INSTITUTO FEDERAL DE SANTA CATARINA

SUYAN MORIEL VIESE MOURA

Uso de chatbot para otimizar a checagem de fake news em redes sociais

São José - SC

Setembro/2022

USO DE CHATBOT PARA OTIMIZAR A CHECAGEM DE FAKE NEWS EM REDES SOCIAIS

Pré-Projeto de trabalho de conclusão de curso apresentado à Coordenadoria do Curso de Engenharia de Telecomunicações do campus São José do Instituto Federal de Santa Catarina para a aprovação do tema perante banca na disciplina de TCC1.

Orientador: Prof. Cleber Jorge Amaral, Dr.

São José - SC

Setembro/2022

Sempre que te perguntarem se podes fazer um trabalho,
respondas que sim e te ponhas em seguida a aprender como se faz.
T. Roosevelt

RESUMO

As redes sociais têm-se tornado um dos principais meios de difusão de informação em massa, e com isso a difusão de informações falsas também. Atualmente um dos maiores problemas virtuais é a dificuldade de se obter fontes confiáveis. A proposta deste trabalho é o desenvolvimento de uma ferramenta, integrável às redes sociais, capaz de fornecer fontes confiáveis ao usuário, sem que ele precise sair da plataforma a qual está utilizando para fazer a busca.

Palavras-chave: Fake News. ChatBot. Redes sociais.

LISTA DE ILUSTRAÇÕES

Figura 1 – Gráfico de uso de redes sociais 9
Figura 2 – fluxograma do sistema 17
Figura 3 – Arquitetura do sistema 18

SUMÁRIO

1 INTRODUÇÃO . . . 9
1.1 Objetivo Geral . . . 10
1.2 Objetivos Específicos . . . 10
1.3 Organização do texto . . . 11

2 FUNDAMENTAÇÃO TEÓRICA . . . 13
2.1 Fact Checking . . . 13
2.2 API . . . 13
2.3 Twitter . . . 14
2.4 Google Fact Checking Tools . . . 14
2.4.1 Explorer . . . 14
2.5 Micro-Serviços . . . 15

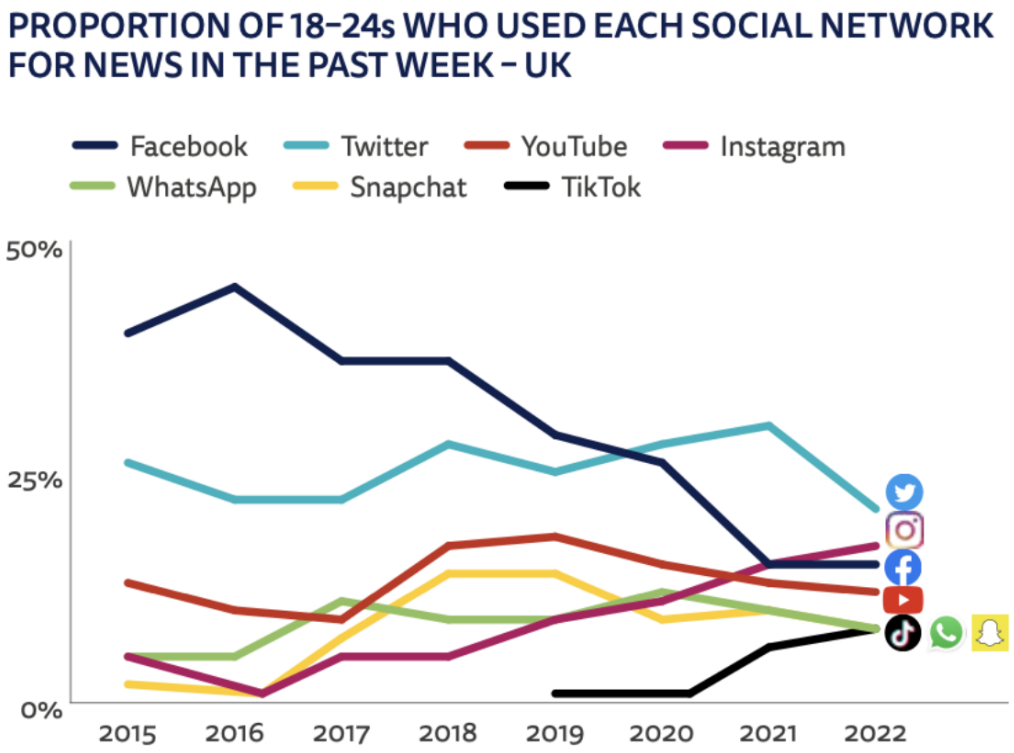
3 PROPOSTA . . . 17
3.1 Modularidade da arquitetura . . . 17
3.1.1 Padronização da arquitetura . . . 18
3.2 Agente fact-checker . . . 18
3.3 Agente Twitter . . . 19

