



UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação

Departamento de Sistemas de Computação

Estudo para criação de ferramenta
de apoio na análise de tendências
nas grandes plataformas da internet

Lucas Albano de Oliveira

São Carlos - SP

Estudo para criação de ferramenta de apoio na análise de tendências nas grandes plataformas da internet

Lucas Albano de Oliveira

Orientador: Marcelo Garcia Manzato

Monografia referente ao projeto de conclusão de curso dentro do escopo da disciplina SSC0670 – Projeto de Formatura I do Departamento de Sistemas de Computação do Instituto de Ciências Matemáticas e de Computação – ICMC-USP para obtenção do título de Engenheiro de Computação.

Área de Concentração: Engenharia de Software e Sistemas de Informação

USP – São Carlos
23/11/2021

“Quando olhemos para o futuro, no próximo século, veremos que os líderes serão os que favoreçam o desenvolvimento de todos os outros”

(William Henry Gates III)

Resumo

Informação sempre foi um recurso crucial para auxiliar na decisão correta de planejamentos e soluções. As grandes plataformas sociais da internet concentram uma quantidade expressiva de dados sobre seus usuários, utilizando ferramentas oferecidas por estes serviços é possível extrair uma parte desses dados e gerar análises valiosas. O tempo de análise também é um fator muito relevante em áreas que necessitam estar constantemente atualizadas, uma tarefa difícil considerando a velocidade com que a internet se transforma. Este trabalho foca nesta problemática de se manter atualizado em um mundo digital onde a cada minuto surge uma nova tendência em alguma comunidade da internet. O projeto consiste em uma ferramenta de busca de tendências, e seu objetivo é facilitar a análise inicial de uma tendência através de um resumo rápido e informativo sobre ela, garantindo que o usuário possa filtrar de forma mais eficiente os assuntos relevantes dentro das prováveis palavras-chaves analisadas.

Índice

LISTA DE ABREVIATURAS/SIGLAS	5
LISTA DE FIGURAS	6
CAPÍTULO 1: INTRODUÇÃO	7
1.1. CONTEXTUALIZAÇÃO E MOTIVAÇÃO	7
1.2. OBJETIVOS	8
1.3. ORGANIZAÇÃO DO TRABALHO	9
CAPÍTULO 2: REVISÃO BIBLIOGRÁFICA	10
2.1. CONSIDERAÇÕES INICIAIS	10
2.2. CONCEITOS RELEVANTES	10
2.2.1 <i>API</i>	10
2.2.2 <i>REST API</i>	10
2.2.3 <i>Tendência</i>	11
2.2.4 <i>Comunidade</i>	12
2.2.5 <i>Banco de dados</i>	12
2.3. CONSIDERAÇÕES FINAIS	13
CAPÍTULO 3: DESENVOLVIMENTO DO TRABALHO	14
3.1. CONSIDERAÇÕES INICIAIS	14
3.2. DESCRIÇÃO DO PROJETO	14
3.2.1 <i>Objetivo</i>	14

3.2.2. <i>Palavras-Chave e Localização</i>	15
3.2.3. <i>Fluxo de dados</i>	15
3.2.4. <i>Google Trends</i>	16
3.2.5. <i>Twitter</i>	16
3.2.6. <i>TikTok</i>	17
3.2.7. <i>Youtube</i>	17
3.3. DESCRIÇÃO DAS ATIVIDADES REALIZADAS.....	18
3.3.1 <i>Estrutura do Projeto</i>	18
3.3.2 <i>Coleta de dados: Google Trends</i>	20
3.3.3 <i>Coleta de dados: Twitter</i>	20
3.3.4 <i>Coleta de dados: TikTok</i>	21
3.3.5 <i>Coleta de dados: Youtube</i>	22
3.4. RESULTADOS OBTIDOS	23
3.5. DIFICULDADES E LIMITAÇÕES	30
3.6. CONSIDERAÇÕES FINAIS	31
CAPÍTULO 4: CONCLUSÃO	32
4.1. CONTRIBUIÇÕES	32
4.2. TRABALHOS FUTUROS	32
REFERÊNCIAS	33

Lista de Abreviaturas/Siglas

HTTP	<i>Hypertext Transfer Protocol</i>
REST	<i>Representational State Transfer</i>
API	<i>Application Programming Interface</i>
ISO	<i>International Organization for Standardization</i>

Lista de Figuras

Figura 1 - Diagrama do Fluxo de Dados.....	16
Figura 2 - Barra de pesquisas.....	19
Figura 3 - Pedido de Interesse ao longo do tempo	20
Figura 4 - Endereço de busca Twitter.....	21
Figura 5 - Busca de informações sobre Hashtag no TikTok.....	22
Figura 6 - Requisição de Vídeos pela API do Youtube	23
Figura 7 - Pesquisas do Google para Metaverse	24
Figura 8 - Resultados de Tweets na última semana para Metaverse.....	24
Figura 9 - Primeiro vídeo recomendado para Metaverse	25
Figura 10 - Segundo vídeo recomendado para Metaverse.....	26
Figura 11 - Terceiro vídeo recomendado para Metaverse.....	26
Figura 12 - Engajamento da hashtag Metaverse no TikTok	27
Figura 13 - Pesquisas do Google para Giz de Cera	28
Figura 14 - Resultados de Tweets na última semana para Giz de Cera	28
Figura 15 - Vídeo mais relevante proposto pelo Youtube.....	29
Figura 16 - Engajamento da hashtag Giz de Cera no TikTok	29
Figura 17 - Tempo de resposta das requisições.....	30

CAPÍTULO 1: INTRODUÇÃO

1.1. Contextualização e Motivação

Devido à expansão contínua da internet na vida das pessoas é possível notar novas formas de utilizá-la, quando isso ocorre surge uma necessidade de adaptação tanto por parte de consumidor quanto pelo criador dessa funcionalidade. Porém o volume de mudanças que a internet sofre constantemente é significativamente maior que em meios de comunicação utilizados no passado.

Uma das principais mudanças observada na forma de comunicação entre humanos foi com a aparição e massiva adoção das redes sociais, antes o conteúdo geralmente era expositivo e apenas alguns indivíduos realizavam a produção do conteúdo consumido por vários. Com a dinâmica de qualquer um poder criar e distribuir seu conteúdo de forma simples, muitas relações apareceram que mudaram completamente as interações de pessoa para pessoa. Vale ressaltar que uma das principais características apresentadas nessa mudança é o surgimento das comunidades online.

As comunidades da internet que tomaram controle da maior parte da produção do conteúdo atual oferecem aos seus usuários uma forma fácil de alcançar pessoas, sem a necessidade de aprender habilidades técnicas como programação por exemplo ou desembolsar grandes quantias monetárias como os meios de comunicação anteriores. Essas propriedades entre outras garantiram grande sucesso de tais plataformas, fazendo com que muitas pessoas aderissem essa tecnologia, com isso as comunidades acabaram por aumentar sua atratividade, pois mais conteúdo estava sendo gerado, criando assim um ciclo de expansão onde mais e mais usuários as adotam como principal forma de comunicação.

