interfaces de usuário interativas de forma rápida e eficiente. Com o React é possível desenvolver seus aplicativos criando componentes de interface reutilizáveis, desta forma, em vez de lidar com toda a interface do usuário como uma unidade única, o React fornece aos desenvolvedores a facilidade de separar interfaces de usuário complexas em componentes reutilizáveis.

Desta forma, o React combina a velocidade e a eficiência do JavaScript com a sua tecnologia que tornou extremamente eficiente a manipulação do DOM, renderizando páginas Web mais rapidamente e de forma extremamente dinâmica.

<sup>1</sup><https://react.dev/>

<sup>2</sup><https://nodejs.org/>

<sup>3</sup><https://mongodb.com/>

### **4.1.2 Node.js**

Node.js ([www.nodejs.org](http://www.nodejs.org)) é um ambiente de tempo de execução de código aberto para Javascript, que permite aos desenvolvedores trabalharem com Javascript não apenas no lado do cliente, mas também no lado do servidor. O Node.js permite criar aplicativos com uso intensivo de dados e é ideal para construção de aplicações que precisam de interação em tempo real.

O Node.js é construído em cima do motor JavaScript V8 do Google Chrome, utilizando um modelo de programação assíncrono e orientado a eventos, sendo assim, é ideal para o desenvolvimento de aplicativos que exigem alta performance e escalabilidade.

### **4.1.3 Express.js**

O Express.js ([www.expressjs.com](http://www.expressjs.com)) é um renomado framework para desenvolvimento de aplicações web utilizando JavaScript, se destacando pela sua simplicidade, desempenho e flexibilidade. Possui uma estrutura de roteamento flexível, que permite criar rotas personalizadas e direcionar solicitações HTTP para funções específicas.

Uma das principais características do Express.js é a sua abordagem modular oferecendo uma arquitetura flexível que permite a inclusão de funcionalidades adicionais através de plugins e middlewares.

### **4.1.4 MongoDB**

MongoDB ([www.mongodb.com](http://www.mongodb.com)) é um sistema de gerenciamento de banco de dados NoSQL, de código aberto, que pode gerenciar informações orientadas a documentos. NoSQL é uma alternativa aos bancos de dados relacionais tradicionais, pois são bastante úteis quando se precisa trabalhar com um grande volume de dados. É uma alternativa flexível e escalável aos bancos de dados relacionais tradicionais. Ao contrário dos bancos de dados relacionais, o MongoDB não utiliza tabelas e linhas e não requer uma estrutura de dados predefinida, permitindo que os usuários adicionem ou removam campos de um documento sem precisar modificar a estrutura da coleção.

Como em qualquer banco de dados NoSQL, possui uma arquitetura composta

por coleções e documentos. Os documentos são compostos de pares chave/valor, sendo a unidade básica de dados do MongoDB. Já as coleções, seriam o equivalente às tabelas nos banco de dados SQL. Outro fator importante para escolha deste sistema de gerenciamento de banco de dados é o fato de existirem muitas ferramentas e bibliotecas disponíveis para ajudar os usuários a trabalhar com o banco de dados de maneira eficiente e eficaz.

#### **4.1.5 API do React**

A API do React (<https://react.dev/reference/react/apis>) é composta por um conjunto de componentes e métodos que permitem criar e manipular elementos de interfaces de usuário declarativas, reutilizáveis e eficientes. O principal conceito por trás da API é permitir que os desenvolvedores escrevam a interface do usuário com a utilização de componentes reutilizáveis, que podem ser desenvolvidos em hierarquias para criar interfaces mais complexas.

O coração da API do React são os componentes, os quais podem ser criados de duas maneiras, como componentes de classe ou como função. Os componentes de classe estendem a classe `React.Component` e devem implementar o método `render()` do React, retornando uma descrição em forma de árvore dos elementos que compõem o componente. Já os componentes de função, por outro lado, são funções puras que recebem as propriedades (`props`) do componente como argumentos, para então retornarem uma descrição em forma de árvore dos elementos que compõem o componente.

#### **4.1.6 Axios**

O Axios ([www.axios-http.com](http://www.axios-http.com)) é uma biblioteca JavaScript utilizada para realizar requisições HTTP, podendo ser utilizada tanto por aplicações no front-end, quanto no back-end. Um dos principais benefícios do Axios é a sua simplicidade de configuração e de uso para executar os diferentes tipos de requisições HTTP, GET, POST, PUT, DELETE e HEAD. As requisições do Axios são baseadas em promessas do JavaScript, o que permite trabalhar com os resultados das requisições de forma síncrona e assíncrona.

## 4.2 Banco de dados NoSQL

Na definição da estrutura do banco de dados, optou-se por armazenar todas as informações em uma única coleção, permitindo com que as consultas sejam feitas de forma extremamente eficiente, sem a necessidade de fazer agregações ou acessar várias coleções em uma mesma consulta, reduzindo assim a complexidade e melhorando o desempenho das consultas, especialmente quando se tem um grande volume de dados coletados.

Os dados armazenados pela coleção são descritos abaixo:

1. Identificador único da coleção coletada no dia;
2. Data e hora da coleta;
3. Um array de subreddits responsável por armazenar a listagem de subreddits coletados no dia e suas informações relevantes para esta aplicação, conforme os campos descritos abaixo:
  - 3.1 ID único do subreddit;
  - 3.2 Nome de exibição do subreddit;
  - 3.3 URL da imagem do cabeçalho do subreddit;
  - 3.4 Título do subreddit;
  - 3.5 Número de usuários ativos no subreddit;
  - 3.6 Nome do subreddit formatado;
  - 3.7 Número de contas ativas no subreddit;
  - 3.8 Número de inscritos no subreddit;
  - 3.8 Descrição pública do subreddit;
  - 3.9 URL do ícone.
4. Um array com a amostra de postagens coletadas, armazenando as seguintes informações:
  - 4.1 Nome do subreddit da postagem;
  - 4.2 Nome completo do autor da postagem;
  - 4.3 Título da postagem;
  - 4.5 Nome do subreddit formatado;
  - 4.6 Número de votos negativos da postagem;

- 4.7 Nome da postagem;
- 4.8 Taxa de votos positivos da postagem;
- 4.9 Tipo do subreddit;
- 4.10 Número de votos positivos da postagem;
- 4.11 Total de prêmios recebidos pela postagem;
- 4.12 Pontuação da postagem;
- 4.13 ID do subreddit;
- 4.14 Número de comentários da postagem;
- 4.15 URL da postagem;
- 4.16 Número de inscritos no subreddit.

Figura 11 – Banco de dados NoSQL

```

{
  "_id": ObjectId,
  "date": Date,
  "subreddits": [
    {
      "_id": ObjectId,
      "display_name": String,
      "header_img": String,
      "title": String,
      "active_user_count": Number,
      "display_name_prefixed": String,
      "accounts_active": Number,
      "subscribers": Number,
      "user_flair_richtext": Array,
      "public_description": String,
      "community_icon": String,
    },
  ],
  "posts": [
    {
      "subreddit": String,
      "author_fullname": String,
      "title": String,
      "link_flair_richtext": Array,
      "subreddit_name_prefixed": String,
      "downs": Number,
      "name": String,
      "upvote_ratio": Number,
      "ups": Number,
      "total_awards_received": Number,
      "user_reports": Array,
      "score": Number,
      "author_flair_richtext": Array,
      "subreddit_id": String,
      "num_comments": Number,
      "url": String,
      "subreddit_subscribers": Number,
    },
  ],
}

