

UNIVERSIDADE ANHANGUERA – UNIDERP

POLO TABOÃO DA SERRA

SUPERIOR DE TECNOLOGIA EM ANÁLISE EM DESENVOLVIMENTO DE
SISTEMAS

ANDERSON ALVES DO NASCIMENTO

FAKE NEWS

ARQUITETURA ORIENTADA À MICROSERVIÇOS

TUTOR JOBSON YONAH GIMENEZ

SÃO PAULO

15 DE MAIO DE 2021

ANDERSON ALVES DO NASCIMENTO

FAKE NEWS

ARQUITETURA ORIENTADA À MICROSERVIÇOS

Trabalho de projeto integrado apresentado ao curso de tecnologia em análise em desenvolvimento de sistemas da Universidade Anhanguera – Uniderp do polo Taboão da Serra com requisito parcial da conclusão do primeiro semestre em Curso tecnólogo.

Tutor: Jobson Yonaha Gimenez

SÃO PAULO

2021

Resumo

A fake News tem sido um problema que gera consequências danosas a humanidade em meio a conflitos atuais. Notícias se propagam em alta velocidade para um número desproporcional de pessoas frente a uma grande quantidade de informação, estamos perdendo a noção de falso e verdadeiro, de certo ou errado ou ainda mais, de bom ou mal.

Falsas notícias estão em toda a parte, seja no ambiente político, midiático, religião, justiça, educacional e entre outros, com isso tudo acaba se tornando muito relativo, pois as informações verídicas não tem mais base sólida o suficiente para serem transmitidas de forma segura e diferenciada através da internet.

Para solucionar essa situação ou parte dela, existem diversas pesquisas que apontam ferramentas e meios de combater fake News, utilizando a própria tecnologia, seja por aplicativo ou web, através de um processo bem planejado e estruturado para minimizar os efeitos devastadores das notícias mentirosas e eliminar suas principais causas.

Os recursos de software e hardware serão indispensáveis para o desenvolvimento das ferramentas e serviços propostos, através de uma arquitetura recomendável, por meio de sistema distribuído e descentralizado, microsserviços ordenando funções de filtrar as informações trazendo inúmeras vantagens em sua implementação. Utilizaremos também uma metodologia ágil que permita flexibilidade e organização na gestão e gerenciamento desse projeto e soluções de segurança para possíveis ataques de hackers ou vulnerabilidades do sistema.

Através destes recursos podemos aplicar meios de combater a fake News, considerando também os trabalhos correlacionados que contribuem com a implementação do projeto e que permitem ressaltar pontos relevantes nesse processo prático.

Palavras chaves: Sistemas distribuídos, microsserviços, Scrum, frameworks, PNL e chatbots.

ÍNDICE

1. Introdução	5
1.1. Fake News e o covid-19	5
1.2. Combate a fake News na TSE	6
2. Recursos de Software e Hardware para combate à fake News	6
2.1. Arquitetura recomendável	7
2.2.1. Sistema distribuído	7
2.3. Descentralização	7
2.4. Arquitetura orientada à Microserviços	8
3. Métodos e definições	9
3.1. Processamento Natural de Linguagem e <i>Chatbots</i>	10
3.2. Clusterização, Vetorização, Tokenização e Classifiers	10
4. Vantagens do sistema	13
5. Requisitos	14
6. Frameworks para detecção de fake News	14
7. Metodologia ágil	15
7.1. Scrum	16
7.2. Equipe no Scrum	17
8. Principais ataques e defesas	18
9. Conclusão	20
10. Referências bibliográficas	21

1. Introdução

Com todo avanço tecnológico conquistado, onde podemos conversar com pessoas de qualquer parte do mundo, o fake News também se torna realidade pleno século 21. Seja por vídeos, imagens, por textos jornalísticos ou via chat, as notícias falsas se propagam em alta velocidade pelas redes computacionais, causando vários problemas políticos, distorções sociais, falsas acusações com efeitos de grande proporção.

1.1. Fake news e o covid-19

Com a pandemia do novo corona vírus e o avanço de pessoas infectadas por covid-19, as fake News não deixaram de se envolver, informações equivocadas foram direcionadas até em nome de grandes fundações como a Fiocruz que tem procurado de todas as formas combater as notícias falsas alertando ainda mais sobre essa propagação de informações.

Segundo a própria fundação, as redes sociais tem sido uma das maiores propagadoras de notícias falsas sobre o novo corona vírus, sendo 10,5% por parte do Instagram, 15,8% Facebook e 73,7% atribuídas ao whatsapp. Os 26,6% das fake News do Facebook e 71,47% atribui essas notícias falsas a Fiocruz como responsável por estas informações.¹

Através do apoio da Fundação amparo a pesquisa do Rio de Janeiro (Faperj), Claudia Galhardi junto com Cecília Minayo como sua orientadora do projeto, disponibiliza um aplicativo chamado *Eu Fiscalizo*, tem como principal foco o combate à fake News que possibilita que usuários notifiquem conteúdos impróprios em veículos de comunicação. Além de registrar conteúdos inapropriados, o aplicativo informa as datas das notificações e permite o envio de fotos, vídeos e mensagens de texto. O aplicativo está disponível para download na *Playstore* e *Appstore*.

¹ Site: <https://portal.fiocruz.br/noticia/pesquisa-revela-dados-sobre-fake-newsrelacionadas-covid-19> acessado dia 20 de abril de 2021 às 06:30.

1.2. Combate a fake News na TSE

As eleições também tem sido alvo constante de fake News, pensando nisso, o Tribunal Superior Eleitoral (TSE) através de uma parceria junto ao Mensageiro, criaram um site com assistente virtual contra falsas notícias e também tiram dúvidas sobre eleições.

O site apresenta uma enquete como *Fato ou Boato?* Que tenta esclarecer os fatos de através de conteúdo verídico, além de dar opção para denuncia de contas suspeitas de realizar disparos em massa pelo whatsapp. O limite de encaminhamento de mensagens é outro fator importante para barrar a fake News, já que o compartilhamento pode ser permitido até de cinco pessoas, não sendo mais possível encaminhamento massivo de conteúdos.²

2. Recursos de Software e Hardware para combate à fake News

Os métodos para identificar e diferenciar notícias falsas das verdadeiras aumentaram de acordo com a necessidade de se combater esse mal. Mais de trezentos trabalhos vem sendo inscritos para obter meios de detecção e discernimento das informações através do processamento de algoritmos, códigos e sistema de detecção automatizado.