O tamanho dessas comunidades é uma vantagem, porém ao mesmo tempo um problema, o volume de postagens muitas vezes acaba por dificultar a descoberta de conteúdos em específico. A velocidade em que novos assuntos são criados também exige um tempo considerável de seus membros para ficarem atualizados, a demanda é ainda maior para usuários de fora da comunidade. E extrapolando esse conceito para uma pessoa

que acompanha um número elevado de comunidades é possível perceber que essa operação pode ficar impraticável rapidamente.

A constante demanda por atualização gera um desgaste para quem necessita manter-se constantemente atualizado, encontrar a informação desejada não é uma tarefa simples quando ela está em meio a incontáveis outras. Helvoort (2010) indicou que existe uma maior demanda por parte dos chamados trabalhadores do conhecimento para o desenvolvimento de competência informacional, competência essa que engloba desde a aquisição até a utilização adequada da informação.

Um estudo realizado com professores estadunidenses que trabalham vendendo suas aulas em cursos online indica que o processo de divulgação do material é uma parte significativamente estressante do processo de docência online, segundo pesquisa realizada por Shelton & Archambault (2019), 64% (261) dos entrevistados responderam que concordam plenamente e 26% (106) concordam em partes que a tarefa de tutoria online é desafiadora, com tarefas como criar conteúdo, marketing, usar redes sociais, etc. Outro dado relevante da mesma pesquisa foi que 49% (198) concordam plenamente e 35% (143) acham desafiadora a tarefa de se manter atualizados com todas as novas tendências que aparecem para tutorias online.

Este trabalho visa contribuir para o estudo de tendências online através de uma ferramenta intuitiva para o usuário buscar informações rápidas e relevantes que ajudam o interessado a discernir o que é a tendência em específico e qual é seu nível de engajamento através de plataformas da internet.

1.2. Objetivos

Este trabalho visa criar uma ferramenta de busca para auxiliar na tarefa específica de manter-se atualizado em novas tendências que utiliza os dados presentes nas grandes plataformas da internet e de suas ferramentas para recolhê-los em tempo real. Gerando um resumo para contextualizar o usuário sobre o que é esta nova tendência e como está seu engajamento a curto e longo prazo expondo os dados através de gráficos construídos

instantaneamente, concentrando informações que estariam dispersas através da internet em apenas um lugar em um instante.

1.3. Organização do Trabalho

O capítulo 2 deste trabalho reúne e explica as noções-chave necessárias para o entendimento das atividades desenvolvidas nesta obra. O capítulo 3 contém de forma detalhada os processos realizados para a produção do projeto assim como os resultados obtidos. Por fim o capítulo 4 compreende as conclusões retiradas da realização do trabalho.

CAPÍTULO 2: REVISÃO BIBLIOGRÁFICA

2.1. Considerações Iniciais

Este capítulo contém explicações sobre os conceitos necessários no entendimento sobre como o projeto foi elaborado, de qual forma os dados são coletados, como eles foram expostos e quais são as opções oferecidas por tais soluções. Também está contido nesta parte uma explicação sobre comunidades online e tendências que são o foco desta obra.

2.2. Conceitos Relevantes

Nesta seção está descrito os conceitos relevantes para o entendimento deste trabalho.

2.2.1 API

O termo API (*Application Programming Interface* – Interface de Programação de Aplicação), esta obra se refere mais especificamente à Web API, uma interface que é estabelecida entre *softwares* e tem como propósito intermediar a transmissão de dados sem a necessidade de um código se aprofundar no outro utilizando dos protocolos disponíveis (e.g. HTTP) (SOHAN; ANSLOW; MAURER, 2015), oferecendo uma conexão de fácil acesso e confiável para os membros.

2.2.2. REST API

A arquitetura REST (o inglês *Representational State Transfer* – Transferência de Estados Representacional) é uma forma de estruturar um *software* para garantir sua eficiência, simplicidade, portabilidade, performance, modificabilidade, visibilidade, confiabilidade e escalabilidade para lidar com sistemas distribuídos (FIELDING, 2000). Neste conceito o *software* segue a estrutura de cliente/servidor. O servidor é responsável por gerir e resolver pedidos chamados de *requests*, essa tarefa pode incluir ações como buscas em bancos de dados, atividades computacionais específicas como operações matemáticas complexas, etc. Os clientes são os pontos de comunicação responsáveis por

gerenciar e enviar as *requests* para o servidor, que no caso deste trabalho são criadas por um usuário humano, qualquer pedido gerado pelo usuário deve ser propriamente adaptado para o formato de comunicação REST pelo cliente e enviado em uma mensagem HTTP (*Hypertext Transfer Protocol* – Protocolo de Transferência de Texto) para o servidor. Outros princípios que constituem a arquitetura REST são a ausência do armazenamento de estados que garante que cada *request* é autossuficiente possuindo todas as informações necessárias para o receptor realizar a tarefa; toda comunicação pode ser armazenada em *cache* potencialmente reduzindo a carga de trabalho dos servidores pois reduz o número de pedidos; o sistema é separado e organizado em camadas, possibilitando assim a fácil adição ou subtração de intermediários durante a comunicação; conter uma interface uniforme que garante a dissociação entre as partes da arquitetura facilitando a troca de componentes internos e a capacidade de adaptação a diversos tipos de dados enviados.

Essa arquitetura é fundamental para estabelecer comunicação com os dados oferecidos pelas plataformas da internet através de suas APIs e comunicação através do protocolo HTTP.

2.2.3. Tendência

Uma tendência é uma classificação de um determinado assunto quando ele atinge um nível de interesse que ultrapassa o seu padrão. Tendências estão intrinsecamente conectadas com a estrutura das redes sociais, tendo como início geralmente em meio a uma comunidade (RECALDE; et al, 2017).

Tendências geralmente estão associadas a um texto curto, assim como descrito na seguinte frase: “Quando uma palavra, frase, ou uma *hashtag* é usada com alta frequência, é dito que ela está relacionada com um *trending topic*” (RECALDE; et al, 2017, p. 205, tradução minha) e este pedaço de texto ajuda aos usuários encontrarem os tópicos através de ferramentas de busca e usar como uma identificação para determinar qual tendência estão se referindo, e é desse pedaço de texto que pode ser retirada a *keyword* (Palavra-Chave), palavra que pode ser usada em uma pesquisa para encontrar sua tendência correspondente.

2.2.4. Comunidade

O conceito de comunidades já existia antes do advento da internet, porém assim como em outros campos essa tecnologia foi disruptiva o suficiente para alterar profundamente a forma de comunicação entre esses grupos. Uma definição não formal do conceito das comunidades online é “Um agrupamento de entidades, indivíduos ou organizações que se reúnem tanto temporariamente quanto permanentemente através de um meio eletrônico para interagir em um problema comum ou espaço de interesse” (PLANT, 2004, p. 54, tradução minha).

As comunidades no contexto da internet possuem papel fundamental na forma como as pessoas interagem e influenciam umas às outras (DHOLAKIA; BAGOZZI; PEARO, 2004). Os laços formados nesses grupos facilitam a disseminação de informações, aumentando a velocidade com que o conteúdo se espalha e influenciando seu destaque em algoritmos de recomendação instalados nas plataformas sociais (FIGUEIREDO; et al, 2014), inicialmente quem define se um conteúdo gerado por um usuário é relevante ou não é a própria comunidade em que o usuário está inserido, assim se sua recepção for significativa o suficiente os algoritmos de recomendação reconhecerão esse destaque e irão espalhar a postagem para fora da comunidade inicial, dessa forma as comunidades são curadoras de conteúdo e tem grande influência na sucesso ou não de uma tendência (FIGUEIREDO; et al, 2014).