```

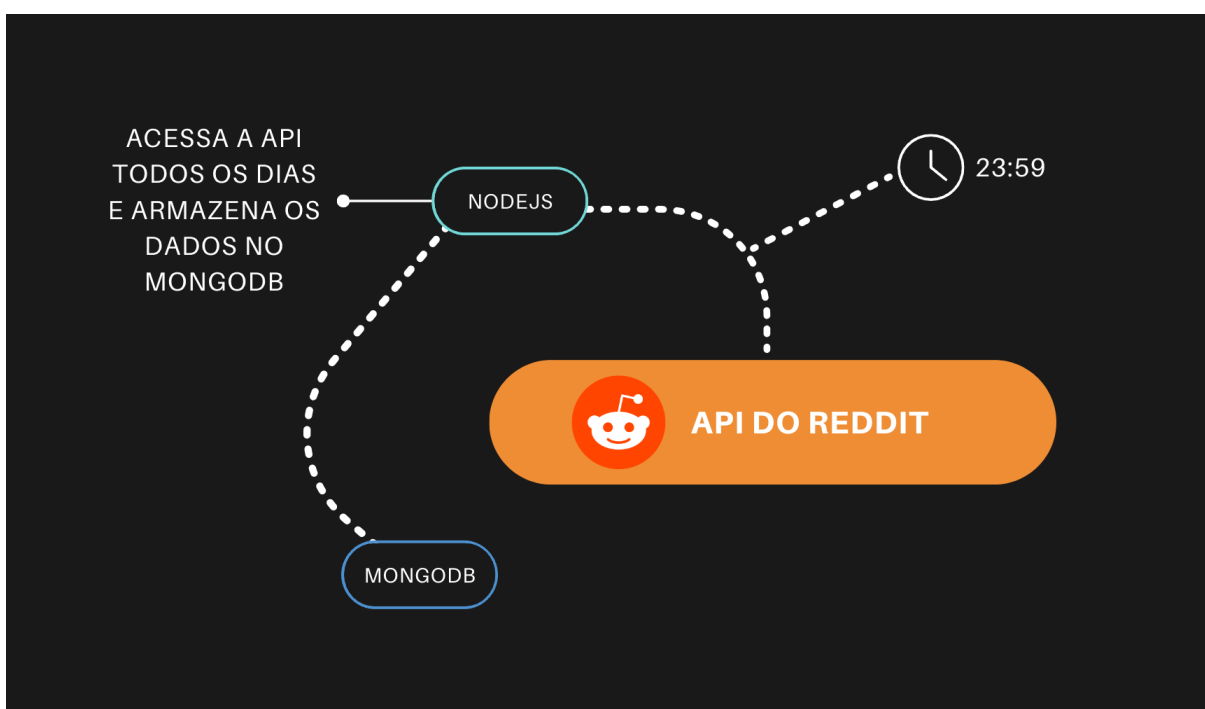
Fonte: Autor



### 4.3 Desenvolvimento do software

O desenvolvimento da aplicação web iniciou com a etapa de configuração do ambiente de desenvolvimento, que incluiu a instalação e configuração das ferramentas essenciais ao projeto e definidas em escopo, o Node.js, a biblioteca ReactJS e o MongoDB. Em seguida, foi criada uma nova aplicação Node.js utilizando o framework Express.js, dando início ao desenvolvimento do backend da aplicação. A solução foi modelada para que o backend faça as requisições para API do Reddit uma vez ao dia e armazene os dados no banco de dados MongoDB para uma análise de evolução temporal dos dados. Dessa forma foi criado um script que é executado todos os dias no último minuto de cada dia, armazenando assim os dados necessários sobre as comunidades analisadas.

Figura 12 – Consultas a API do Reddit

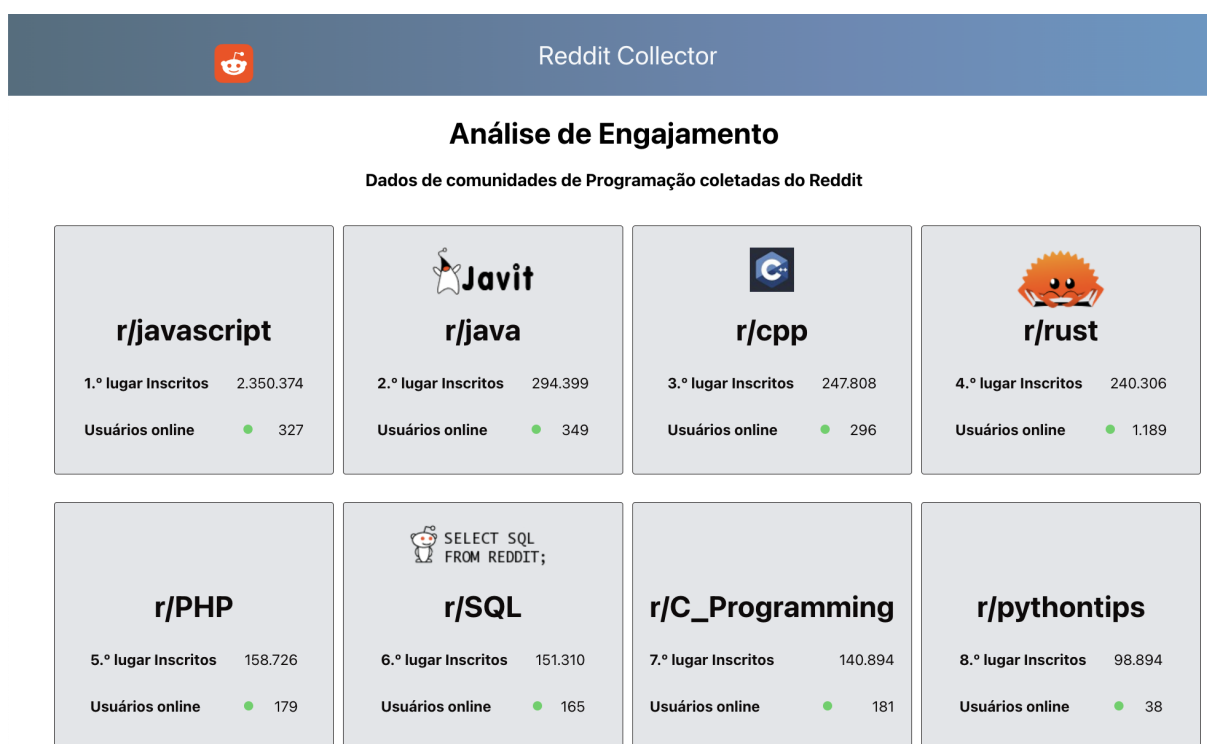


Fonte: Autor

Desta forma são coletados os dados das principais comunidades públicas de linguagens de programação da rede social, com postagens de usuários de todo o mundo, em diversos idiomas.

Durante os testes realizados com a API notou-se a limitação nas requisições por postagens, existindo um limite de até 100 postagens por requisição, conforme indicado na documentação da plataforma. Sendo assim, o sistema implementado para este projeto trabalha com uma amostra das últimas 100 postagens por comunidade de linguagem de programação analisada, onde também existe a coleta de dados sobre as comunidades, como número de inscritos, usuários online, flairs utilizados, postagens e engajamento.