REFERÊNCIAS . . . 21

1 INTRODUÇÃO

Cada vez mais as redes sociais estão se tornando parte de nossas vidas, não só como meio de interação social, mas também como meio de obter informações. No Reino Unido, por exemplo, 39% dos jovens entre 18 e 24 anos preferem se informar através de redes sociais (NEWMAN et al., 2022). Das redes sociais, o Twitter se destaca como fonte de informação jornalística para estes jovens, conforme a Figura 1.

Figura 1 – Gráfico de uso de redes sociais



Fonte: (NEWMAN et al., 2022)

Nos últimos anos têm-se intensificado a discussão referente às fake-news, que são notícias intencionalmente falsas e passíveis de verificação (LAZER et al., 2018).

Em redes sociais como o Twitter, qualquer pessoa pode criar de uma a várias contas e publicar mensagens, chamadas de tweets, que podem conter qualquer tipo de conteúdo e difundir qualquer tipo de informação. O Twitter também permite o retweet, isto é, que um usuário compartilhe a informação publicada por outro usuário. Dessa forma é muito fácil o espalhamento de notícias falsas. No entanto, para descobrir a veracidade referente a algo, é relativamente complexo, às vezes, demanda horas de pesquisa em fontes confiáveis.

Hoje há empresas e ferramentas que facilitam a verificação de fatos, são chamadas de

agências verificadoras de fatos. Estas agências realizam o processo *fact-checking* (verificação de fatos) de uma noticia que geralmente é difundida massivamente em redes sociais.

Atualmente o Twitter permite o uso de *bots*, que são programas de computador que se comportam como usuários convencionais (humanos) da rede social. Estes *bots* podem realizar as mesmas operações que usuários humanos, como ver os tweets nos quais foram marcados, publicar e responder tweets. Assim como usuários humanos, estes robôs estão sujeitos as regras da rede social.

Uma das formas de simplificar esse processo de pesquisa, seria um *bot* integrado ao Twitter, uma das redes mais populares usadas para consumo de notícias, de forma a qual o usuário simplesmente citasse *bot* em um tweet e o *bot* faria essas pesquisas e forneceria os resultados ao usuário.

Dessa forma se tornará simples consultar os fatos como também será simples difundir uma análise averiguada sobre a notícia do tweet em questão. Pois uma vez que o bot respondesse ao tweet no qual ele fora marcado, as referências publicadas pelo bot ficariam disponíveis para todos os usuários da rede.

Sabe-se que identificar fake news pode ser uma tarefa complexa, e mesmo com agências verificadoras, assim há um fator humano interpretativo, e tendo ciência desse grau de complexidade, a proposta deste projeto não é acusar de fake news ou não uma notícia passível a interpretações humanas, mas sim, indicar artigos de sites de checagem, de tal forma que fique fácil ao usuário encontrar fontes sobre algo.