Trabalho como dos estudantes Jugo Queiroz e Janaina Ignácio revela destaque no meio acadêmico, onde busca diferenciar notícias legítimas das falsas e irônicas. Com uma avaliação de 85% de acerto, sendo que 94% referem-se a textos somente em português apresenta um enorme avanço contra a fake News. O foco maior tem sido conteúdo textual baseado nas palavras e também na forma escrita (estilometria).³

Pesquisadores da USP e da UFSCar criaram também ferramenta de detecção de falsas notícias que pode ser acessado via whatsapp ou web. Ainda em fase de teste, revela 90% de eficiência com base em inteligência artificial. *O detector de fake News* permite os usuários detectar a veracidade das informações através de imagens, vídeos e sons manipulados.⁴

² Site: <https://gizmodo.uol.com.br/whatsapp-tse-fake-news-eleicoes-2020/> acessado dia 20 de abril de 2021 às 06:40.

³ Site: http://www.uel.br/com/agenciaueldenoticias/index.php?arq=ARQ_not&id=28656 acessado dia 20 de abril de 2021 às 07h40min.

⁴ Site: <https://gizmodo.uol.com.br/usp-e-ufscar-criam-ferramenta-que-promete-identificar-fakenews-via-whatsapp/> acessado dia 20 de abril de 2021 às 07h43min.

2.1.Arquitetura recomendável

Compreender alguns recursos de software e hardware necessários para uma arquitetura de computadores que pode combater a fake news de forma proeminente e aplicável é fundamental. No entanto é preciso compreender que um bom funcionamento de um sistema requer uma estrutura que alicerça e forneça condições para um desempenho otimizado.

A partir da compreensão do sistema distribuído, da descentralização, dos microsserviços e seus recursos, este projeto precisa de uma metodologia ágil que atenda o combate a fake news e disponibilize boas recomendações de segurança de informação.

2.2.1. Sistema distribuído

Um sistema distribuído já é realidade, mesmo que não aparente ele funciona em diversas aplicações, passando uma impressão de sistema único e aberto, como navegadores, redes sociais, streamings, entre outros. Segundo Tanenbaum um sistema distribuído é um conjunto de computadores que são interligados via rede, porém, para o usuário final das aplicações que são executadas através deles, aparenta ser um sistema único, como uma única máquina ou único software⁵.

Este sistema traz grandes vantagens quanto ao vasto acesso de usuários às informações. No entanto a rede também pode sofrer ataques de diversos meios, principalmente pelas notícias falsas que são propagadas em alta velocidade.

2.3. Descentralização

O sistema distribuído contém uma arquitetura descentralizada que é colaborativa, onde os todos os computadores conectados a rede trocam informações e não provem de uma hierarquia centralizada, onde um único servidor transmita informações para os computadores interligados em rede⁶. A arquitetura descentralizada é comum em vários ambientes virtuais e se apresenta também como referência no que se refere à segurança das informações, como a rede de criptomoedas com o sistema de blockchain, os computadores conectados validam as

⁵ TANENBAUM; STEEN, 2017. Citado por Caique Silva Pereira no livro sistema distribuído unidade 3 seção 1 <https://www.colaboraread.com.br/integracaoAlgetec/> acessado dia 15 de abril de 2021 às 06h39min.

⁶Gomes, Roberta Lima. Arquiteturas distribuídas para sistemas colaborativos. Site: <https://sistemascolaborativos.uniriotec.br/wp-content/uploads/sites/18/2019/06/SC-cap20-arquiteturas.pdf>. Acessado dia 15 de abril de 2021 às 07h09min.

transações e atualizações ocorridas em rede. Para serem realizadas, é necessário que 50%+1 dos integrantes concorde que a ação é legítima, pois do contrário ela não pode ser concretizada.⁷

2.4. Arquitetura orientada à Microserviços

Um sistema distribuído contém computadores com diferentes configurações, apresentando dessa forma uma interoperabilidade no sistema, ou seja, os computadores conectados a rede podem ter arquiteturas e sistemas operacionais diferentes entre si, fazendo com que o sistema seja heterogêneo.

Para um desenvolvimento eficiente de um sistema é necessário agilidade e flexibilidade para atender as expectativas em meio a tantas mudanças internas e externas no processo de construção do software ou hardware. Com isso, a comunicação entre os DevOps também precisa acompanhar o nível em que esses eventos mudam intensivamente. As trocas de informações no processo de desenvolvimento acabam sendo comprometidas quando o sistema operacional e outros impedimentos interferem no andamento dos projetos, como naquela frase típica ‘mas no meu não funcionou’, assim podemos trabalhar com meios que fornecem soluções para esse tipo de situação.

A containerização é um processo onde é possível encapsular um aplicativo em um recipiente com seu próprio ambiente operacional, promovendo um sistema mais leve e ágil. Cada computador pode colocar sua aplicação em um contêiner, que fica isolado do sistema operacional da máquina de cada DevOps. Assim é possível uma comunicação eficiente que proporciona uma arquitetura orientada a microserviços.

Através de Caique Silva Pereira podemos ter uma melhor compreensão do conceito de microserviços e suas principais características:

Os sistemas distribuídos fazem uso extensivo dos contêineres no contexto de microserviços. A ideia dos microserviços está associada a empresas que possuem sistemas altamente dinâmicos e ao termo modularidade. Um portal de notícias por exemplo, é composto por vários elementos, como um (ou mais) banco de dados, um (ou mais) framework front-end utilizado para desenvolver interfaces, framework back-end para desenvolver a parte dinâmica do sistema, frameworks para gerenciar mensagens entre servidores e clientes.⁸

⁷ Site: <https://cointimes.com.br/o-que-e-blockchain-como-funciona/> acessado dia 15 de abril de 2021 às 07h16min.