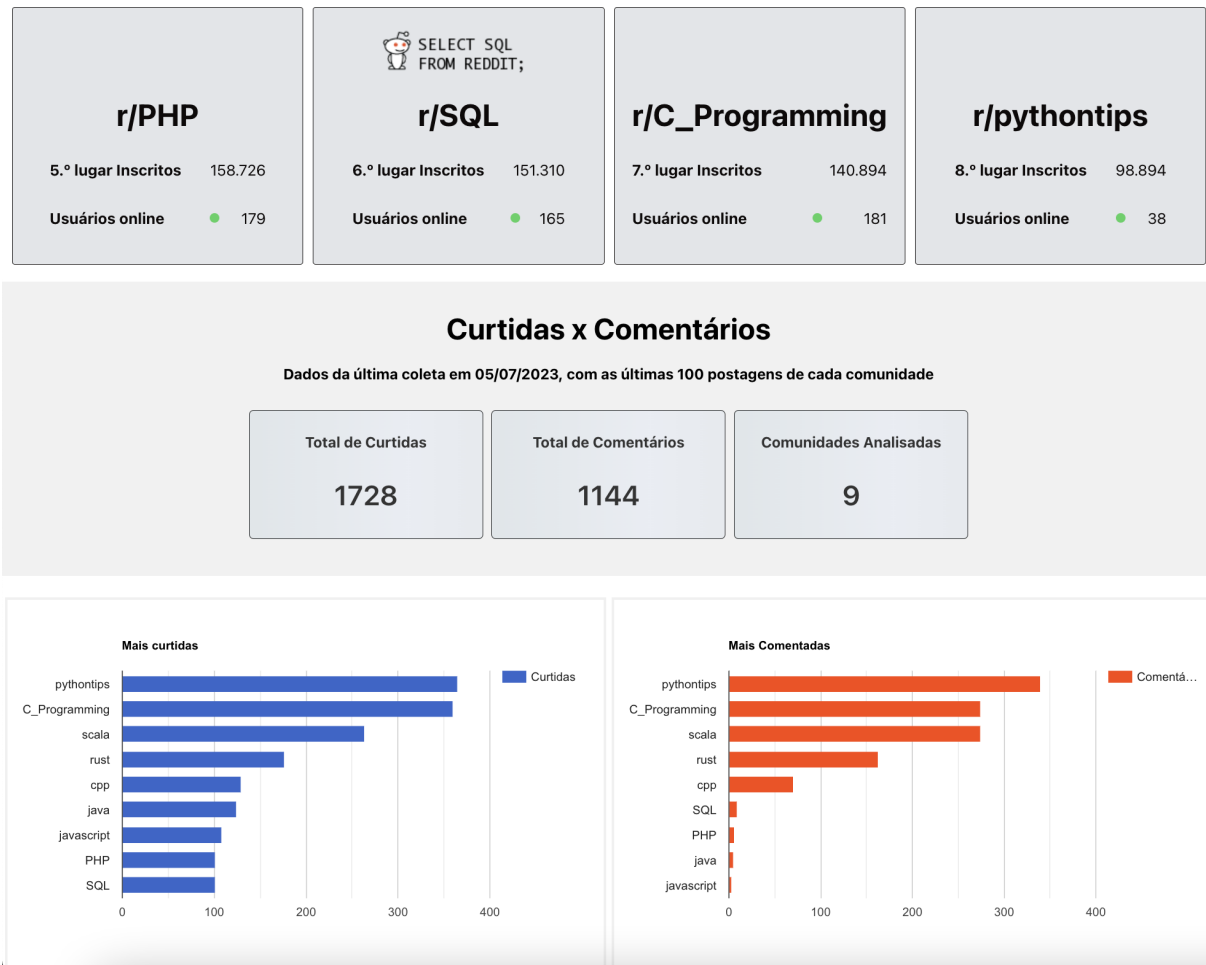
Figura 13 – Tela de análise de comunidades



Fonte: Autor

Para desenvolver o frontend, utilizou-se a biblioteca ReactJS, permitindo a integração com a API desenvolvida e disponibilizada pelo backend. A biblioteca permitiu a criação da interface da aplicação, permitindo a interação do usuário com o ranking de comunidades, métricas de engajamento e visualização dos detalhes de cada comunidade com a visualização das últimas postagens e análise de *flairs*.

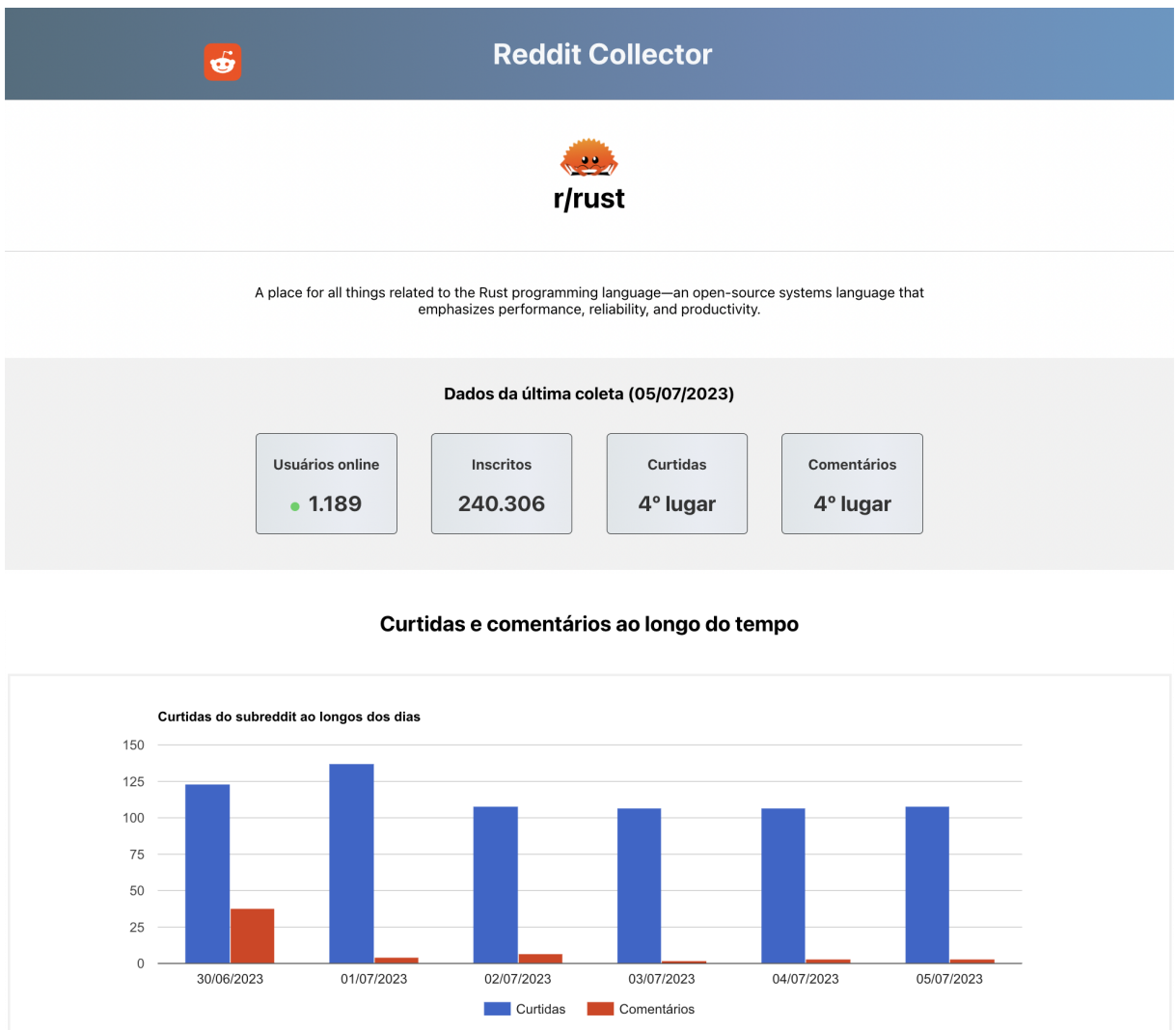
Figura 14 – Tela de análise de engajamento



Fonte: Autor

O sistema também gera gráficos a partir do engajamento dos usuários, em cima das quantidades de curtidas e comentários. O gráfico é construído de acordo com a amostra coletada de cada linguagem de programação, conforme pode ser visualizado na figura 14. No caso, por conta das limitações impostas pela API do Reddit, a amostra é limitada a 100 registros por consulta, onde optou-se para este trabalho a consulta das últimas 100 postagens até o momento da requisição diária.

Figura 15 – Detalhes da comunidade



Fonte: Autor

Na tela de detalhes da comunidade indicada na figura 15, é possível obter mais informações da mesma, com as descrições e objetivos da comunidade, alguns indicadores, visualização das postagens podendo ser filtradas por mais curtidas e maior score.