1.1 Objetivo Geral

Criar um bot integrado ao twitter que facilite a obtenção de artigos de fact-checking referentes a um texto de um tweet.

1.2 Objetivos Específicos

1. Desenvolver um bot integrado ao Twitter.

2. Desenvolver um serviço integrado à *Application Programming Interface* (API) de indexação de fact-checking da Google.

3. Desenvolver uma arquitetura, orientada a micro-serviços, que permita facilmente a integração a outras redes sociais e outros serviços de fact-checking.

1.3 Organização do texto

No capitulo 2 é apresentada a fundamentação teórica e no capitulo 3 é apresentada a proposta para este projeto.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Fact Checking

Com a facilidade da difusão de informações em massa, e com a democratização do acesso e divulgação de informações, todo mundo com acesso a internet, pode publicar qualquer conteúdo, sem que haja qualquer forma de verificação. Sendo assim, os usuários de redes sociais estão sujeitos a consumirem fake news, achando que são verdade.

Cada vez mais está aparente o problema das fake news. Com o intuito de resolver tal problema, *surgiram* as chamadas "agências checadoras de fato", as quais têm como objetivo evitar a disseminação de notícias falsas ou imprecisas.

O Fact Checking, tradicionalmente dentro do jornalismo, se relaciona com procedimentos e técnicas para verificação de fatos antes da publicação de uma notícia, isso se chama de verificação interna de fatos, e por volta das décadas de 1920 e 1930 tornou-se uma função distinta dentro das redações jornalísticas.

No entanto, com o crescimento do acesso a internet e a facilidade de compartilhamento de informações, surgiu a necessidade de uma verificação externa de fatos, a qual consiste na publicação de uma análise baseada em evidências e fatos sobre a acurácia de um determinado texto público (GRAVES; AMAZEEN, 2019).

Atualmente existem diversas agências as quais fazem esse serviço de checagem externa de fatos. A fim de indexar as publicações dessas agências, surgiram ferramentas como o Google Fact Checking Tools, onde o usuário pesquisa por algo, e obtém como resultado diversas publicações de agências verificadoras de fatos.

2.2 API

A proposta desse trabalho é integrar, de forma automatizada, as informações de agências verificadoras de fatos com redes sociais. Para isso é necessário a comunicação entre ferramentas indexadoras (como o Google Fact Checking Tools) e as redes sociais. Uma das formas a qual aplicações se comunicam são as APIs. As APis (Application Programming Interface) permitem que um serviço se comunique com outros serviços sem saber como eles foram implementados (FERREIRA, 2021).

As APIs ficam entre uma aplicação e o servidor web, agindo como uma camada intermediária a qual processa dados entre os sistemas. A comunicação por API opera da seguinte forma: Um cliente obtém informações de um servidor através de uma chamada,

a qual é denominada de *request*. O *request* é processado pelo servidor através da URI (Uniform Resource Identifier) e inclui também um verbo, que representa uma ação solicitada (GET, POST ou DELETE, por exemplo) um cabeçalho e às vezes um corpo. O servidor então processa a requisição e retorna ao cliente uma resposta, podendo ser ela de sucesso ou até mesmo de falha (FERREIRA, 2021).

2.3 Twitter

O Twitter é uma das redes sociais mais populares do mundo (NEWMAN et al., 2022). O Twitter permite que usuários interajam publicamente através de mensagens curtas de texto. Ainda há a possibilidade de envios mensagens privadas e outros tipos de mídia (imagens e vídeos, por exemplo).

Neste projeto serão utilizados as seguintes funcionalidades do Twitter:

- **Tweet:** É a forma a qual uma publicação no Twitter é chamada.
- **Menções:** No twitter, um usuário pode citar outro usuário em um tweet. O usuário mencionado recebe uma notificação de que fora marcado em uma determinada publicação. Dessa forma, os usuários podem convidar outras pessoas para visualizarem uma postagem.
- **Thread:** O Twitter organiza as publicações e suas respectivas respostas em formas de thread, ou seja, uma sequência de publicações agrupadas.
- **Reply:** Permite que um usuário responda ao tweet de outro usuário. Assim uma publicação pode se dividir em vários subtópicos, formando uma thread.
- **Cabeça da thread:** É o tweet que dá inicio a uma thread.