⁸ Pereira, Caique Silva. Sistema distribuído unidade 4 seção 2.

<https://www.colaboraread.com.br/integracaoAlgetec/> acessado dia 15 de abril de 2021 às 08h00min.

Os microsserviços permitem um maior dinamismo o que facilita a comunicação de todos envolvidos. Outro conceito de microsserviço é fornecido por Laura Betti Monteiro que enfatiza a containerização das aplicações e a comunicação entre serviços request e response:

A arquitetura de microsserviços são serviços pequenos e autônomos que funcionam juntos. São entidades separadas, de maneira que devem ser implementados como serviços isolados em plataformas isoladas ou com o seu próprio sistema operacional. A comunicação entre os serviços ocorre entre chamados de *request* e *response*.⁹

A comunicação dos serviços pode ser *síncrona*, quando há uma necessidade de aguardar uma resposta, porém é preciso lidar com a questão da latência onde ela fica fortemente acoplada ao endpoint requerido. Ou *assíncrona*, onde não depende de aguardar uma resposta que faz com que a troca de mensagens seja rápida e com baixo acoplamento.¹⁰

3. Métodos e definições

Estes recursos podem funcionar como se fossem filtros de informação que ficam entre as informações vindas dos servidores até o dispositivo do usuário. Também se aplicam ao método como ferramentas importantes para caracterizar a utilização dos microsserviços em contêineres. Laura Betti propõe os seguintes métodos para o projeto: As redes neurais, Processamento Natural de Linguagem (PNL) e *Chatbots*, e as subtarefas que se aplicam aos microsserviços como a Clusterização, Vetorização, Tokenização e Classifiers.

Redes Neurais A Inteligência Artificial (IA) é uma área da ciência da computação que surge logo após a Segunda Guerra Mundial. Pode ser definida em quatro categorias diferentes, sendo elas: pensando como um humano; pensando racionalmente; agindo como seres humanos e agindo racionalmente. Em cada uma destas categorias existem modelos representativos práticos e teóricos sendo respectivamente: a abordagem do teste de Turing, a estratégia de modelagem cognitiva, a abordagem das leis do pensamento e a abordagem de agente racional. Atualmente é uma área interdisciplinar e abrange diversos subcampos principalmente *machine learning* e redes neurais. Tem como objetivo a criação e desenvolvimento de técnicas computacionais e a construção de *softwares* que sejam capazes de aprender de maneira independente. Redes Neurais é um desdobramento da área da inteligência artificial e podem ser definidas como uma máquina desenvolvida para replicar a maneira como o cérebro humano realiza determinada

⁹ Monteiro, Laura Betti. Barion, Michele Cristiani . Ferreira, Adriano. Arquitetura de *machine learning* para análise de reportagens textuais em redes sociais para a detecção de *fake news*. Pag.4. São Paulo-SP. 31 de março de 2020.

¹⁰ Site: <https://medium.com/avecnews/comunicacao-entre-microservices-async-ed3e5897ba6> acessado dia 15 de abril de 2021.

tarefa. Esta rede pode ser implementada através de componentes eletrônicos ou através de um algoritmo de programação.¹¹

Com a contribuição de Alan Turing, as redes neurais contam um método dinâmico que contempla mais de perto a cognição humana, chegando a uma possível compreensão de atitudes que são exclusivamente do ser humano.

3.1.Processamento Natural de Linguagem e *Chatbots*

Alan Turing, o pai da computação moderna, quando criou o teste de Turing (1950), afirmou que para que o computador fosse capaz de passar no teste ele deveria possuir as seguintes capacidades: processamento de linguagem natural, aprendizado de máquina, representação do conhecimento e raciocínio automatizado [Russell e Norvig 2013]. O Processamento Natural de Linguagem (PNL) é um conjunto de técnicas variadas da computação para o processamento de textos, sejam eles orais ou escritos, com a finalidade de obter um processamento de linguagem para uma variedade de tarefas e aplicativos.

Já *chatbots* podem ser conceituados como *softwares* que interagem com usuários utilizando linguagem natural. Seu objetivo é simular uma conversação humana e sua arquitetura normalmente integra um modelo de linguagem associado a um algoritmo computacional para emular a comunicação entre usuário e máquina.

A PNL e a chatbots tem buscado uma maior proximidade da linguagem humana e seus diversos desdobramentos na comunicação através de emulação.

3.2.Clusterização, Vetorização, Tokenização e Classifiers

Através da PNL cientistas e especialistas propuseram subdividir tarefas principais de compreensão da linguagem em subtarefas como a clusterização, a vetorização e a tokenização e a classifiers:

Clusterização é o agrupamento de objetos similares. Objetos que diferem em detalhes insignificantes recebem o mesmo nome e podem ser tratados da mesma forma. As principais funções da clusterização são: nomear, exibir, resumir e prever. Desta maneira, todos os objetos que estão dentro de um mesmo *cluster* recebem o mesmo nome.¹²

¹¹ Monteiro, Laura Betti. Barion, Michele Cristiani . Ferreira, Adriano. Arquitetura de *machine learning* para análise de reportagens textuais em redes sociais para a detecção de *fake news*. Pag.5. São Paulo-SP. 31 de março de 2020.

¹² Monteiro, Laura Betti. Barion, Michele Cristiani . Ferreira, Adriano. Arquitetura de *machine learning* para análise de reportagens textuais em redes sociais para a detecção de *fake news*. Pag.5 e 6.São Paulo-SP. 31 de março de 2020.

Esse modo de classificação organiza as palavras compreendendo sua semelhança, relacionando entre elas como iguais, facilitando mais adiante uma análise aprofundada da informação transmitida.

Vetorização é a distribuição representativa de dados em vetores a fim de auxiliar algoritmos de aprendizado a obter uma melhor performance agrupando itens similares.

Outro modo de agrupamento buscando referência de outras semelhanças.

Tokenização é o processo de dividir blocos de textos em frases, palavras, símbolos ou elementos significativos chamados de *tokens* [Verma *et al.* 2014]. Seu principal objetivo é explorar palavras dentro de contextos frasais para identificação de palavras-chaves.