2.2.5. Banco de dados

Bancos de dados são uma estrutura criada para armazenar de forma eficiente e garantir que vários usuários tenham acesso ao mesmo tempo um conjunto de dados. Essas estruturas são divididas entre itens que são a menor unidade de dado possível contendo uma informação singular; registros que são o agrupamento de itens relacionados; um arquivo que é o conjunto de registros pertencentes a uma mesma categoria (TEOREY; LIGHTSTONE; NADEAU, 2007).

As plataformas online utilizadas nesse projeto armazenam seus dados na estrutura de um banco de dados e disponibilizam a retirada desses dados através de Web APIs

2.3. Considerações Finais

Neste capítulo foi tratado sobre os conceitos relevantes para compreender esta obra, da coleção de dados através das Web APIs com uma arquitetura REST API, armazenamento em bancos de dados, até a compreensão sobre tendências e comunidades que possuem relação intrínseca e são o foco da análise deste trabalho. No capítulo a seguir está descrito como foi produzido o trabalho vinculado a esta monografia.

CAPÍTULO 3: DESENVOLVIMENTO DO TRABALHO

3.1. Considerações Iniciais

Neste capítulo está reunido de forma completa os passos utilizados para a realização deste projeto, assim como a discussão dos resultados obtidos após a elaboração da ferramenta. Também está presente nesta seção quais foram as dificuldades e limitações encontradas durante o desenvolvimento.

Inicialmente é tratado sobre as definições do projeto, como o problema foi modelado e que forma a solução foi estruturada, na seção seguinte é descrito com detalhes como foi realizado a implementação na prática do modelo elaborado que contém ferramentas como Web API, uma aplicação em React (FACEBOOK, 2021) seguindo a arquitetura REST para facilitar a distribuição da ferramenta.

3.2. Descrição do Projeto

Nesta seção está listado a forma como o projeto foi construído pensando nas necessidades do problema.

3.2.1. Objetivo

O objetivo deste trabalho é desenvolver uma ferramenta de uso imediato para facilitar o processo de análise de tendências em tempo real, fornecendo para o usuário um resumo sobre o desempenho de um determinado tópico escolhido e inserido em uma barra de pesquisa, para este trabalho foi definido a coleta de informações de quatro plataformas populares da internet, Google Trends, Twitter, TikTok e Youtube, sendo que o Google Trends contém as informações de pesquisas da ferramenta de busca Google. O número de plataformas foi definido apenas pelo escopo do projeto e a forma como o projeto foi estruturada disponibiliza a opção de adicionar mais resultados em uma expansão futura.

3.2.2. Palavras-Chave e Localização

Para a identificação de uma tendência utiliza-se uma palavra-chave que é solicitada pelas plataformas para realizar uma busca em seu banco de dados e retornar o conjunto de dados desejado. Essa palavra-chave está atrelada a um idioma o que limita qual o resultado obtido, porém algumas plataformas necessitam a definição de uma localização geográfica sem considerar em qual língua está escrita a pesquisa, portanto foi selecionado neste trabalho a localização “BR” que de acordo com a ISO 3166 (ISO, 1999) define a região brasileira como a preferência para respostas.

3.2.3. Fluxo de dados

O fluxo dos dados atende as especificações estipuladas por cada plataforma, e as restrições de projeto da arquitetura REST, dessa forma o algoritmo detém uma base confiável para sustentar requisições com diferentes plataformas e possibilitar alterações e expansões futuras.

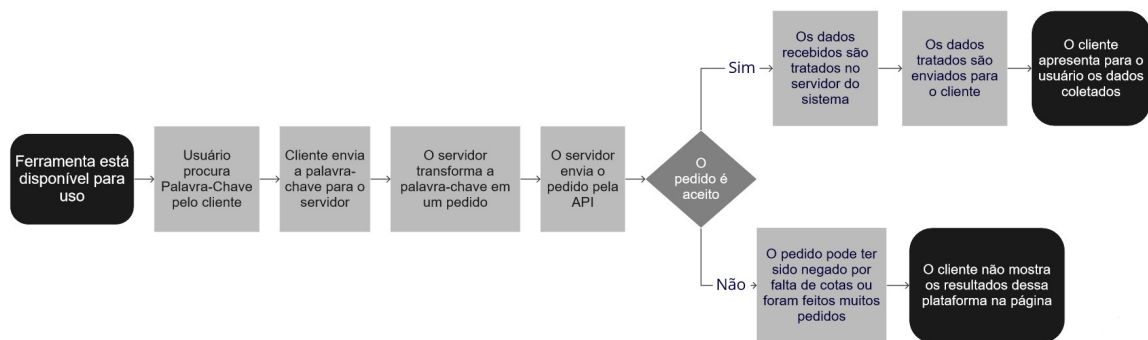
O início do fluxo de dados começa com o usuário fazendo uma consulta na barra de pesquisas do cliente, o cliente envia a palavra para o servidor criar um pedido no formato adequado para cada uma das APIs consultadas.

A coleta dos dados necessários é feita através das APIs oferecidas pelas plataformas, cada resposta de API é diferente uma da outra de acordo com a estrutura dos dados armazenados em seus bancos, entretanto a forma de comunicação é a mesma para todas, com o uso do protocolo de comunicação HTTP, desse protocolo os principais *verbos* utilizados neste projeto são os métodos POST e GET.

Após os dados serem recebidos é necessário um tratamento para torná-los apresentáveis para o usuário, são removidos os valores nulos (valores recebidos como *null*, os zeros são considerados como valores válidos), os dados de tabelas são extraídos de seus encapsulamentos para serem utilizados na criação dos gráficos, e informações não relevantes para o escopo deste trabalho são descartadas.

A figura 1 ilustra a estrutura por completo do fluxo de dados abstraindo a interação com apenas uma API, para atender a mais consultas basta extrapolar o conceito para cada uma das interfaces.

Figura 1 - Diagrama do Fluxo de Dados



Fonte: Elaborado pelo autor

3.2.4. Google Trends

A decisão de coletar os dados do Google Trends foi tomada pois a plataforma reúne informações sobre os tópicos de pesquisas realizadas no Google. Esses dados são processados previamente e o usuário tem acesso a um valor normalizado entre 0 e 100 baseado no volume de pesquisa em proporção a outras pesquisas na época e no local determinado (ROGERS, 2016).

A resposta recebida contém o nome da tendência procurada, uma série de datas e uma quantidade denominada de “Interesse” que corresponde ao valor normalizado de buscas.

3.2.5. Twitter

No Twitter o conteúdo é limitado em mensagens curtas de no máximo 280 caracteres (previamente este valor era de 140), este foco em mensagens curtas pressiona a comunidade a se comunicar de forma direta e criativa (FERREIRA; TAVARES; ABREU, 2010). Podendo conter um arquivo de mídia como fotos ou vídeo, as tendências circulam

através das chamadas *hashtags* que são palavras-chave que os usuários indicam em suas postagens para determinar em qual tópico aquele conteúdo pode se enquadrar.

Os dados coletados do Twitter são referentes ao volume de *Tweets* (postagens) ao longo da última semana em função da palavra-chave, diferente do Google Trends os valores desta categoria são absolutos. A estrutura dos dados é composta por uma série temporal do volume de pesquisas e um valor com o número total de Tweets no tópico durante a última semana (TWITTER, 2021).