Figura 16 – Análise das últimas postagens

**Postagens da última coleta em 05/07/2023**

FILTRAR POSTAGENS

Discussion FOR HIRE - Programmer Programming Offer Suggestion Question Humor / Meme Image/Album Self Promotion  
Academics Gone Wild Tech Support Profile Review Form check Question Open | Windows Software Help Question | Help  
CS/EECS SQL Server C# Question Other Help Limited Time Limited Time Limited Time match-1 SCREENSHOT  
Pytania i Dyskusje Future Tiger Question Book Club Advice Needed! Politics Limited Time Limited Time Question | Help  
Question FOR HIRE - Programmer Question News Limited Time Verified Chemistry and Computer Science Expert  
Limited Time

**Autor:**

**Título:** [FOR HIRE][Paid] 4 Years EXP - Unity Developer LF short/long term projects remote work

**Texto da postagem:**

My name is Mario I have been A Game Developer creating mobile and desktop games for 4 years. I provide full game and app development including design and publishing to Apple Store and Google Play store. Platforms: Apple app store / Google Play Store / Web / PC(steam)  
CONTACT Portfolio: [check all my projects here](https://mario-gamedev.com/) Discord: silver91#3662 My skills: • 4 years of experience in Unity • Unity 2D / 3D / C# / • I got experience with making mobile games - implementing ads and publishing them to the play store • NPC Design Programming • User Interface • Unity Services Implementation • Prototyping (Core Gameplay for Rapid Iteration • Performance Optimization and Target Platforms • Unity Services Implementation • Git • Unit testing • I worked as a lead gameplay programmer on a prototype Dinosaur Survival Rpg game for Norian Games(you can dm me if you want to know more about the project) • Worked on online games using Mirror & Photon(Quantum) • Sport Games(Lacrosse game) • Open World FPS Game Specialties: Game Programming: • Gameplay Programming • AI Programming • Tools My hourly rate: • 20 \$ / hour

**Gostei: 1 Não gostei: 0**

[Ver postagem na íntegra](#)

**Autor:**

Fonte: Autor

Na figura 16 é possível visualizar melhor as postagens da tela de detalhes da comunidade, onde é possível fazer o filtro tanto por texto quanto por *flair* listados abaixo do *input* para filtrar as postagens. Ao clicar no *flair*, irá filtrar para que apareçam apenas as postagens que inseriram o *flair* selecionado.

#### 4.3.1 Coleta diária dos dados

O código abaixo utiliza o método *cron* do Javascript, para executar o método de coleta todo dia às 23:59. O método de coleta em questão faz a requisição GET utilizando a biblioteca Axios, que é um cliente HTTP baseado em promessa, que no caso irá executar a chamada para API em questão, nos retornando sucesso da promessa, ou não. Em caso de sucesso, os dados são armazenados no banco de dados da aplicação.

Figura 17 – Script de coleta diária dos dados

```
const scheduledFunction = async () => {
  console.log('Executando função diária de coleta dos dados...');

  let data = { ...
}

let subreddits = [ ...
]

for (let index = 0; index < subreddits.length; index++) { ...
};

const newData = new Data(data);

try {
  const savedData = await newData.save();
  console.log('Data inserted:', savedData);
} catch (error) {
  console.error('Error saving data:', error);
}
};

cron.schedule('23 59 * * *', scheduledFunction);
```

You, 1 second ago • Uncommitted changes

Fonte: Autor

## **5 CONCLUSÃO**

Com o desenvolvimento deste trabalho, concluiu-se que a aplicação desenvolvida torna-se mais uma ferramenta que pode auxiliar usuários que desejam analisar e fazer comparações entre as principais linguagens de programação do mercado, no que diz respeito ao seu engajamento, popularidade e assuntos recentes dentro de uma comunidade online da rede social Reddit.

A plataforma permite que os usuários analisem os números relevantes de diferentes comunidades de programação ao longo do tempo, além de poder medir o engajamento destas comunidades dentro da plataforma, como número de inscritos, usuários online, curtidas, comentários e análise de flairs. O usuário também tem a opção de visualizar e filtrar entre os comentários coletados de cada comunidade, dessa forma, é possível buscar por assuntos e temas que desejam que possam estar sendo debatidos por outros usuários dentro da comunidade de programação de seu interesse.

Analisando os resultados dos gráficos de engajamento gerados ao longo dos dias de coleta, verificou-se que os resultados normalmente mantiveram as comunidades relacionadas às linguagens PYTHON, JAVA, C, C++ e C# nas primeiras colocações. Também vale ressaltar que em nenhum dos casos o Javascript esteve no topo como indicado na pesquisa do Stack Overflow.

Vale ressaltar também as dificuldades encontradas com as constantes mudanças nas regras das APIs fornecidas pelas plataformas Twitter e Reddit. Primeiramente com o Twitter, onde o consumo de certos recursos da API que seriam necessários para este trabalho passaram a ser pagos, o que fez com que fosse feita a migração para coletar dados da plataforma Reddit. Que também passou por mudanças nas regras para consumo da API ao longo do desenvolvimento deste projeto.

### **5.1 TRABALHOS FUTUROS**

Para trabalhos futuros, há várias melhorias e possibilidades de expansão da plataforma, como por exemplo a análise de outras comunidades online além do Reddit como Stack Overflow e GitHub, proporcionando uma visão mais abrangente

do engajamento e popularidade das linguagens em diferentes contextos de comunidades online.

Outra funcionalidade interessante seria a análise de sentimento e opiniões. Isso envolveria técnicas de processamento de linguagem natural para identificar a polaridade dos comentários e avaliar a opinião dos usuários em relação a determinadas linguagens ou tópicos. Recursos de visualização aprimorados também seria de grande valor, pois torna a plataforma ainda mais amigável e intuitiva.

Ou seja, mesmo cumprindo com o objetivo proposto no início do trabalho, a plataforma desenvolvida possui pontos que podem ser melhorados para proporcionar uma melhor análise das informações e uma melhor usabilidade da aplicação por parte do usuário.



## REFERÊNCIAS

VASILESCU, Bogdan. **Human Aspects, Gamification, and Social Media in Collaborative Software Engineering**. 2014. 4 f. Tese (Doutorado) - Curso de Computer Science, Department Of Mathematics And Computer Science, Eindhoven University Of Technology, Eindhoven, 2014.

TIOBE. TIOBE Index for Dec 2022. 2022. Disponível em: <https://www.tiobe.com/tiobeindex/> Acesso em: 05 dec 2022.

STACK OVERFLOW. **Developer Survey**: 2021. New York: Stack Overflow, 2021.

WIKIPEDIA. WIKIPEDIA for Dec 2022. 2022. Disponível em: <https://en.wikipedia.org/wiki/JavaScript> Acesso em: 15 nov 2022.

REACT. **Getting Started**. Disponível em: <https://reactjs.org/docs/getting-started.html>. Acesso em: 01 abr. 2023.

W3SCHOOLS. **React Tutorial**. Disponível em: <https://www.w3schools.com/react/>. Acesso em: 01 abr. 2023.

VASILESCU, Bogdan. **Human Aspects, Gamification, and Social Media in Collaborative Software Engineering**. 2014. 4 f. Tese (Doutorado) - Curso de Computer Science, Department Of Mathematics And Computer Science, Eindhoven University Of Technology, Eindhoven, 2014.

ZAPALOWSKI, Vanius. **Análise quantitativa e comparativa de linguagens de programação**. 2011. 45 f. Tese (Doutorado) - Curso de Ciência da Computação, Instituto de Informática, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2011.

FREECODECAMP. **React**. Disponível em: <https://www.freecodecamp.org/learn/front-end-libraries/react/>. Acesso em: 01 abr. 2023.

STACK OVERFLOW. **Perguntas com a tag reactjs**. Disponível em: <https://stackoverflow.com/questions/tagged/reactjs>. Acesso em: 02 abr. 2023.

NODEJS. **Documentação oficial do Node.js**. Disponível em:

<https://nodejs.org/en/docs/>. Acesso em: 03 mar. 2023.

Wikipedia. Reddit. Wikimedia Foundation, Inc. Disponível em: <https://en.wikipedia.org/wiki/Reddit>. Acesso em: 06 abr. 2023.