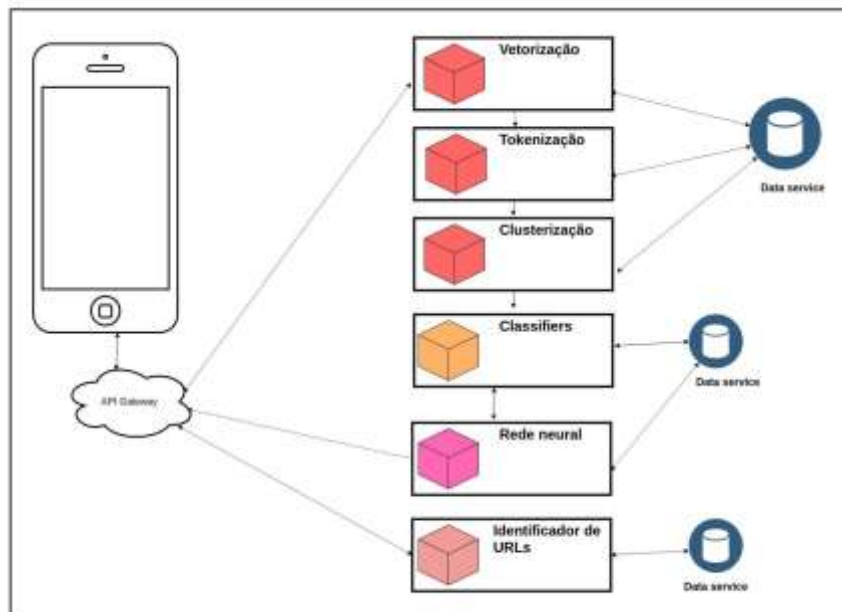
Essa busca tem relação com blocos maiores de referência semântica, está atrelado à sua semelhança ou diferença, compreendendo palavras-chaves como parte de um contexto.

Classifiers são algoritmos de *machine learning* que mapeiam dados e os classificam de acordo com categorias específicas. A classificação preditiva é a tarefa de aproximação de uma função de mapeamento (f) de variáveis de *input* (x) para variáveis de output (y).

Machine learning é o aprendizado da máquina, uma área da ciência da computação que permite automatizar respostas ao usuário a partir de inteligência artificial e big data.

Essas subtarefas têm funções específicas que agem sob uma estrutura de microsserviços que pode fazer o mapeamento das informações, analisando diferentes aspectos textuais.

A arquitetura orientada a microsserviços, conforme apresentada na Figura 6, realiza as seguintes operações:



Para uma melhor compreensão das subtarefas aplicáveis aos microserviços podemos dividi-la em etapas: ¹³

1. O *front-end* faz requisições para a *API gateway*, dependendo de qual informação for enviada - *URL* ou trecho textual - diferentes microserviços serão acionados.
2. Caso seja inserido uma *URL*, a *API gateway* irá requisitar via serviço *REST*, o serviço de identificador de *URLs* que irá acessar seu respectivo banco de dados.
3. Será executada uma *query*, caso a *URL* esteja no banco é devolvido para o usuário a informação de que este *site* não é uma fonte de notícias confiáveis.
4. Uma vez que o usuário insere um trecho textual, o primeiro serviço a ser chamado é o de vetorização. Para ser classificado pela *machine learning* os dados precisam ser transformados em vetor.
5. Logo em seguida a informação é tratada de acordo com os métodos PNL. Assim, se houver necessidade serão utilizados os métodos de tokenização e clusterização. Isto se dá ao fato de muitas vezes existirem palavras no trecho inserido que já foram tokenizadas anteriormente.
6. Por fim os dados serão classificados e enviados para a rede neural que irá fazer a classificação de veracidade da notícia.

¹³ Monteiro, Laura Betti. Barion, Michele Cristiani . Ferreira, Adriano. Arquitetura de *machine learning* para análise de reportagens textuais em redes sociais para a detecção de *fake news*. Pag.11 a 13. São Paulo-SP. 31 de março de 2020.

4. Vantagens do sistema

Os principais parâmetros de uma arquitetura em microsserviços é que pode ser avaliada. Segundo o site do Red Hat essas vantagens estão de acordo com sua rapidez, escalabilidade, resiliência, implementação, acessibilidade e open source.

Lançamento no mercado com mais rapidez: Como os ciclos de desenvolvimento são reduzidos, a arquitetura de microsserviços é compatível com implantações e atualizações mais ágeis.

Altamente escalável: À medida que a demanda por determinados serviços aumenta, você pode fazer implantações em vários servidores e infraestruturas para atender às suas necessidades.

Resiliente: Os serviços independentes, se construídos corretamente, não afetam uns aos outros. Isso significa que, se um elemento falhar, o restante da aplicação permanece em funcionamento, diferentemente do modelo monolítico.

Fácil de implementar: Como as aplicações baseadas em microsserviços são mais modulares e menores do que as aplicações monolíticas tradicionais, as preocupações resultantes dessas implantações são invalidadas. Isso requer uma coordenação maior, mas as recompensas podem ser extraordinárias.

Acessível: Como a aplicação maior é decomposta em partes menores, os desenvolvedores têm mais facilidade para entender, atualizar e aprimorar essas partes. Isso resulta em ciclos de desenvolvimento mais rápidos, principalmente quando também são empregadas as tecnologias de desenvolvimento ágil.

Mais open source: Devido ao uso de APIs políglotas, os desenvolvedores têm liberdade para escolher a melhor linguagem e tecnologia para a função necessária.¹⁴

Através dos recursos disponíveis conseguimos uma flexibilidade maior ao atender a demanda relacionada à comunicação podendo usufruir das vantagens dos microsserviços. Podemos destacar pontos relevantes para maior compreensão dos parâmetros apresentados pelo Red Hat correlacionando com as observações de Laura Betti.