3.2.6. TikTok

O TikTok é uma rede social que está crescendo de forma muito acelerada e recentemente atingiu a marca de 1 Bilhão de usuários ativos por mês segundo a própria empresa (TIKTOK, 2021), sua estrutura de conteúdo é focada em vídeos curtos e por consequência alta rotatividade de consumo por parte dos usuários. Dentro da plataforma é possível encontrar as categorias que funcionam de forma similar ao Twitter, concentrando o conteúdo em torno de uma *Hashtag*, mas também pode separar conteúdo pela música ligada ao vídeo ou um *Challenge* que é um desafio sugerido pelos os usuários da plataforma geralmente incentivando outros usuários a participar da proposta.

Do TikTok os dados coletados são em função de suas categorias, é utilizado a palavra-chave para pesquisar quais seções são compatíveis, desta forma seu resultado contém o número total de vídeos e de visualizações.

3.2.7. Youtube

O Youtube é uma das plataformas de compartilhamento vídeo mais populares da internet, muitos criadores de conteúdo a usam como principal fonte de renda. De acordo com uma pesquisa realizada pelo Youtube em conjunto com o instituto Oxford Economics a plataforma contribui direta e indiretamente com aproximadamente R\$ 3,4 bilhões para o produto interno bruto (PIB) brasileiro (GALLOWAY; GOODWIN, 2021). Essa grande quantia monetária é um indício do quão abrangente a influência do Youtube se espalha pela vida dos brasileiros e da sociedade como um todo.

Todo esse tamanho demonstrado pela plataforma indica uma qualidade muito importante oferecida, com muitos usuários existe uma pletora de conteúdo disponível para os consumidores, com o conceito de comunidades e como os algoritmos de recomendação são desenvolvidos para adaptar-se as preferências de seus usuários é inevitável que exista muito conteúdo curado de boa qualidade filtrado pela plataforma, o ambiente competitivo fomenta a proatividade de seus criadores, em suma, no Youtube é possível encontrar conteúdo com boa qualidade mesmo para tópicos recentes devido ao ecossistema presente na plataforma, portanto foi definido neste projeto o Youtube como plataforma para suplementar informações sobre o tópico em questão para o usuário.

A pesquisa no Youtube é feita através da palavra-chave para encontrar os vídeos mais relevantes para o assunto, a categorização de relevância dos vídeos é feita pela própria plataforma para cada termo de pesquisa além de considerar uma diversa gama de parâmetros, a resposta é constituída pelos vídeos mais bem classificados pela plataforma (YOUTUBE, 2019).

3.3. Descrição das Atividades Realizadas

Nesta seção está descrito de forma detalhada como o projeto foi construído, quais ferramentas, quais algoritmos, e quais técnicas foram utilizadas. O projeto está armazenado no site Github e foi disponibilizado publicamente (OLIVEIRA, 2021).

3.3.1 Estrutura do Projeto

O projeto foi criado seguindo a estrutura de Cliente/Servidor constituído majoritariamente pela linguagem de programação JavaScript. O cliente engloba a interação com o usuário, como a entrada e saída dos dados, através da biblioteca *open source* React que utiliza o conceito de componentes para construir interfaces modulares com alta compatibilidade com dispositivos diversos e que está em constante desenvolvimento e manutenção. Para o servidor foi utilizada a tecnologia Node.js que “foi projetada para lidar com aplicações escaláveis em redes de computadores” (OPENJS FOUNDATION, 2021).

No cliente foi criada uma interface interativa com a biblioteca de componentes e estilização chamada React Material UI (MATERIAL-UI SAS, 2021) para o usuário inserir a palavra-chave em uma barra de pesquisa apresentado na figura 2.

Figura 2 - Barra de pesquisas



Fonte: Elaborado pelo autor

Assim que um pedido de busca é recebido pela interface o cliente envia para o servidor o termo de pesquisa proposto através do protocolo de comunicação HTTP com o uso do verbo POST onde o corpo da requisição contém a palavra-chave. Essa estrutura garante ao sistema que mais de um cliente possa atuar concomitantemente e não impacte desproporcionalmente o desempenho do servidor.

A partir do momento em que o servidor recebe a palavra-chave ele a adapta para as necessidades de cada uma das plataformas, de acordo com os modelos estipulados por cada uma das APIs.

3.3.2 Coleta de dados: Google Trends

A coleta dos dados do Google Trends consiste em converter o termo de pesquisa para uma requisição, utilizando a biblioteca Google-Trends-API (TRASBORG, 2016) no formato *interest over time* demonstrado na figura 3, assim é recuperado uma série temporal de 2004 até três dias antes da data da pesquisa contendo o valor normalizado do interesse de pesquisas. Com essa resposta o servidor extraí os valores criando uma matriz de duas dimensões para enviar para o cliente.

Figura 3 - Pedido de Interesse ao longo do tempo

```
async function interestTime(keyword){
  return new Promise(resolve => {
    googleTrends.interestOverTime({keyword: keyword}, function(err, results){
      if(err) console.error('there was an error on google trends!', err);
      else {
        resolve(results);
      }
    })
  });
}
```

Fonte: Elaborado pelo autor

No cliente os dados recebidos são apresentados com o uso da biblioteca Chartist.Js (KUNZ, 2014) e a React-Chartist (XU, 2014) que geram os gráficos e os componentes React respectivamente. Os gráficos são gerados em tempo real pelo próprio navegador do usuário e, portanto, não necessitam enviar imagens através de internet reduzindo o consumo de banda.

3.3.3 Coleta de dados: Twitter

Para o Twitter é utilizada a API para desenvolvedores fornecida pela plataforma, o acesso é garantido para contas do Twitter que foram aprovadas para desenvolvimento.

Com a conta confirmada o Twitter fornece chaves de segurança para a validação da sessão, cada usuário possui um número de cotas de acordo com sua categoria de conta de desenvolvedor e permissão ou não de acessar requisições especializadas. A conta criada para o desenvolvimento deste projeto foi uma conta de desenvolvedor simples, porém com uma validação de conta para pesquisa acadêmica ou empresa é possível ampliar o acesso a plataforma.

Do Twitter foi retirado uma série temporal dos volumes de postagens referentes ao termo chave durante a última semana, para isso é construído um endereço de pesquisa como ilustrado na figura 4 desta forma, basta realizar um pedido HTTP GET para o endereço gera e a resposta obtida contém os dados desejados, nesse caso uma conta com mais permissões poderia acessar um período limitado apenas pelos dados armazenados pela plataforma.

Figura 4 - Endereço de busca Twitter

```
function createUrlTwitter(query){  
    var url = "https://api.twitter.com/2/tweets/counts/recent"  
    url += '?query=' + query  
    return url  
}
```

Fonte: Elaborado pelo autor

3.3.4 Coleta de dados: TikTok

No TikTok os dados são coletados através da biblioteca TikTok-API (CRAIG, 2019), as informações retiradas são dependentes das da palavra-chave e buscam as estatísticas da *Hashtag* encontrada assim como apresentado na figura 5. Os dados interessantes encontrados englobam o volume de vídeos postados o que pode ser utilizado para averiguar o engajamento dos criadores de conteúdo com o tópico e o número de visualizações da *Hashtag* que pode ser um indício do nível de engajamento dos usuários da plataforma com a tendência procurada.