Lifewire. **What is Reddit and How Does It Work?** Disponível em: <https://www.lifewire.com/what-is-reddit-4163576>. Acesso em: 06 abr. 2023.

JURASK, Gabriel. **UM ESTUDO SOBRE O USO DE LINGUAGENS DE PROGRAMAÇÃO E SOFTWARES UTILIZADOS NA INDÚSTRIA DE JOINVILLE E REGIÃO E SUA RELAÇÃO COM O PERFIL DO EGRESSO DO BACHARELADO EM CIÊNCIA E TECNOLOGIA DA UFSC CAMPUS JOINVILLE**. 2019. 27 f. TCC (Graduação) - Curso de Ciência e Tecnologia, Centro Tecnológico, Universidade Federal de Santa Catarina, Joinville, 2019.

W3SCHOOLS. **Node.js Tutorial**. Disponível em:

<https://www.w3schools.com/nodejs/default.asp>. Acesso em: 04 de abril de 2023.

MONGODB. **Documentação oficial do MongoDB**. Disponível em: <https://docs.mongodb.com/>. Acesso em: 05 de abril de 2023.

TUTORIALS POINT. **MongoDB Tutorial**. Disponível em: <https://www.tutorialspoint.com/mongodb/index.htm>. Acesso em: 05 de abril de 2023.

UML. **ABOUT THE UNIFIED MODELING LANGUAGE SPECIFICATION VERSION 2.5.1**. Disponível em: <https://www.omg.org/spec/UML/2.5.1/PDF>. Acesso em: 11 de julho de 2023.

## REPOSITÓRIO DO CÓDIGO

O código fonte da aplicação está disponível no repositório no GitLab: <https://codigos.ufsc.br/marcos.demetrio/tcc>. Contendo o projeto *backend* e *frontend* da aplicação e as instruções necessárias para rodar o projeto.

## APÊNDICE A

# Coleta de dados do Reddit para análise das linguagens de programação mais populares do mercado\*

Marcos Demétrio<sup>1</sup>, José Eduardo De Lucca<sup>2</sup>

<sup>1</sup>Departamento de Informática e estatística  
Universidade Federal de Santa Catarina (UFSC) – Florianópolis, SC - Brasil

marcosdemetrio@gmail.com

***Abstract.** This article summarizes the course completion work, for the Information Systems course at the Federal University of Santa Catarina, entitled Reddit data collection for analysis of the most popular programming languages on the market.*

***Resumo.** Este artigo resume o trabalho de conclusão de curso, para o curso de Sistemas de Informação pela Universidade Federal de Santa Catarina, intitulado Coleta de dados do Reddit para análise das linguagens de programação mais populares do mercado.*

## 1. Introdução

Os desenvolvedores têm várias opções para se conectarem e trocarem informações, como redes sociais e fóruns na internet. Um desses ambientes é o Reddit, uma rede social que também funciona como um agregador de notícias onde os usuários podem compartilhar e discutir uma ampla variedade de tópicos, incluindo programação e desenvolvimento de software.

Essas comunidades online oferecem uma forma para os desenvolvedores compartilharem conhecimentos e experiências, podendo fazer perguntas, responder dúvidas de outros usuários e participar de discussões sobre tópicos relacionados à programação. Além disso, essas comunidades são um excelente recurso para os desenvolvedores se manterem atualizados sobre as últimas tendências na área da programação, onde os usuários compartilham notícias, artigos, tutoriais e recursos úteis para os desenvolvedores.

\* Trabalho de Conclusão de Curso – TCC, do curso de Sistemas de Informação, UFSC.

## **1.1 Objetivo**

O objetivo do trabalho foi desenvolver uma ferramenta que permita a análise das linguagens de programação mais populares do mercado. Para isso, utilizou-se da coleta de dados da plataforma do Reddit, permitindo a comparação entre a popularidade das comunidades de linguagens de programação presentes na plataforma, com indicadores dos engajamento de cada comunidade, incluindo comentários e curtidas ao longo do tempo.

Para conclusão do objetivo foi necessário levantar e analisar pesquisas sobre a popularidade das linguagens de programação, avaliar e selecionar uma plataforma relevante com discussões sobre linguagens de programação, comparar a popularidade de cada linguagem com base na frequência de postagens relacionadas a elas e construir uma interface interativa para visualizar os resultados.

## **2. Estudos Correlatos**

Neste capítulo, são apresentados dois estudos extremamente relevantes sobre a popularidade das linguagens de programação: a Pesquisa Anual de Desenvolvedores do Stack Overflow e o TIOBE Index. Ambos os estudos têm como objetivo avaliar a popularidade das linguagens de programação e fornecer informações sobre as tendências do mercado de tecnologia.

### **2.1 Pesquisas correlatas**

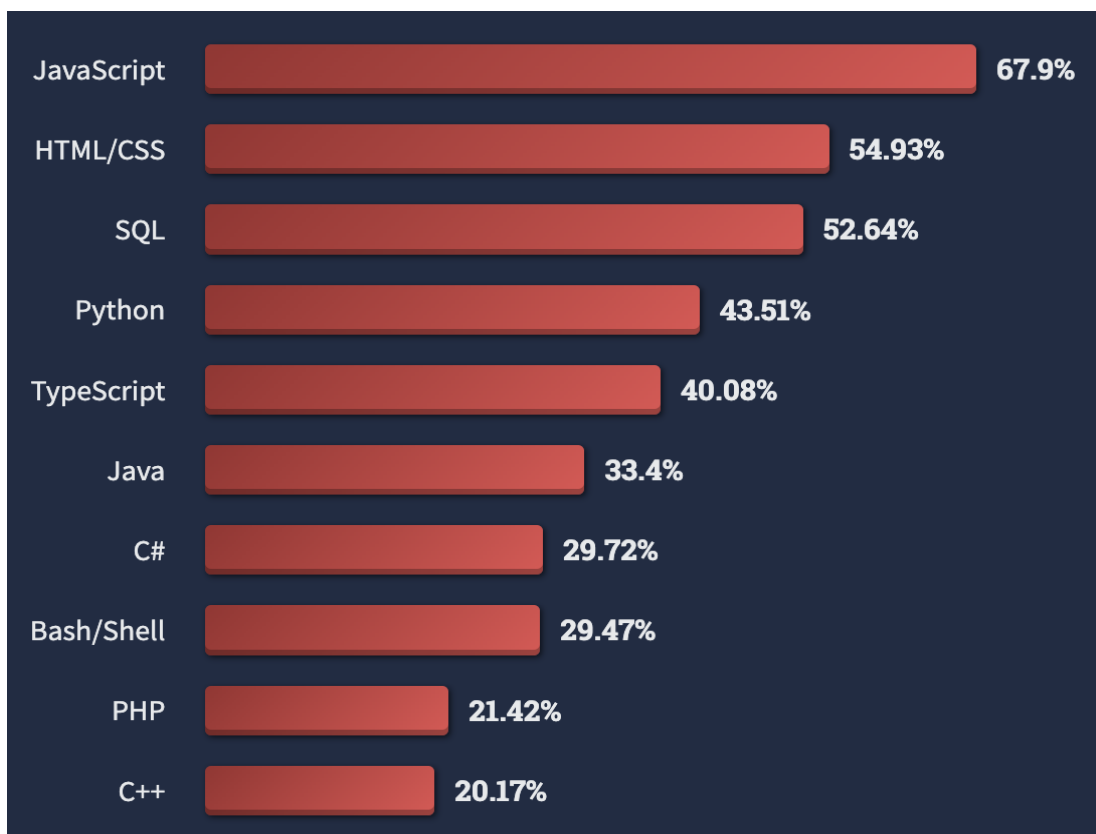
Este capítulo examina duas das pesquisas mais relevantes quando se trata em medir a popularidade das linguagens de programação, que são a Pesquisa Anual de Desenvolvedores do Stack Overflow e a pesquisa TIOBE Index.