2.4 Google Fact Checking Tools

Trata-se de duas (pelo menos até o momento da escrita deste trabalho) ferramentas que visam facilitar o trabalho de jornalistas e pesquisadores. Estas permitem a pesquisa e a publicação de análises de fatos (fact checking). Este projeto utilizará a ferramenta de pesquisa, o Explorer.

2.4.1 Explorer

Permite que o usuário pesquise por um texto qualquer e ela retornará uma lista com os respectivos artigos (GOOGLE, 2023).

O Google Explorer fornece uma API a qual recebe como parâmetros de entrada um texto, e retorna, em caso de sucesso, uma lista de *claims* (forma com a qual a API da google chama os objetos de retorno).

As *claims* contém informações relevantes sobre os artigos, como nome, data de publicação, autor e fonte.

O tipo da resposta é JSON(JavaScript Object Notation), e as *claims* trazem dados referentes a cada fact-checking obtido, como por exemplo URL da fonte, autor, veículo emissor, palavras-chave e etc.

2.5 Micro-Serviços

Uma arquitetura orientada a micro-serviços implica em aplicações que podem ser implementadas, escaladas e testadas de forma independente e que, idealmente, possuem um única funcionalidade (WASEEM; LIANG; SHAHIN, 2020).

Segundo Amaral et al. (2015) as principais razões para se usar micro-serviços são:

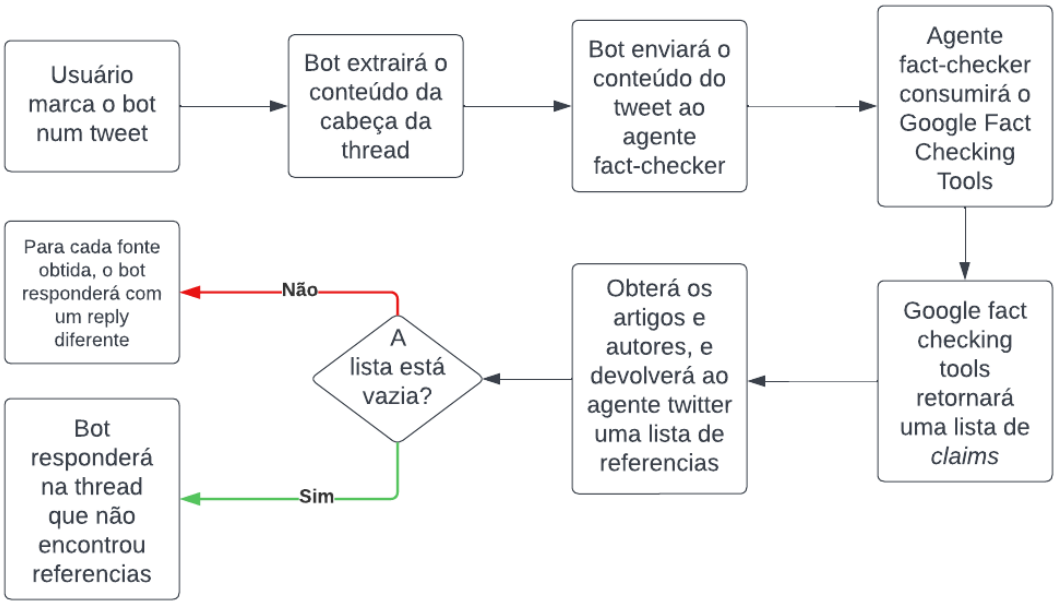
- Reduzir a complexidade de um projeto, utilizando pequenos serviços;
- Escalabilidade;
- Flexibilidade para usar diferentes frameworks e ferramentas.