Figura 5 - Busca de informações sobre Hashtag no TikTok

```
async function getTikTokHashtag(keyword){  
    return new Promise(resolve =>{  
        const hashtag = TikTokScraper.getHashtagInfo(keyword);  
        resolve(hashtag);  
    })  
}
```

Fonte: Elaborado pelo autor

3.3.5 Coleta de dados: Youtube

Na coleta de dados do Youtube o Google oferece a Youtube Data API que é uma interface que possui acesso a todo o conteúdo presente na plataforma do Youtube, para requisitar acesso a esta API é necessário a criação de uma conta de desenvolvedor Google e especificar que será utilizado a Youtube Data API. Da mesma forma que o Twitter a interface do Youtube fornece ao desenvolvedor chaves de segurança que são fundamentais para a autorização do pedido e devem ser incluídas no envio da requisição.

O pedido contém os parâmetros de pesquisa que são utilizados para delimitar como a resposta será construída, o primeiro parâmetro é sobre a pesquisa de vídeos, portanto a *string* (Pedaço de texto) 'video', o segundo parâmetro da pesquisa é a própria palavra-chave, o terceiro serve para indicar quais propriedades do vídeo serão retornadas, para este trabalho foi selecionado 'Snippets' e 'Id' que são informações sobre o vídeo como título, descrição, visualizações, por exemplo, e o código de identificação do vídeo, após isso são indicados quantos vídeos serão retornados por página, que para este trabalho foram selecionados 3 vídeos, a localização foi definida como 'BR' e foi requerido vídeos que possam ser embutidos em sites fora do Youtube, essa requisição é ilustrada pela figura 6.

Figura 6 - Requisição de Vídeos pela API do Youtube

```
function getVideo(keyword) {  
  return new Promise(resolve => {  
    var url = createUrlSearch('video',keyword,'snippet,id','3','BR','true')  
    axios.get(url)  
      .then(response => {  
        resolve(parseData(response.data));  
      })  
      .catch(error =>{  
        console.log(error)  
      })  
  });  
}
```

Fonte: Elaborado pelo autor

3.4. Resultados Obtidos

Após a construção da ferramenta foi realizado alguns testes com palavras que são consideradas tendências e com palavras comuns que não possuem uma tendência atrelada para a verificação do funcionamento da ferramenta.

A primeira palavra inspecionada foi: “*Metaverse*” que recentemente apareceu na mídia após a mudança de nome do Facebook para Meta, e suas propostas para a criação de um metaverso, que em suma seria um ecossistema digital da empresa. Porém o termo também é utilizado por alguns membros da comunidade de criptomoedas para definir um conceito similar, então era esperado encontrar respostas seguindo estes dois aspectos. Na figura 7 é possível observar o resultado de pesquisas na ferramenta Google ao longo do tempo, e também é possível observar o crescimento expressivo de buscas após o pronunciamento do Facebook.

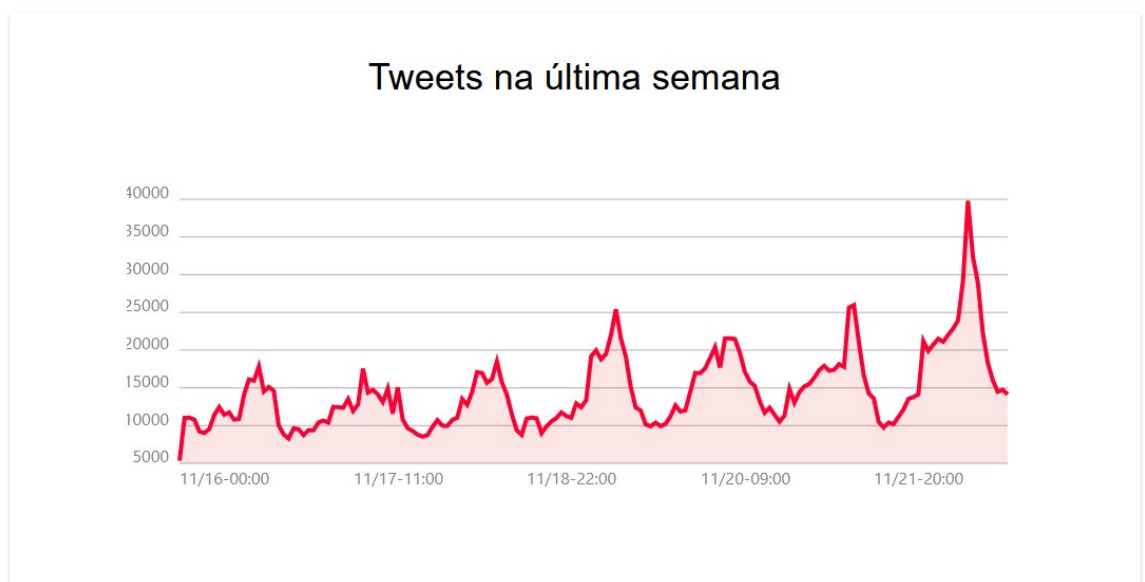
Figura 7 - Pesquisas do Google para *Metaverse*



Fonte: Elaborado pelo autor

É possível ver que na figura 8 está contido o volume de *Tweets* ao longo da semana, esta busca foi realizada utilizando o *endpoint* de *tweet count*, que busca na plataforma postagens que contém a palavra-chave pesquisada.

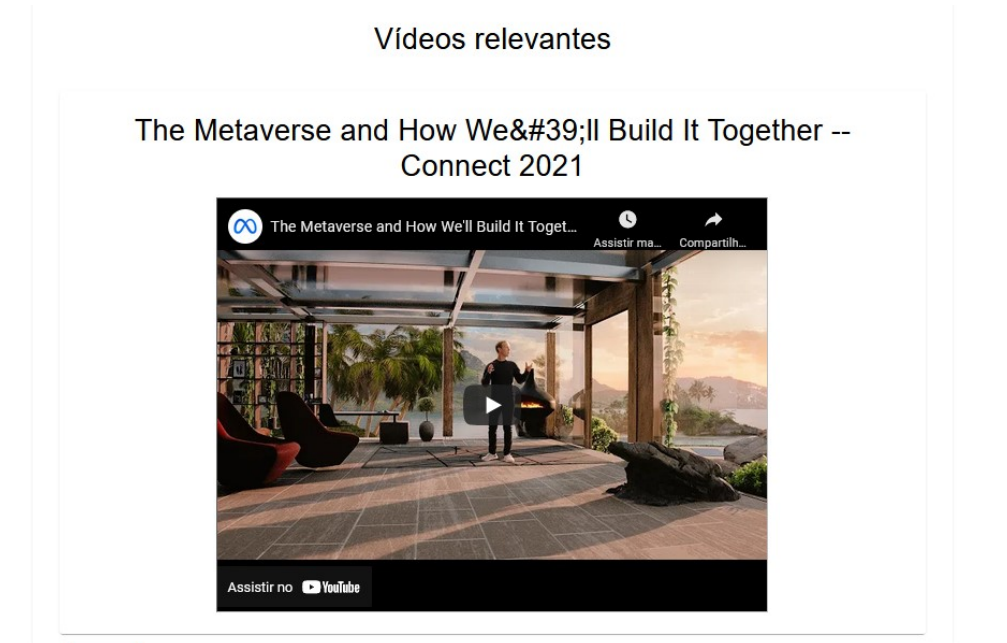
Figura 8 - Resultados de Tweets na última semana para *Metaverse*



Fonte: Elaborado pelo autor

Na figura 9 aparece o primeiro resultado para a pesquisa na ferramenta do Youtube utilizando a palavra-chave *Metaverse*, que é o resultado mais relevante para a pesquisa de acordo com o algoritmo de recomendação da plataforma, este vídeo é a apresentação oficial publicada pelo Facebook onde eles anunciam seus planos para o rumo da empresa.