#### **2.1.1 Pesquisa Anual de Desenvolvedores do Stack Overflow**

A plataforma Stack Overflow é um site popular de perguntas e respostas na área de tecnologia e recebe milhões de usuários mensalmente. A Pesquisa Anual de Desenvolvedores do Stack Overflow é realizada desde 2011 e analisa diferentes aspectos das preferências dos desenvolvedores, onde em 2022 a pesquisa recebeu mais de 70 mil respostas de desenvolvedores de mais de 180 países. Os países com maior número de respostas foram Estados Unidos, Índia, Alemanha, Reino Unido, Canadá, França, Brasil, Polônia, Países Baixos e Itália.

Os Estados Unidos lideram a lista com mais de 9.000 desenvolvedores participantes, seguido pela Índia, em segundo lugar com mais de 4.500 participantes. O Brasil aparece em sétimo lugar, com mais de 1400 participantes.

Na Pesquisa Anual de Desenvolvedores do Stack Overflow de 2022, o Javascript foi classificado como a linguagem de programação mais utilizada pelos desenvolvedores profissionais, mantendo essa posição ao longo dos últimos 10 anos, refletindo a ampla utilidade do Javascript em diversos tipos de projetos de desenvolvimento de software.



**Figura 1. Ranking das linguagens da Pesquisa Anual do Stack Overflow.**

### 2.1.2 Tiobe Index

O índice TIOBE é um índice de popularidade das linguagens de programação de acordo com o número de pesquisas relacionadas a cada linguagem de programação nos motores de busca como Google, Bing, Yahoo!, Wikipedia, Amazon, YouTube, entre outros. O índice abrange centenas de linguagens de programação e fornece uma visão abrangente das tendências da indústria de tecnologia.

Programming Language	2022	2017	2012	2007	2002	1997	1992	1987
Python	1	5	8	7	13	28	17	-
C	2	2	1	2	2	1	1	1
Java	3	1	2	1	1	16	-	-
C++	4	3	3	3	3	2	2	5
C#	5	4	5	8	12	-	-	-
Visual Basic	6	16	-	-	-	-	-	-
JavaScript	7	7	10	9	8	21	-	-
Assembly language	8	12	-	-	-	-	-	-
SQL	9	-	-	-	7	-	-	-
PHP	10	8	6	5	6	-	-	-

**Figura 2. Ranking das linguagens do Tiobe Index.**

Na figura 4, é apresentado o ranking das linguagens de programação de acordo com o TIOBE Index. O TIOBE Index é uma pesquisa realizada desde 2001 que classifica as linguagens de programação com base no total de buscas realizadas em 25 motores de busca, conseguindo divulgar os resultados pegando também anos anteriores a sua fundação, coletando resultados que vão até 1987.

## 2.2 Redes sociais com concentração de desenvolvedores

### 2.2.1 Reddit

O Reddit é uma rede social que agrega notícias e fóruns de discussões na internet, contendo uma ampla variedade de tópicos, permitindo aos usuários da plataforma compartilharem conteúdos e participarem de fóruns de discussões. Os tópicos são diversos, entre uma ampla gama de temas como tecnologia, notícias, política, ciência, entre outros. Os usuários podem avaliar as postagens dos outros usuários de forma positiva ou negativa, através de votação das postagens, que faz com que os usuários aumentem ou diminuam a relevância de uma postagem dentro da plataforma.

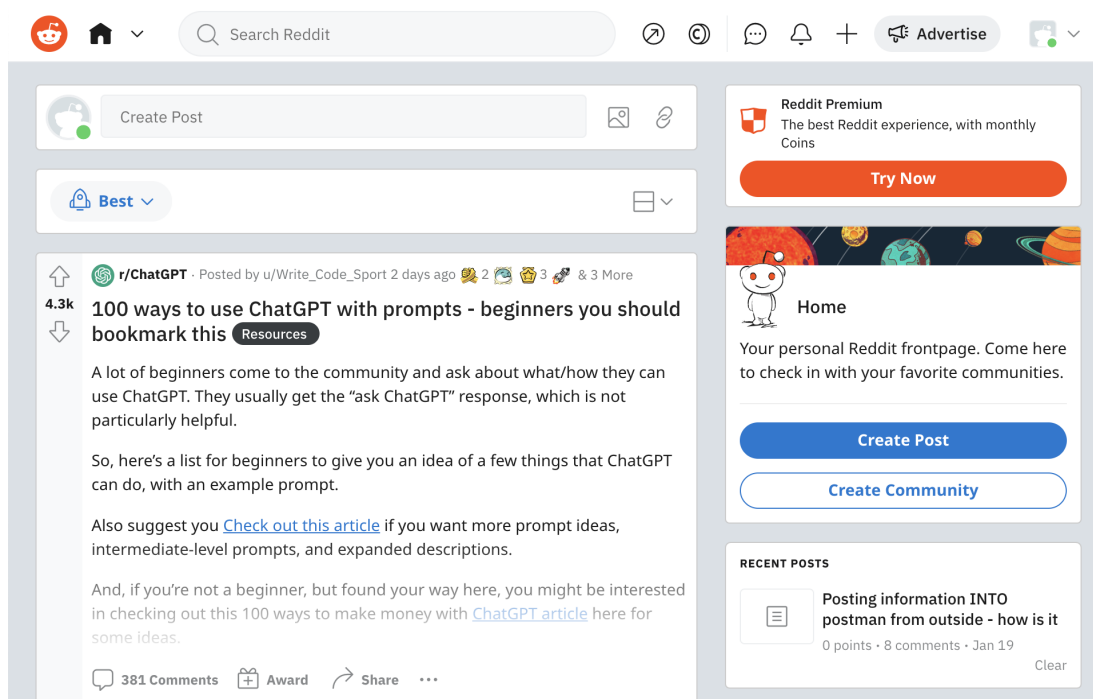


Figura 3. Plataforma do Reddit.

No Reddit, as postagens são organizadas em *subreddits*, fóruns temáticos criados pelos usuários. Os *subreddits* permitem que os usuários adicionem *tags* chamadas de "*flairs*" para categorizar o conteúdo. Essas *flairs* podem ser definidas pelos moderadores ou pelos próprios usuários, dependendo das configurações do *subreddit*. Qualquer usuário pode criar *subreddits*, resultando em uma ampla variedade de tópicos disponíveis na plataforma.

## 2.2.1 Twitter

O Twitter é uma rede social popular onde os usuários compartilham tweets, que são postagens contendo informações, notícias e opiniões em geral. Os usuários podem seguir outros perfis de usuário para acompanharem e serem notificados de suas postagens. No Twitter, é possível usar hashtags para categorizar e encontrar tweets relacionados a um determinado tema ou assunto.

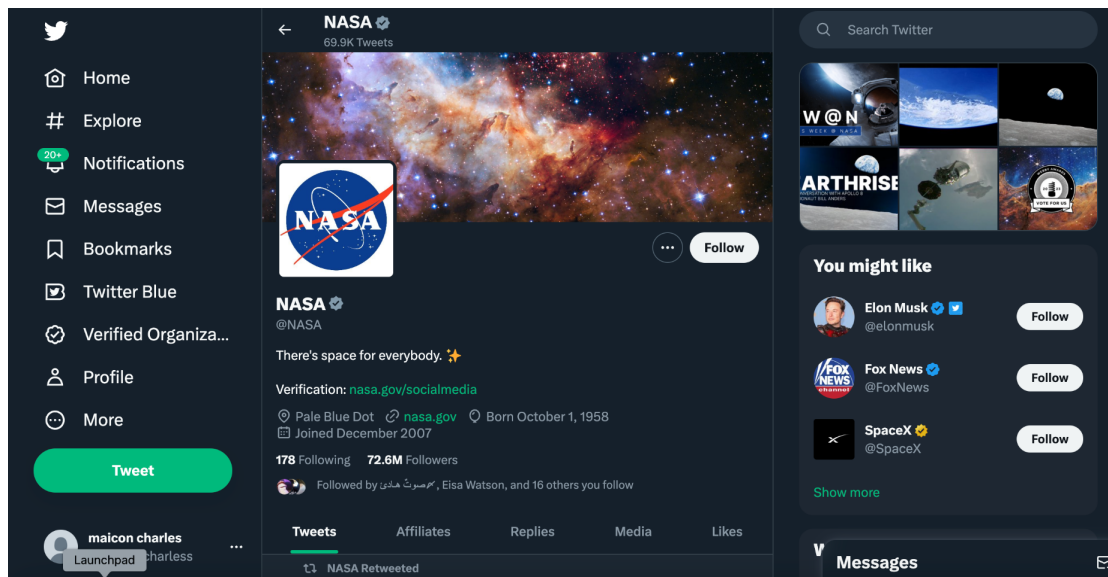


Figura 4. Plataforma do Twitter.