3 PROPOSTA

A proposta deste TCC é o desenvolvimento de dois *agentes* (micro-serviços) para que operando juntos, sejam capazes de integrar o Google Fact Checking Tools ao Twitter. Será criado um usuário no twitter, este usuário será um bot. Toda vez que este usuário for marcado em uma thread, o agente twitter buscará a cabeça da mesma, extrairá seu conteúdo e o enviará ao agente fact-checker. O Agente fact-checker irá enviar o texto do tweet a Google, a qual retornará a ele uma lista de artigos publicados por agencias verificadoras de fatos credenciadas. O Agente fact-checker por fim irá enviar essa lista de artigos ao agente twitter, o qual responderá ao tweet no qual fora marcado com estes artigos, a fim de indicar fontes ao usuário que o marcou inicialmente.

O fluxograma a seguir demonstra como o sistema se comportará.

Figura 2 – fluxograma do sistema



Fonte: Desenvolvida pelo autor

3.1 Modularidade da arquitetura

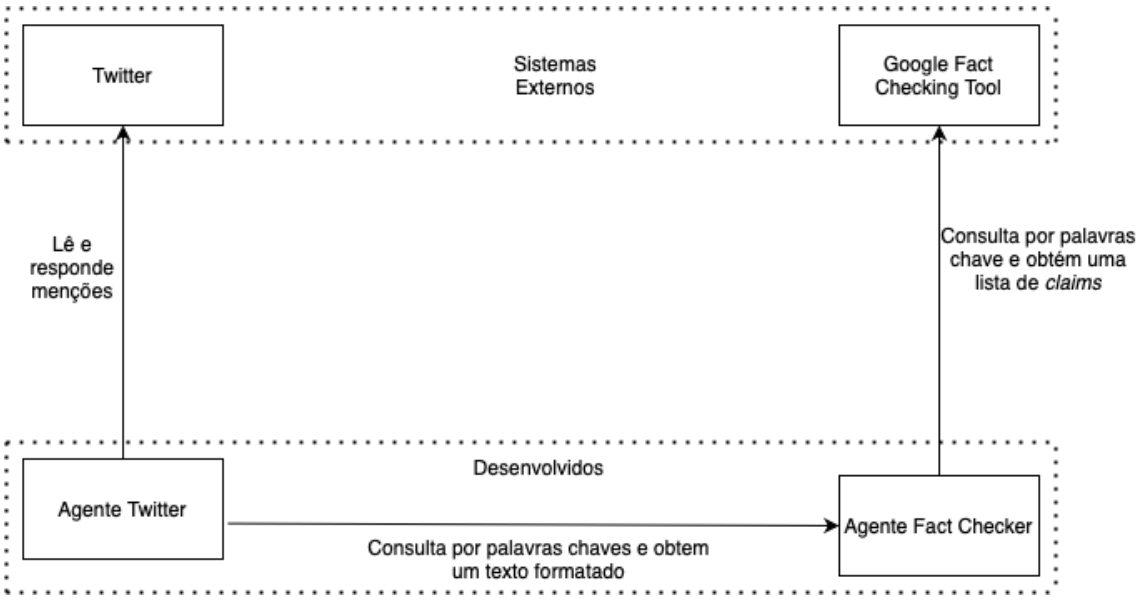
O projeto utilizará a arquitetura de micro-serviços. Será desenvolvido duas aplicações, as quais serão chamadas de agentes, uma vez que possuem um certo grau de autonomia. (RAHWAN et al., 2015).

Estes agentes estarão integrados à uma tecnologia remota. Um agente estará integrado ao Twitter e o outro ao Google Fact Checking Tools. Eles serão chamados respectivamente de "Agente Twitter"e "Agente Fact Checker".