O uso da arquitetura orientada a microsserviços permite uma melhor definição dos domínios do artefato. Esta arquitetura permite que sejam utilizadas, se necessário, diferentes linguagens de programação para escrever cada um dos serviços. Permite a **escalabilidade** de forma individual, ou seja, é possível escalar determinados serviços quando necessário. Possui alto nível de **agilidade** uma vez que qualquer defeito na aplicação afeta apenas um serviço, desta maneira as mudanças e correções acontecem de maneira que não afetam a aplicação inteira.

Na arquitetura orientada a microsserviços existe uma complexidade extra, uma vez que microsserviços é um sistema distribuído e cabe ao desenvolvedor escolher e criar as conexões entre os módulos e os bancos de dados. A arquitetura é um sistema complexo de múltiplos sistemas e banco de dados, desta maneira todas as conexões devem ser tratadas com cuidado. Quando é criada uma aplicação com microsserviços podem ocorrer problemas com configuração, métricas, *health checks*

¹⁴ Site: <https://www.redhat.com/pt-br/topics/microservices/what-are-microservices> acessado dia 17 de abril de 2021 às 09h44min.

entre outros componentes. Como é uma arquitetura composta por componentes diferentes, criar testes *end-to-end* pode ser algo complexo.¹⁵

Essas funcionalidades permitem uma melhor utilização dos recursos quando se refere a um desenvolvimento atrelado a mudanças constantes, um sistema desse porte estará apto a melhorar a comunicação de todos envolvidos na rede, diminuindo drasticamente as distorções quando se trata de falsas notícias.

5. Requisitos

O sistema tem um modo de operar de modo criterioso de acordo com o tipo de plataforma, segurança e lei que o protege, banco de dados e serviços.

- O artefato deve funcionar nas plataformas *web* e *mobile*;
- O artefato deve ser seguro, a proteção de dados de usuário deve estar em conformidade com a Lei Geral de Proteção de Dados (LGPD)¹⁶
- O *dataset* de notícias deve estar em formato *JSON* (*JavaScript Object Notation*) e deve ser armazenado em um banco de dados NoSQL;
- O processo de aprendizagem da *Machine Learning* deve ser protegido com camadas de segurança;
- A conexão entre *back-end* e *front-end* deve ser feita via serviços *REST* ;
- A base de dados com *URLs* deve ser um banco de dados, podendo ser relacional, NoSQL ou uma tabela do Excel.¹⁷

Todos esses procedimentos estão diretamente ligados as subtarefas dos microsserviços, com requisitos que possibilita um método mais transparente.

6. Frameworks para detecção de fake News

Framework é um conjunto de códigos prontos que podem ser usados no desenvolvimento de aplicativos e sites. Um framework distribuído poderia ser utilizado no desenvolvimento de sistema de detecção de *fake News*.

Laura Betti Monteiro propõe principais frameworks distribuídos orientados à microsserviços e detecção de fake News que podemos considerar.¹⁸

¹⁵ Monteiro, Laura Betti. Barion, Michele Cristiani . Ferreira, Adriano. Arquitetura de *machine learning* para análise de reportagens textuais em redes sociais para a detecção de *fake news*. Pag.12. São Paulo-SP. 31 de março de 2020.

¹⁶ LGPD é uma lei que demanda a proteção e confidencialidade de dados sensíveis. Disponível em Site: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.html acessado dia 15/04/2021 às 10h04min.

¹⁷ Monteiro, Laura Betti. Barion, Michele Cristiani . Ferreira, Adriano. Arquitetura de *machine learning* para análise de reportagens textuais em redes sociais para a detecção de *fake news*. Pag.9. São Paulo-SP. 31 de março de 2020

Desenvolvimento do *Front-End* com Angular Framework que funciona em múltiplas plataformas, escalável e permite ao desenvolvedor um maior controle sob o código.¹⁹

Clusterizar *Dataset K-means*, o método *k-means* é utilizado para a quantificação de vetores. O algoritmo tem como objetivo particionar *n* observações em *k clusters* de tal maneira que, cada observação pertence ao *cluster* com o significado mais próximo.

Vetorizar *dataset word2vec*. Este modelo é uma rede neural de duas camadas treinada para reconstruir o contexto linguístico das palavras.²⁰

Remover Stopwords do dataset NLTK. A biblioteca é implementada como uma coletânea independente de módulos, onde cada um define os tipos básicos de dados a serem processados.²¹

Implementar Rede neural com *TensorFlow*. É um sistema de *Machine Learning* que opera escalado e em ambientes heterogêneos.²²

7. Metodologia ágil

Os métodos ágeis são ferramentas para gerenciamento e desenvolvimento de softwares, considerado *métodos lights* em contraposição aos métodos tradicionais ou “métodos pesados”, como era chamado pelos engenheiros de software. O método ágil é caracterizado pela rapidez, qualidade, interação com o cliente e tem como benefícios o ganho de produtividade e diminuição de custos.

O método cascata precisava seguir rigorosamente as etapas previstas no ciclo de vida, como a iniciação, planejamento, controle e monitoração, execução e por fim, encerramento. Se caso ocorresse algum problema durante as etapas, ficava difícil de aplicar ao processo de adaptações em relação às mudanças. Com um contexto de modificações constantes, seria necessário um método mais flexível que valorizasse mais as pessoas ao invés de ferramentas, algo mais dinâmico e simples que seja aplicável sem muitas burocracias.

Em 2001 ocorre o manifesto ágil, uma reunião com dezessete pessoas que buscavam compreender as mudanças que viam ocorrendo no processo da elaboração de projetos e

¹⁸ Monteiro, Laura Betti.Barion, Michele Cristiani . Ferreira, Adriano. Arquitetura de *machine learning* para análise de reportagens textuais em redes sociais para a detecção de *fake news*. Pag.14 .São Paulo-SP. 31 de março de 2020.

¹⁹ Site: <https://angular.io/> acessado dia 18 de abril de 2021 às 12:25.

²⁰ Site: <https://pathmind.com/wiki/word2vec> acessado dia 18 de abril de 2021 às 12:28.

²¹ Site: http://www.nltk.org/howto/portuguese_en.html acessado dia 18 de abril de 2021 às 12:29.

²² Site: <https://www.tensorflow.org/> acessado dia 18 de abril de 2021 às 12:31.

pretendiam padronizar nova metodologia que fosse aplicável de acordo com a necessidade vigente de um método ágil.