Apesar da plataforma do Twitter ser muito interessante para fazer a coleta das informações, sendo uma das maiores redes sociais do planeta, contendo um enorme número de usuários ativos, a sua nova gestão vem gerando dúvidas sobre a continuação das APIs de acesso gratuito.

Por tais fatores, o projeto optou por fazer coleta dos dados da plataforma do Reddit, por sua API ser totalmente gratuita e aberta para uso, permitindo fácil integração, sem custo algum. Além disso, o Reddit é uma plataforma muito popular e ativa, possuindo uma grande quantidade de desenvolvedores ativos, comunidades relacionadas às linguagens de programação analisadas, sendo assim, possuindo uma grande quantidade de dados disponíveis para análise.



### 3 Modelagem do Projeto

#### 3.1 Etapas do projeto

Adotou-se as seguintes etapas para conclusão do projeto: Levantamentos de requisitos, diagrama de caso de uso da aplicação, prototipação das telas e modelagem do banco de dados e etapa de desenvolvimento da aplicação.

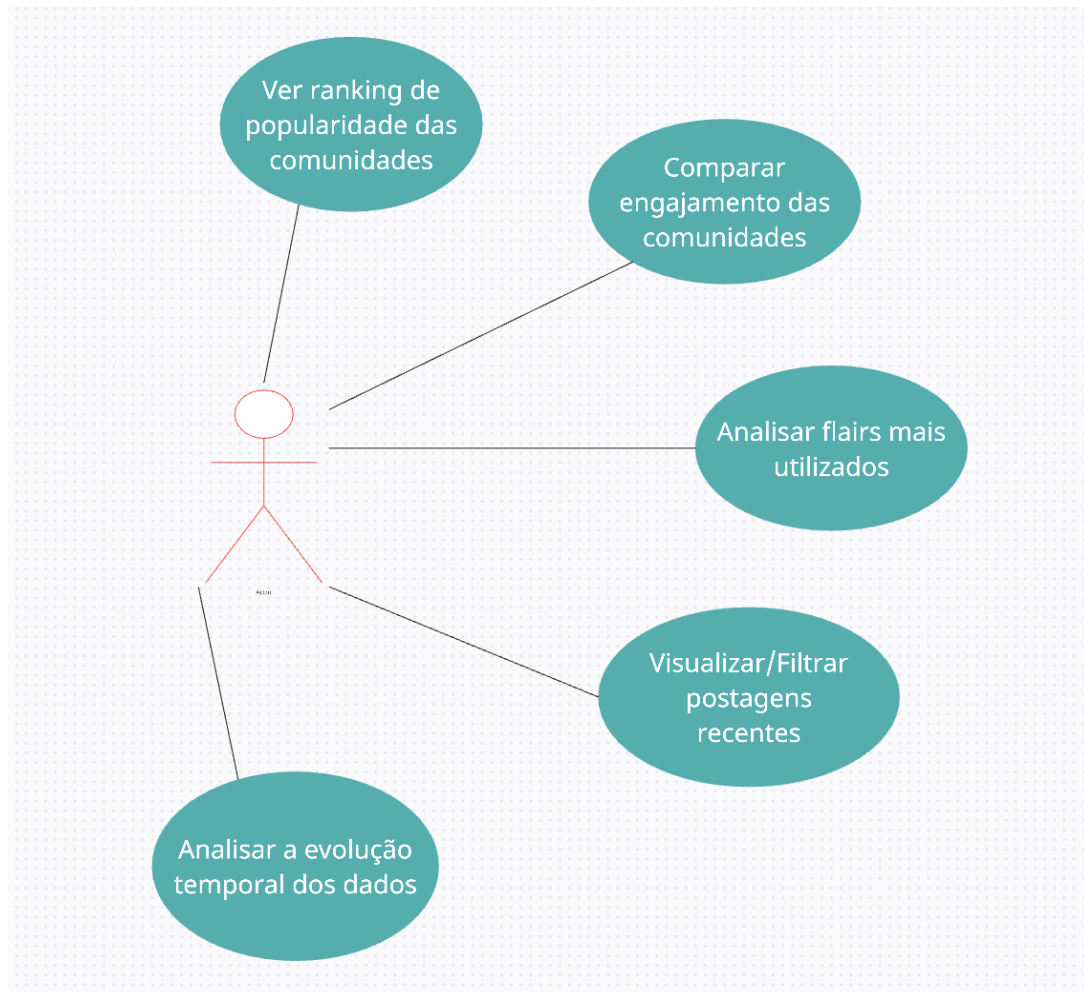


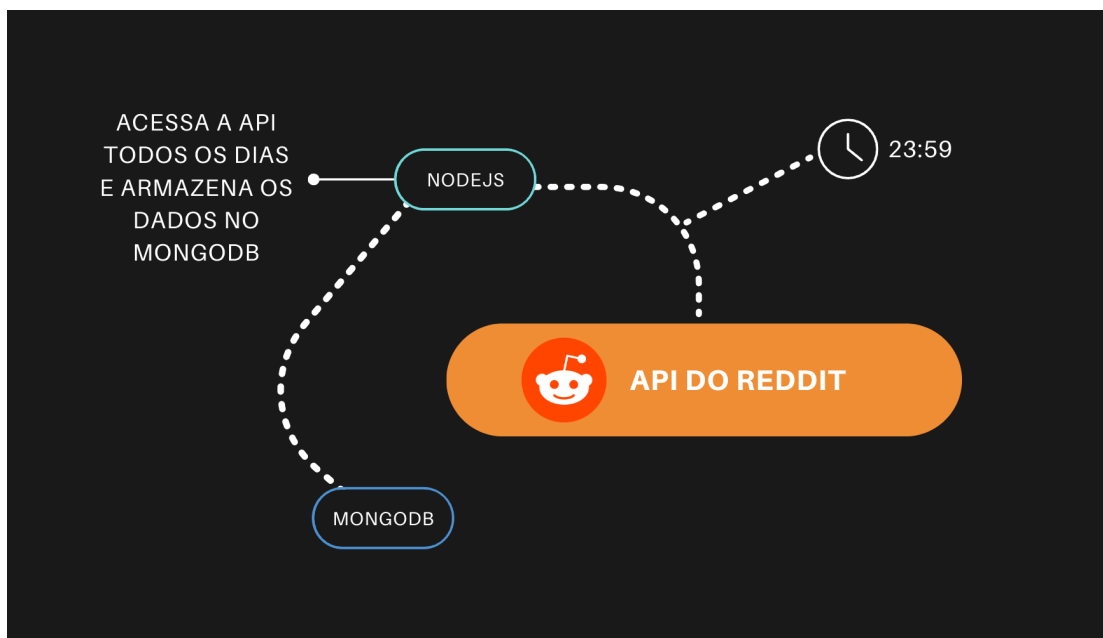
Figura 5. Diagrama de casos de uso da aplicação.

## 4 DESENVOLVIMENTO

Para desenvolvimento da aplicação, optou-se pelo ReactJS para compor o frontend da aplicação, o NodeJS para o backend e MongoDB para o banco de dados, fazendo uso do Javascript tanto frontend e no backend da aplicação.