Figura 9 - Primeiro vídeo recomendado para *Metaverse*

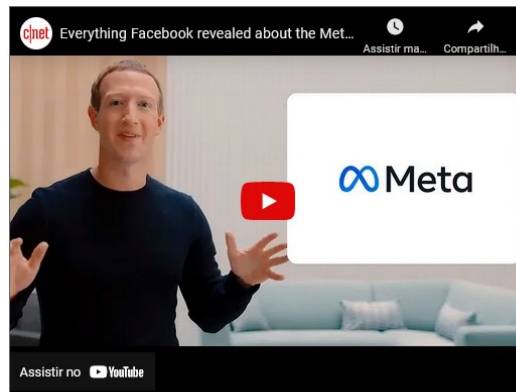


Fonte: Elaborado pelo autor

A figura 10 mostra o segundo resultado recebido pelo Youtube, um vídeo que resume os tópicos importantes da apresentação de mais de uma hora, para apenas 11 minutos, este tipo de conteúdo é bem popular na plataforma, e podem ser encontrados em diversos tópicos por diferentes criadores.

Figura 10 - Segundo vídeo recomendado para *Metaverse*

Everything Facebook revealed about the Metaverse in 11 minutes

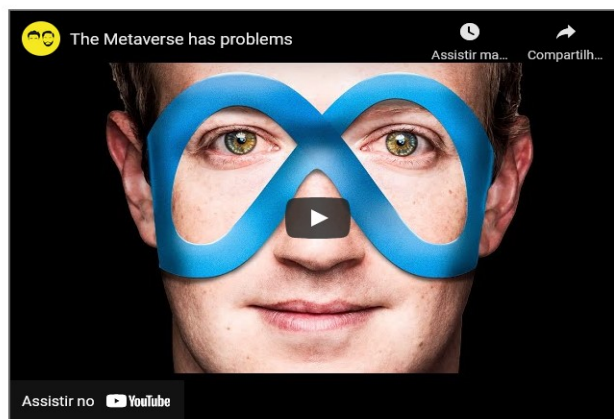


Fonte: Elaborado pelo autor

Por fim na figura 11 é recomendado um vídeo que faz uma crítica à proposta do Facebook, e elenca quais são os pontos negativos desta ideia de Metaverso. Críticas são uma forma de encontrar opiniões diferentes sobre um determinado assunto que podem auxiliar ao usuário da ferramenta a tomar uma decisão mais informada.

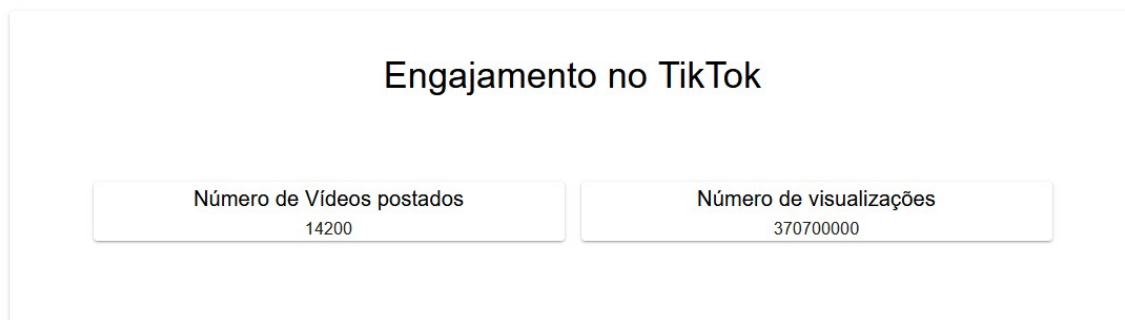
Figura 11 - Terceiro vídeo recomendado para *Metaverse*

The Metaverse has problems



Fonte: Elaborado pelo autor

Figura 12 - Engajamento da *hashtag* Metaverse no TikTok



Fonte: Elaborado pelo autor

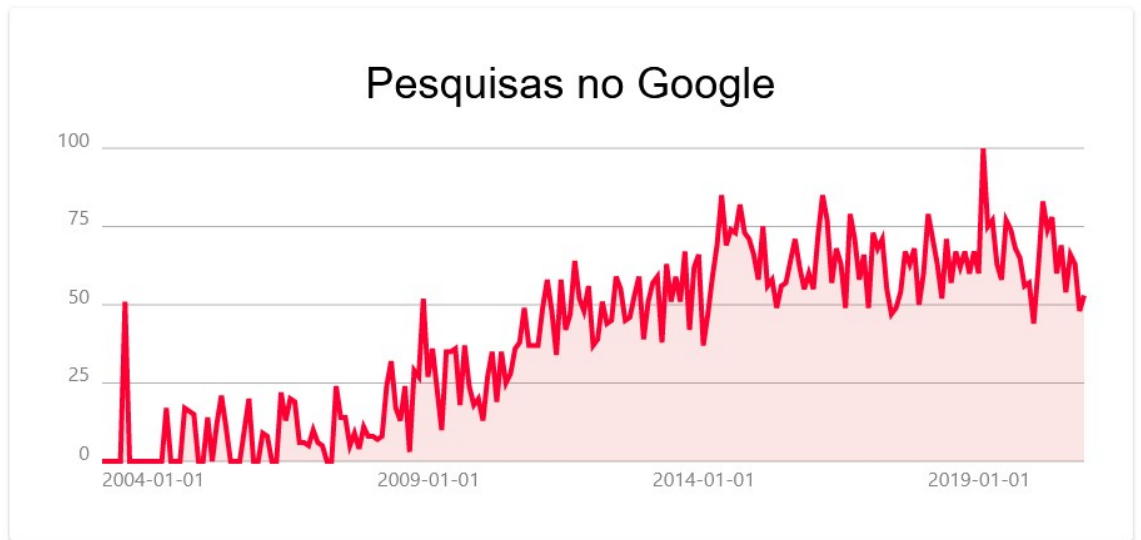
A figura 12 mostra o resultado obtido pela busca da palavra *Metaverse* na rede social do TikTok, e nela é possível observar um grande número de visualizações e uma alta quantidade de vídeos contidos na *hashtag*.

Um resultado observado desta tendência é que mesmo especificando a região como “BR” os vídeos sugeridos foram tão populares que superaram esta limitação e foram recomendados mesmo que sejam na língua Inglesa. Outro ponto observado é que o Termo Metaverse criado pelo Facebook é tão popular que obscurece as outras variações, este caso específico onde um termo oclui outros resultados demanda um estudo mais aprofundado para avaliar os méritos de se expandir ou não a ferramenta para acomodar esta situação, porém essa discussão está fora do escopo proposto por este projeto.