Esse método compromete-se com doze princípios que norteiam todo o projeto sendo eles:

- Nossa maior prioridade é satisfazer o cliente por meio da entrega contínua e adiantada de software com valor agregado.
- Mudanças nos requisitos são bem-vindas, mesmo tardiamente, no desenvolvimento. Processos ágeis tiram vantagem das mudanças visando à vantagem competitiva para o cliente.
- Entregar, frequentemente, software funcionando, de poucas semanas a poucos meses, preferindo a menor escala de tempo.
- Pessoas de negócio e desenvolvedores devem trabalhar diariamente em conjunto durante todo o projeto.
- Construa projetos em torno de indivíduos motivados. Dê a eles o ambiente e o suporte necessários e confie neles à realização do trabalho.
- O método mais eficiente e eficaz de transmitir informações para e entre uma equipe de desenvolvimento é por meio de conversa face a face.
- Software funcionando é a medida primária de progresso.
- Os processos ágeis promovem desenvolvimento sustentável. Os patrocinadores, desenvolvedores e usuários devem ser capazes de manter um ritmo constante indefinidamente.
- Contínua atenção a excelência técnica e ao bom design aumenta a agilidade.
- Simplicidade - a arte de maximizar a quantidade de trabalho não realizado é essencial.
- As melhores arquiteturas, os melhores requisitos e designs emergem de equipes organizadas.
- Em intervalos regulares, a equipe reflete a respeito de como se tornar mais eficaz e, então, refina e ajusta seu comportamento.²³

Todos os métodos ágeis são condizentes com a filosofia ágil, como o Extreme programming (XP), Scrum, DSDM, FDD, Cristal clear, ASD, Lean, Design Thinking e o MPV, podem ser aplicados de acordo com a necessidade do cliente e a situação interna e externa em relação à empresa.

Podemos notar que devido à evolução da tecnologia, esses processos se tornaram complementares um do outro em ótimas ferramentas de trabalho.

7.1.Scrum

Scrum é um processo de desenvolvimento iterativo para gerenciamento de projetos, usando para trabalhos complexos nos quais é impossível prever tudo o que irá ocorrer. Segundo o site *Devmedia* o **Scrum** é uma metodologia ágil de trabalho onde é usada para estabelecer conjuntos de regras e práticas de gestão para conseguir o sucesso de um projeto.

²³ Moraes, Izabelly Soares. Site: <https://www.colaboraread.com.br/integracaoAlgetec/> . Projeto de Software. Unidade 2, seção 1. Acessado dia 18 de abril de 2021 às 15:01.

Com o foco no trabalho em equipe, ocorre uma melhora na comunicação e maximiza o apoio de todos, fazendo com que todo o time se esforce e se sinta seguro com que estão fazendo e isso acaba gerando posteriormente um aumento na produtividade.²⁴

As principais características do SCRUM são:

- É um processo ágil para gerenciar e controlar o desenvolvimento de projetos;
- É um wrapper para outras práticas de engenharia de software;
- É um processo que controla o caos resultante de necessidades e interesses conflitantes;
- É uma forma de aumentar a comunicação e maximizar a cooperação;
- É uma forma de detectar e remover qualquer impedimento que atrapalhe o desenvolvimento de um produto;
- É escalável desde projetos pequenos até grandes projetos em toda empresa.²⁵

Podemos ver nas características do Scrum a coerência em relação aos princípios norteadores da filosofia ágil proposta no manifesto em 2001. Principalmente no que se refere à cooperação, comunicação, foco em interesses e necessidades.

7.2. Equipe no scrum

Uma equipe Scrum consiste em uma equipe de seis a dez funcionários, não incluindo o product owner ou o Scrum máster, trabalham juntos em um projeto enquanto aderem a processos e práticas Scrum. Embora o tipo de projeto e seus requisitos ditarão o tamanho da equipe Scrum, o objetivo é manter a equipe suficientemente reduzida para permanecer ágil — e grande o suficiente para completar todo o trabalho descrito no Sprint.

Quanto à estrutura da equipe scrum, ela não são hierárquicas, mesmo que tenhamos líderes envolvidos, cada membro é reconhecido igualmente como importante e essencial. Cada um colabora no projeto com suas respectivas habilidades, experiências e conhecimentos necessários para concluir o projeto.

As equipes Scrum são projetadas para serem flexíveis e criativas com foco na produtividade. Valorizando o feedback, a autorreflexão e avaliação contínua durante o projeto, maximando seu potencial.²⁶

²⁴ Site: <https://www.devmedia.com.br/agile-development-xp-e-scrum-em-uma-abordagem-comparativa/30808> acessado dia 14 de abril de 2021 às 13:00.

²⁵ Site: <https://www.devmedia.com.br/agile-development-xp-e-scrum-em-uma-abordagem-comparativa/30808> acessado dia 14 de abril de 2021 às 13:00.

²⁶ Site: <https://monday.com/blog/pt/como-criar-uma-equipe-scrum-bem-sucedida/>. Acessado dia 14/04/2021 às 13:08.

8. Principais ataques e defesas

Com todo o avanço tecnológico e facilidade que internet proporciona a todos os usuários, o sistema distribuído requer uma maior atenção quanto as suas vulnerabilidades que precisam ser consideradas, um lado negativo que precisa de monitoramento constante, os ataques proporcionado por hackers que podem explorar vulnerabilidades e falhas do sistema.

Podemos listar os principais ataques que o software e o hardware podem sofrer e quais as principais defesas que há para protegê-lo. Os pontos de atenção em relação à segurança no projeto de sistema distribuídos são:

Portas são expostas: Sistemas distribuídos são construídos com base em um conjunto de processos que oferecem serviços e compartilham informação. As portas de comunicação nas quais esses serviços se comunicam são, intrinsicamente, abertas (para que clientes possam acessar tais serviços) e, dessa forma, um hacker pode enviar mensagem a qualquer uma dessas portas.