### 4.1 Desenvolvimento do software

O desenvolvimento da aplicação web iniciou com a etapa de configuração das ferramentas essenciais ao projeto, o Node.js, a biblioteca ReactJS e o MongoDB. O backend faz as requisições para a API do Reddit uma vez ao dia e armazena os dados no banco de dados.



**Figura 6. Coleta diários dos dados.**

A plataforma coleta todos os dias dados das principais comunidades públicas relacionadas às linguagens de programação escolhidas.

A biblioteca do React permitiu a criação da interface da aplicação de maneira extremamente ágil, permitindo que o usuário interaja com a plataforma, fazendo a análise do ranking de comunidades, métricas de engajamento e visualização dos detalhes de cada comunidade, conforme as figuras 7 e 8, que contém a tela inicial e a tela de detalhes da comunidade.

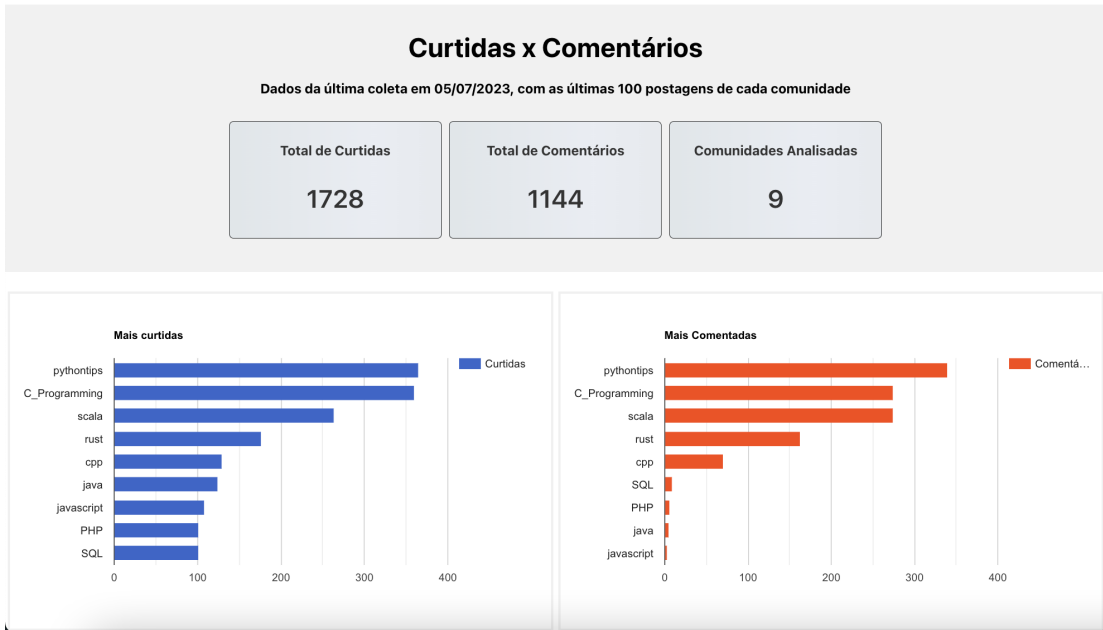


Figura 7. Tela de análise de engajamento.

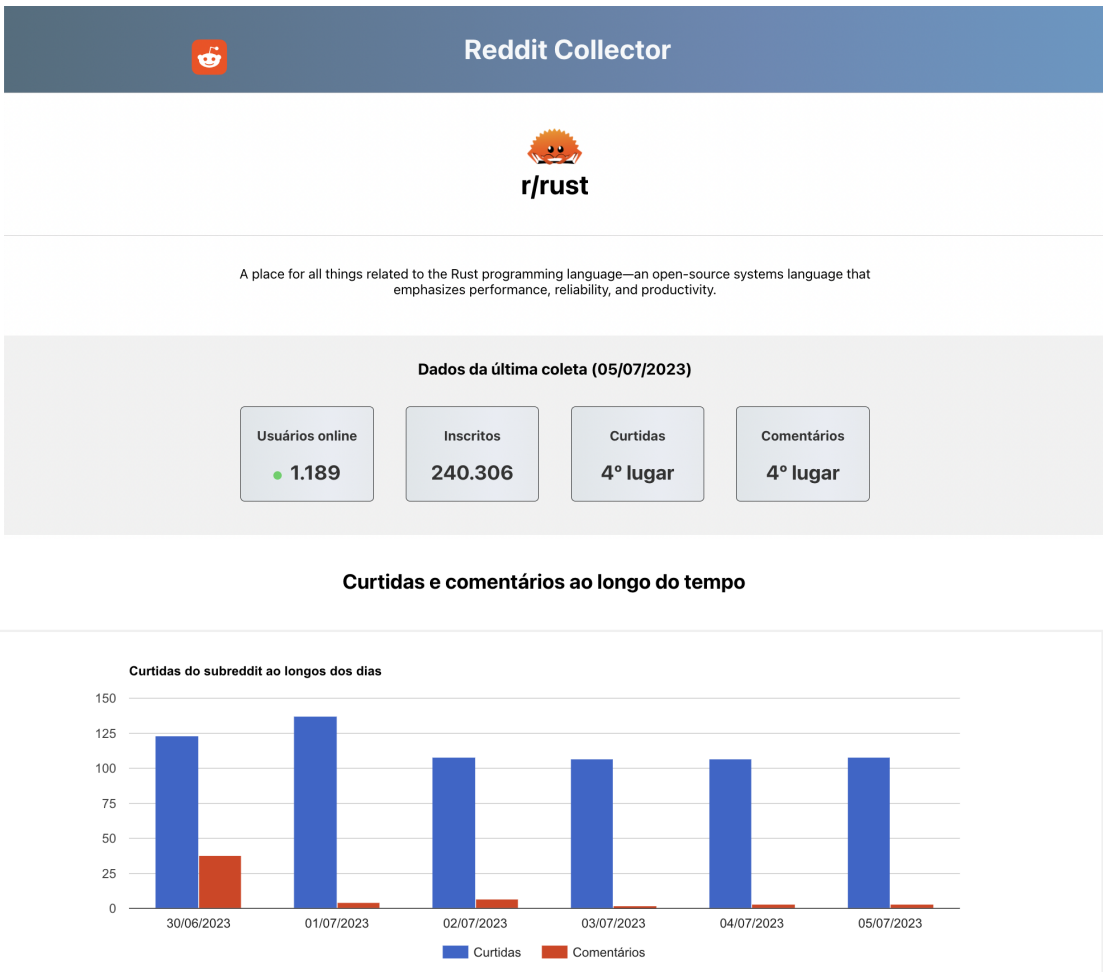


Figura 8. Tela de detalhes da comunidade.

## 5 Conclusão

A plataforma permite que os usuários analisem números e dados relevantes de diferentes comunidades de programação ao longo do tempo, além de poder medir o engajamento destas comunidades, como número de inscritos, usuários online, curtidas, comentários, *flairs* e postagens.

## Referências

TIOBE. TIOBE Index for Dec 2022. 2022. Disponível em: <https://www.tiobe.com/tiobeindex/> Acesso em: 05 dec 2022.

STACK OVERFLOW. Developer Survey: 2021. New York: Stack Overflow, 2021.

WIKIPEDIA. WIKIPEDIA for Dec 2022. 2022. Disponível em: <https://en.wikipedia.org/wiki/JavaScript> Acesso em: 15 nov 2022.

REACT. Getting Started. Disponível em: <https://reactjs.org/docs/getting-started.html>. Acesso em: 01 abr. 2023.

NODEJS. Documentação oficial do Node.js. Disponível em: <https://nodejs.org/en/docs/>. Acesso em: 03 mar. 2023.

Wikipedia. Reddit. Wikimedia Foundation, Inc. Disponível em: <https://en.wikipedia.org/wiki/Reddit>. Acesso em: 06 abr. 2023.

MONGODB. Documentação oficial do MongoDB. Disponível em: <https://docs.mongodb.com/>. Acesso em: 05 de abril de 2023.