Estes agentes se comunicarão através de APIs. O Agente Fact Checker consome a API da Google e também implementa sua própria API, enquanto o Agente Twitter consome a API do Twitter e a API do Agente Fact Checker.

Cada parte é um módulo que pode ser facilmente trocado por outro, desde que sejam respeitados os parâmetros de entrada e saída esperados dentro da arquitetura. Portanto, caso em trabalhos futuros seja desejado integrar a ferramenta da Google ao WhatsApp, Facebook ou qualquer outra rede social, será possível, assim como também será possível integrar o agente twitter a qualquer outra ferramenta com caráter similar.

Figura 3 – Arquitetura do sistema



Fonte: Desenvolvida pelo autor

3.1.1 Padronização da arquitetura

Para que a solução seja mais modular, será criado um padrão de entrada e saída de cada agente. Pois assim o desenvolvimento de integrações futuras será mais simples.

3.2 Agente fact-checker

A funcionalidade deste agente será obter uma lista de artigos, de uma fonte confiável, a partir de um texto. Para que isso seja possível é necessário que este serviço consuma a API da Google. Consumindo esta API, este agente conseguira obter os artigos. Este agente também implementará uma API própria. Este agente, portanto, irá receber um

texto de entrada por sua API, consumirá a API da Google e retornará os artigos recebidos pela Google. Dessa forma, este serviço encapsula a complexidade de obter artigos a partir de um texto. Assim é possível trocar facilmente o serviço que fornece os artigos. A Nível de API, este agente implementará um método o qual receberá como entrada um texto e retornará uma lista de textos.

3.3 Agente Twitter

Este agente será responsável por controlar as ações do usuário bot, ele irá ler e responder tweets. Toda vez que o usuário for citado em um tweet, este agente irá obter o corpo do tweet e chamar o Agente fact-checker. O Agente fact-checker devolverá uma lista de artigos. O Agente Twitter então irá replicar o tweet no qual fora citado com os artigos. Caso o agente fact-checker não obtenha nenhum artigo referente a busca, ele então notificará o usuário informando-o que não encontrou nenhuma fonte. Este agente não precisará implementar nenhuma API.

REFERÊNCIAS

AMARAL, M. et al. Performance evaluation of microservices architectures using containers. In: *2015 IEEE 14th International Symposium on Network Computing and Applications*. [S.l.: s.n.], 2015. p. 27–34. Citado na página 15.

FERREIRA, A. G. *Interface de programação de aplicações (API) e web services*. [S.l.: s.n.], 2021. Citado 2 vezes nas páginas 13 e 14.

GOOGLE. *Fact Check Tools*. 2023. Disponível em: <https://toolbox.google.com/factcheck/about#fce-goal>. Citado na página 14.

GRAVES, L.; AMAZEEN, M. A. *Fact-Checking as Idea and Practice in Journalism*. Oxford University Press, 2019. Disponível em: <https://oxfordre.com/communication/view/10.1093/acrefore/9780190228613.001.0001/acrefore-9780190228613-e-808>. Citado na página 13.

LAZER, D. M. J. et al. The science of fake news. *Science*, v. 359, n. 6380, p. 1094–1096, 2018. Disponível em: <https://www.science.org/doi/abs/10.1126/science.aao2998>. Citado na página 9.

NEWMAN, N. et al. *Reuters Institute Digital News Report 2022*. [S.l.]: Reuters Institute for the Study of Journalism, 2022. Citado 2 vezes nas páginas 9 e 14.

RAHWAN, T. et al. Coalition structure generation: A survey. *Artificial Intelligence*, v. 229, p. 139–174, 2015. ISSN 0004-3702. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0004370215001198>. Citado na página 17.

WASEEM, M.; LIANG, P.; SHAHIN, M. A systematic mapping study on microservices architecture in devops. *Journal of Systems and Software*, v. 170, p. 110798, 2020. ISSN 0164-1212. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0164121220302053>. Citado na página 15.