Redes de computadores não são seguras: Remetentes de mensagens podem ser falsificados, ou seja, um e-mail enviado por caique@caique.com pode não ter sido enviado pelo Caique; endereços IP podem estar duplicados, de forma que alguém malicioso possa receber as mesmas mensagens de um destinatário válido, etc.

A validade das chaves criptográficas deve ser limitada: Quanto mais tempo uma mesma chave estiver válida e ativa, maiores são as chances de esta estar comprometida, por ter maiores chances de ser conhecida (e explorada) por uma quantidade maior de pessoas e sistemas.

Algoritmos de criptografia podem ter falhas: nos algoritmos de criptografia podem ter falhas: na atualidade, a melhor prática é de divulgar publicamente os algoritmos de criptografia para que a comunidade e entidades especialistas possam validar o algoritmo e sugerir melhorias, de forma que a privacidade esteja garantida pela chave criptográfica, e não pela inacessibilidade ao algoritmo utilizado.

Hackers podem ter acesso a recursos poderosos: O custo dos recursos computacionais tem diminuído cada vez mais, de forma que máquinas poderosas estão acessíveis para a maioria da população. Assim, certifique-se de considerar que ataques podem ocorrer de inúmeras fontes, e que podem explorar vulnerabilidades utilizando inclusive ataques do tipo força-bruta (que tentam descobrir senhas por tentativa e erro, através de simples “chutes”).²⁷

Os sistemas distribuídos podem ser afetados por essas invasões e conseqüentemente os microsserviços podem ser prejudicados com ações danosas que abusam de permissões do sistema. Segundo o Westcon-comstor, a arquitetura de microsserviços ainda se difere dos mecanismos tradicionais de segurança, consegue minimizar a superfície de ataque e permitir

²⁷Pereira, Caique Silva. Sistema distribuído unidade 3 seção3.

<https://www.colaboraread.com.br/integracaoAlgetec/> acessado dia 18 de abril de 2021 as 16:06.

uma maior segmentação dos componentes de aplicação. Entretanto, esse mecanismo é mais difícil de monitorar através de ferramentas tradicionais de segurança.

Algumas providências podem ser tomadas, as propostas podem ser classificadas com pontos de atenção direcionados a microsserviços:

1 - Proliferação de endpoints: cada instância do serviço ainda é um endpoint de uma rede exclusiva. Embora a superfície de ataque da aplicação não seja mais concentrada em servidores isolados, aumenta de acordo com a propagação de serviços, pois muito mais portas estão abertas, APIs expostos e a autenticação se tornou apenas uma questão de distribuição.

2 - IPs efêmeros e perímetros indefiníveis: contêineres podem ser movidos em questão de segundos, assim a noção de perímetro está desaparecendo. Além disso, IPs dinâmicos e escalonamento de serviços não são mais viáveis para configurar IPs estáticos ou para orientar o tráfego de perímetros de segurança de aplicativos tradicionais.

3 - Entropia e mascaramento do contêiner: contêineres aumentam a entropia da infraestrutura, porque podem ser orquestradas de forma dinâmica em todos os serviços, ou nuvens. Outro desafio é que o endereço IP de um contêiner pode ser invisível fora de onde é executado, assim a configuração de segurança em uma aplicação externa torna-se um problema, já que as fontes de tráfego não podem ser identificadas.

4 - Arquiteturas de microsserviços e contêineres: possui facilidade de implantar novos serviços e instâncias dinâmicas de serviços, mas pode complicar a definição de políticas de segurança, caso os administradores não saibam as regras que regem as suas interações. Além disso, o aumento de serviços no interior do Data Center ou na nuvem significa que mais transações requerem interações entre os serviços para serem cumpridas.

5 - Os princípios de DevOps (amálgama de Desenvolvedor e Operações - que é a reação à interdependência entre desenvolvimento de software e operações de TI): são conceitos promissores para a linha de negócio, tanto para desenvolvedores como para administradores de sistemas, pois assim, os desenvolvedores conseguem manter o foco na sua atividade principal, enquanto os times de operações se beneficiam da flexibilidade, de menores áreas ocupadas no Data Center e de custos mais baixos. Porém, processos e ferramentas de segurança ainda estão se desenvolvendo para acompanhar a agilidade desse desenvolvimento.

Outras práticas podem ainda ser recomendadas no tocante a segurança para microsserviços:

Monitorar as comunicações entre serviços: os administradores de segurança, muitas vezes, não sabem as interações permitidas entre os diferentes serviços nas aplicações de microsserviços. Por isso, precisam configurar políticas de segurança que governam as interações com base nas instruções de implantação dos aplicativos dos desenvolvedores.

Para garantir a segurança, deve-se visualizar o tráfego entre os serviços, para entender o comportamento da aplicação e ser possível estabelecer as bases da política. Assim, identificam-se automaticamente novos serviços na aplicação. Além de ajudar as equipes de segurança a notarem padrões de comunicação anômalos ou identificar possíveis violações.

Segmentação de serviços isolados: a segmentação para trabalhar em aplicações de microsserviços deve aplicar políticas no nível de cluster em todas as instâncias de serviço, uma vez que a direção do tráfego baseado no perímetro não é prático.

Criptografia dos dados em trânsito e estáticos: realizar a criptografia do tráfego entre as aplicações ou serviços que compõem uma aplicação para atender aos requisitos de conformidade e para melhorar a segurança, especialmente quando o tráfego atravessa redes públicas.

Automatizar o gerenciamento de políticas: para minimizar erros humanos, é ideal

a automatização de tarefas, tais como a definição de políticas que podem ser replicadas em todos os aplicativos, separando funções de desenvolvimento das administrativas, e gerenciamento de certificados SSL e chaves.²⁸

Com essas práticas podemos beneficiar os usuários com a qualidade e segurança dos serviços prestados para a utilização dessas ferramentas, visando uma arquitetura que facilite o processo comunicativo e a veracidade das informações sem interferências alheias que compromete a evolução e eficácia do software.

9. Conclusão

É preciso ter consciência que as falsas notícias se faz pertinente às relações humanas, seja pelo uso da tecnologia como meio de propagação ou não, porém não precisamos nos conformar com essa atual situação, para isso, o sistema requer o uso de ferramentas que proporcione constante monitoração, controle e proporcione qualidade e segurança aos usuários.