Agora para um termo de pesquisa comum que não está atrelado a nenhuma tendência de imediato, foi determinado o termo Giz de cera para servir como exemplo, na simplificação da visualização apenas um resultado será mostrado sobre os vídeos obtidos.

Para os resultados da figura 13 é possível observar um crescimento nas buscas ao longo do tempo, porém em uma primeira observação, não é detectado nenhuma modificação expressiva e contínua.

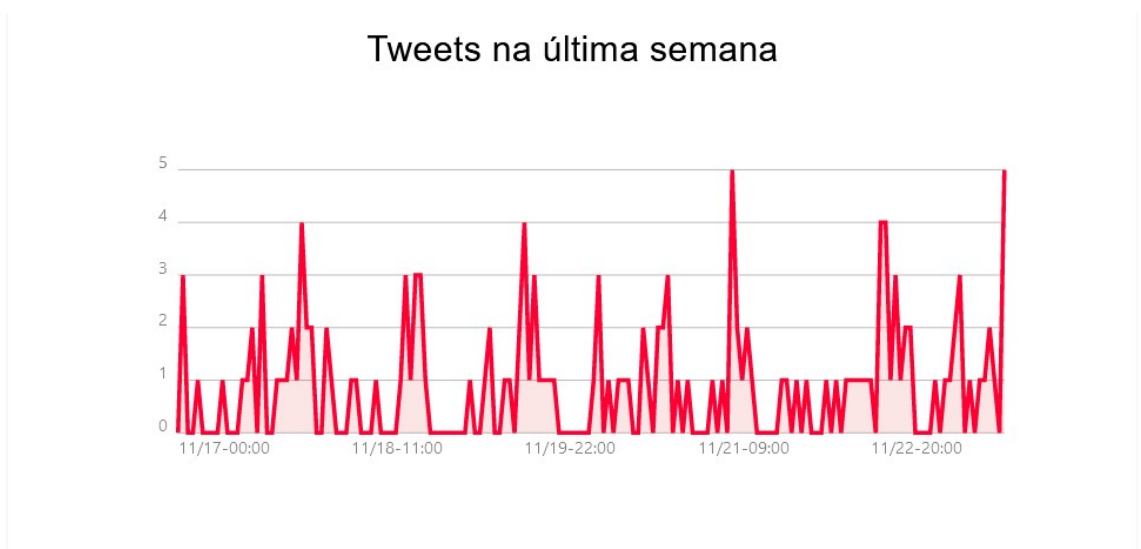
Figura 13 - Pesquisas do Google para Giz de Cera



Fonte: Elaborado pelo autor

Já na figura 14 o volume de *Tweets* dentro de uma semana é bem menor que em comparação com a palavra-chave *Metaverse*, variando entre 0 e 5 *Tweets* a cada hora, portanto não são resultados expressivos para uma tendência.

Figura 14 - Resultados de Tweets na última semana para Giz de Cera



Fonte: Elaborado pelo autor

O vídeo encontrado pela ferramenta na busca de Giz de cera foi a música Chão de Giz, uma gravação ao vivo do cantor Zé Ramalho, evento que foi gravado muito tempo antes da busca, e não representa uma grande tendência na plataforma.

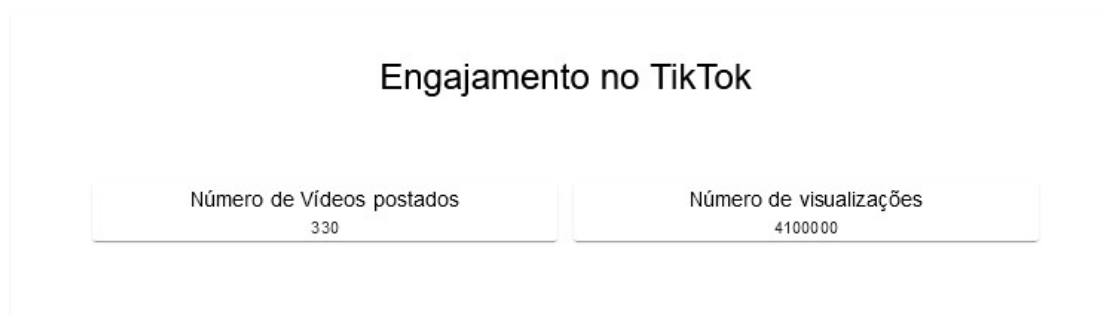
Figura 15 - Vídeo mais relevante proposto pelo Youtube



Fonte: Elaborado pelo autor

Por fim na figura 16 está contido os resultados para a busca no TikTok, e mais uma vez são encontrados números inferiores quando comparados com a tendência *Metaverse*.

Figura 16 - Engajamento da hashtag Giz de Cera no TikTok



Fonte: Elaborado pelo autor

Foi possível observar uma grande variação dos resultados em comparação com o resultado para *Metaverse*, o volume de *Tweets* é expressivamente menor, assim como o número de vídeos postados e de visualizações no TikTok, claramente esta análise rápida não é o suficiente para determinar com total propriedade no desempenho de uma tendência, porém a proposta deste trabalho é criar uma ferramenta que dá ao usuário um indício de como está o comportamento da tendência de forma rápida. Mencionando o tempo de resposta, foi coletado também o tempo que demora para receber as respostas de cada uma das plataformas no cliente da aplicação assim como mostrado na figura 17.

Figura 17 - Tempo de resposta das requisições

200	POST	localhost:3001	youtube	main.chunk.js:365 (fetch)	html	544 B	299 B	1,28 s
204	OPTIO...	localhost:3001	twitter	fetch	plain	357 B	0 B	7 ms
204	OPTIO...	localhost:3001	trends	fetch	plain	357 B	0 B	3 ms
204	OPTIO...	localhost:3001	tiktok	fetch	plain	357 B	0 B	4 ms
200	POST	localhost:3001	twitter	main.chunk.js:374 (fetch)	html	6,82 KB	6,58 KB	672 ms
200	POST	localhost:3001	trends	main.chunk.js:383 (fetch)	html	8,01 KB	7,77 KB	1,89 s
200	POST	localhost:3001	tiktok	main.chunk.js:392 (fetch)	html	287 B	42 B	747 ms

Fonte: Elaborado pelo autor

Os tempos de respostas estão contidos na última coluna, o maior tempo obtido foi 1,89s para o Google Trends, portanto o usuário tem uma resposta em aproximadamente 2 ou 3 segundos após sua pesquisa inicial.

3.5. Dificuldades e Limitações

Durante o desenvolvimento deste trabalho o autor se deparou com algumas limitações impostas pelas plataformas para controlar quantos dados podiam ser recuperados dentro de um limite de tempo, limite de cotas foi o termo designado para referir-se a estas limitações. Esses limites impediram em algumas situações que o autor desenvolvesse métodos de coleta de dados mais aprofundados, como por exemplo uma inspeção mais completa dos termos e variações relacionadas da pesquisa no Youtube, pois a ferramenta de pesquisa do Youtube Data API consome um número elevado de cotas cada

vez que é utilizada, inviabilizando uma busca mais completa que fosse sustentável considerando a utilização por mais de um usuário ao mesmo tempo.

Outra limitação encontrada também pode ser exemplificada pela API do Twitter que não fornece acesso a algumas funcionalidades para contas de desenvolvimento comuns, como era o caso neste projeto. A opção de recuperar o volume de Tweets sem restrição de datas seria muito positiva para a comparação com os dados obtidos pelo Google Trends, porém esta funcionalidade não pode ser acessada por contas comuns e necessita de validações acadêmicas para sua utilização.

3.6. Considerações Finais

Neste capítulo foi tratado sobre como o projeto foi desenvolvido, desde os conceitos da arquitetura até a implementação da coleta e mostra de dados para o usuário, no capítulo a seguir serão apresentadas as conclusões finais obtidas por este trabalho.

CAPÍTULO 4: CONCLUSÃO

4.1. Contribuições

A ferramenta desenvolvida por este trabalho pode ser usada por pessoas que necessitam manter-se atualizadas nas tendências em determinadas comunidades, reduzindo o tempo gasto estudando cada tópico individualmente devido a síntese e agrupamento feito pelo *software* desenvolvido. O algoritmo pode ser disponibilizado em uma plataforma web de forma simples devido a sua construção utilizando a biblioteca *open source* React, permitindo que o *software* ser utilizado por muitas pessoas de forma escalável bastando apenas o distribuidor do algoritmo entrar em contato com as empresas responsáveis pelas APIs para negociar cotas adequadas para seu público-alvo.

4.2. Trabalhos Futuros

A forma como este trabalho foi organizado possibilita a fácil modificação do seu conteúdo para a adaptação considerando situações onde as necessidades dos usuários mudem de forma considerável, utilizando plataformas diferentes, com intensidades variadas. Portanto expandir o trabalho adicionando novas redes sociais para a resposta do algoritmo seria uma tarefa simples. Uma atualização mais aprofundada deste *software* poderia garantir ao usuário sugestões de palavras-chaves que foram coletadas e detectadas como tendências de forma automática utilizando algoritmos avançados de classificação por exemplo. Assim como uma pesquisa utilizando privilégios superiores seria ideal na continuidade deste projeto pois isto aumentaria muito as possibilidades disponíveis para a ferramenta.

REFERÊNCIAS

CRAIG, J. W. **TikTok-API**. 2019. Disponível em: < <https://github.com/szdc/tiktok-api>>. Acesso em: 19 de nov. de 2021.

DHOLAKIA, Utpal M.; BAGOZZI, Richard P.; PEARO, Lisa Klein. A social influence model of consumer participation in network-and small-group-based virtual communities. **International journal of research in marketing**, v. 21, n. 3, p. 241-263, 2004.

FACEBOOK. React, 2021. Biblioteca Open Source em JavaScript React. Disponível em: <https://pt-br.reactjs.org/>. Acesso em: 10, nov. de 2021

FERREIRA, Robson Thiago; TAVARES, Daiane da Silva; ABREU, Karen Cristina Kraemer. O Twitter como ferramenta de comunicação organizacional. **Salvador: RP em revista**, 2010.

FIELDING, Roy Thomas. **Architectural styles and the design of network-based software architectures**. University of California, Irvine, 2000.

FIGUEIREDO, Flavio et al. On the dynamics of social media popularity: A YouTube case study. **ACM Transactions on Internet Technology (TOIT)**, v. 14, n. 4, p. 1-23, 2014.

GALLOWAY, Hamilton; GOODWIN, Andrew P. From Opportunity to Impact: Assessing the Economic, Societal and Cultural Benefits of YouTube in Brazil. **Oxford Economics**, 16 de set. de 2021 Disponível em: < <https://www.oxfordeconomics.com/recent-releases/From-Opportunity-to-Impact-Assessing-the-Economic-Societal-and-Cultural-Benefits-of-YouTube-in-Brazil>>. Acesso em: 15 de nov. de 2021.

ISO. The International Standard for country codes and codes for their subdivisions. **International Organization for Standardization**, Geneva, Switzerland, 1999. Disponível em: <<https://www.iso.org/iso-3166-country-codes.html>>. Acesso em: 10 de nov. de 2021.

KUNZ, Gion. **Chartist-Js**. 2014. Disponível em: <<https://github.com/gionkunz/chartist-js>>. Acesso em: 19 de nov. de 2021.

MATERIAL-UI SAS. **Material UI**, c2021. Biblioteca de componentes de interface para React. Disponível em: <https://mui.com/pt/>. Acesso em: 18, nov. de 2021.

OLIVEIRA, Lucas Albano. **Projeto-Tcc-Trends**. 2021. Disponível em: <<https://github.com/Tornom/Projeto-Tcc-Trends>>. Acesso em: 23 de nov. de 2021.

OPENJS FOUNDATION. **NodeJS**, c2021. Sistema de tempo de execução focado em eventos para JavaScript. Disponível em: <https://nodejs.org/en/about/>. Acesso em: 17, nov. de 2021.

PLANT, Robert. Online communities. **Technology in society**, v. 26, n. 1, p. 51-65, 2004.

RECALDE, Lorena et al. Detection of trending topic communities: Bridging content creators and distributors. In: **Proceedings of the 28th ACM Conference on Hypertext and Social Media**. 2017. p. 205-213.

ROGERS, Simon. What is Google Trends data — and what does it mean? What is Google Trends data — and what does it mean?. **Medium**, 2016 Disponível em: <<https://medium.com/google-news-lab/what-is-google-trends-data-and-what-does-it-mean-b48f07342ee8>>. Acesso em: 12 de nov. de 2021.

SHELTON, Catharyn C.; ARCHAMBAULT, Leanna M. Who are online teacherpreneurs and what do they do? A survey of content creators on TeachersPayTeachers. com. **Journal of Research on Technology in Education**, v. 51, n. 4, p. 398-414, 2019.

SOHAN, S. M.; ANSLOW, Craig; MAURER, Frank. A case study of web API evolution. In: **2015 IEEE World Congress on Services**. IEEE, 2015. p. 245-252.

TEOREY, Toby; LIGHTSTONE, Sam; NADEAU, Tom. Projeto e modelagem de banco de dados. **Campus, 4ª edição–2007**, 2007.

TIKTOK, Um bilhão de vezes obrigado!. **TikTok**, 27 de set. 2021. Disponível em: <<https://newsroom.tiktok.com/pt-br/um-bilhao-de-pessoas-no-tiktok>>. Acesso em: 12 de nov. de 2021.

TRASBORG, Patrick. Google-Trends-API. 2016. Disponível em: <<https://github.com/pat310/google-trends-api>>. Acesso em: 19 de nov. de 2021.

TWITTER, Tweet counts. **Developer Twitter**, c2021. Disponível em: <<https://developer.twitter.com/en/docs/twitter-api/tweets/counts/api-reference/get-tweets-counts-recent>>. Acesso em: 12 de nov. de 2021.

VAN HELVOORT, AAJ Jos. Impact of recent trends in information and communication technology on the validity of the construct information literacy in higher education. In: **International Symposium on Information Management in a Changing World**. Springer, Berlin, Heidelberg, 2010. p. 61-73.

XU, Fraser. **React-Chartist**. 2014. Disponível em: <<https://github.com/fraserxu/react-chartist>>. Acesso em: 19 de nov. de 2021.

YOUTUBE. Search: list. **Developers Google**, 11 de Jun. de 2019. Disponível em: <https://developers.google.com/youtube/v3/docs/search/list?hl=pt_br>. Acesso em: 17 de nov. de 2021.