Com o sistema distribuído, sendo descentralizado e orientado a microsserviços, já são meios que estão sendo utilizados e testados em diferentes ambientes que se enquadram nos requisitos necessários e nas boas práticas para seu funcionamento.

Aliado a uma metodologia ágil, com frameworks adequados, buscando minimizar os riscos e ataques cibernéticos, os microsserviços tem se apresentado como forma mais adequada de se combater a fake News. Essa arquitetura com todas suas ferramentas proporcionam uma quantidade de filtros o suficiente para dificultarem de a falsa informação ser propagada com encaminhamento massivo de conteúdos.

Estamos em constante evolução, a tecnologia aumenta conforme a necessidade ou situação vigente e se torna obsoleta quando os problemas são totalmente superados, pensando assim, estaremos em constante inovação na ampliação e implantação de recursos tecnológicos.

²⁸Site: <https://blogbrasil.westcon.com/seguranca-em-arquiteturas-de-microservicos> acessado dia 14 de abril de 2021 as 13:16.

10. Referências bibliográficas

Agile Development: XP e Scrum em uma Abordagem Comparativa. Devmedia,2019. Disponível no Site: <https://www.devmedia.com.br/agile-development-xp-e-scrum-em-uma-abordagem-comparativa/30808>. Acessado dia 14 de abril de 2021 às 13h00min.

Arquitetura de microserviços. Red Hat, 2021. Disponível no Site: <https://www.redhat.com/pt-br/topics/microservices/what-are-microservices>. Acessado dia 17 de abril de 2021 às 09h44min.

Bassotto, Lucas. O que é blockchain? Como funciona a tecnologia. Coint times,2018. Disponível no Site: <https://cointimes.com.br/o-que-e-blockchain-como-funciona/>. Acessado dia 15 de abril de 2021 às 07h16min.

Botelho, Bia. Pesquisadores elaboram método para identificar fake News. Agência UEL de notícias, 2019. Disponível em Site: http://www.uel.br/com/agenciaueldenoticias/index.php?arq=ARQ_not&id=28656. Acessado dia 20 de abril de 2021 às 07h40min.

Como criar uma equipe Scrum bem-sucedida. Monday português, 2020. Disponível no Site: <https://monday.com/blog/pt/como-criar-uma-equipe-scrum-bem-sucedida/> Acessado dia 14 de abril de 2021 às 13h08min.

Examples for Portuguese Processing. Howto, 2020. Disponível no Site: http://www.nltk.org/howto/portuguese_en.html. Acessado dia 18 de abril de 2021 às 12h29min.

Gonzaga, Rafael. Comunicação entre microserviços: Async. Avec News, 2019. Disponível no Site: <https://medium.com/avecnews/comunicacao-entre-microservices-async-ed3e5897ba6>. Acessado dia 15 de abril de 2021.

Gomes, Roberta Lima. Arquiteturas distribuídas para sistemas colaborativos.

Site: <https://sistemascolaborativos.uniriotec.br/wp-content/uploads/sites/18/2019/06/SC-cap20-arquiteturas.pdf>. Acessado dia 15 de abril de 2021 às 07h09min.

LGPD é uma lei que demanda a proteção e confidencialidade de dados sensíveis. Planalto, 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.html. Acessado dia 15 de abril de 2021 às 10h04min.

Monteiro, Laura Betti. Barion, Michele Cristiani . Ferreira, Adriano. Arquitetura de *machine learning* para análise de reportagens textuais em redes sociais para a detecção de *fake news*. Curso Superior de Tecnologia em Análise e Desenvolvimento de Sistemas-Instituto Federal de Educação, Ciência e Tecnologia de São Paulo (IFSP). São Paulo-SP. Março, 2020.

Moraes, Izabelly Soares. Projeto de Software. Ebook, 2021. Disponível no Site: <https://www.colaboraread.com.br/integracaoAlgetec/> . Acessado dia 18 de abril de 2021 às 15:01.

Nicholson, Chris. A Beginner's Guide to Word2Vec and Neural Word Embeddings. Pathmind, 2020. Disponível no Site: <https://pathmind.com/wiki/word2vec>. Acessado dia 18 de abril de 2021 às 12h28min.

Pereira, Caique Silva. Sistema distribuído. E-book, 2021. Disponível no site: <https://www.colaboraread.com.br/integracaoAlgetec/>. Acessado dia 15 de abril de 2021 as 06h39min.

Pesquisa revela dados sobre 'fake News' relacionadas à Covid-19. Portal Fiocruz, 2020. Disponível em Site: <https://portal.fiocruz.br/noticia/pesquisa-revela-dados-sobre-fake-newsrelacionadas-covid-19>. Acessado dia 20 de abril de 2021 às 06h30min.

Segurança em arquiteturas de microsserviços. CANAL SYNnex WESTCON-COMSTOR, 2020. Disponível no Site: <https://blogbrasil.westcon.com/seguranca-em-arquiteturas-de-microservicos>. Acessado dia 14 de abril de 2021 às 13h16min.

The modern web developer's platform. Angular, 2021. Disponível no Site: <https://angular.io/>. Acessado dia 18 de abril de 2021 às 12h25min.

Uma plataforma completa de código aberto para machine learning. Tensorflow. Disponível no Site: <https://www.tensorflow.org/>. Acessado dia 18 de abril de 2021 às 12h31min.

USP e UFSCar criam ferramenta que promete identificar fake News via Whatsapp. Gismodo Brasil, 2018, disponível no Site: <https://gizmodo.uol.com.br/usp-e-ufscar-criam-ferramenta-que-promete-identificar-fakenews-via-whatsapp/>. Acessado dia 20 de abril de 2021 às 07:43.

WhatsApp e TSE lançam assistente contra fake News e tirar dúvidas sobre eleições. Gismodo Brasil, 2020. Disponível no Site: <https://gizmodo.uol.com.br/whatsapp-tse-fake-news-eleicoes-2020/>. Acessado dia 20 de abril de 2021 às 06h40min.