**1 INTRODUÇÃO**

A área de ciência de dados se faz cada dia mais útil para o cotidiano do brasileiro, pois é com ela que o usuário da internet recebe sugestões baseadas em seus gostos e hábitos online, desde os vídeos recomendados que aparecem em sua página inicial do Youtube, rotas seguras e rápidas no seu aplicativo de carona, até as propagandas mostradas para ele em redes sociais baseadas nos seus dados, documentos e conversas.   
 É verdade que com essa ciência é possível promover e realizar tarefas humanamente impossíveis ou extremamente penosas com uma facilidade e eficácia muito maior, como o supercomputador Watson da IBM que usa inteligência artificial para identificar o melhor tratamento para cada tipo de câncer, os carros da marca estadunidense Tesla que dirigem sozinhos e identificam pessoas, objetos e animais na pista, ou até mesmo os foguetes das bilionárias SpaceX e Blue Origin que são capazes de pousar verticalmente em solo terrestre após de uma expedição espacial.  
 No entanto, o acúmulo de dados para o aprendizado de máquina e inteligência artificial também pode ser usado de forma desonesta e apelativa para um determinado objetivo, seja ele econômico, social ou geopolítico. Para esse último, temos como exemplo um caso recente que ganhou atenção dos holofotes mundiais: Após a vitória do republicano Donald Trump contra a democrata Hillary Clinton nas eleições para presidente dos EUA no ano de 2016, a imprensa estadunidense, em paralelo com diversos grupos de ciberativistas, descobriram uma série de provas que comprovaram que ambos os candidatos (Donald Trump com maior destaque) haviam utilizado bots russos para disparos de mensagens em favor de suas próprias propostas, além de também direcionar ofensas e notícias falsas relacionadas aos seus oponentes políticos [[[1]](#footnote-0)][[[2]](#footnote-1)].  
  
 Mesmo sendo evidente, é importante salientar que o uso de bots é feito de maneira criteriosa. A maioria dos bots para disparo de mensagens, como os do caso exposto acima, são programados para simular humanos e atingir de maneira mais orgânica possível o maior número de pessoas de um público específico. Dessa forma, o uso da ciência de dados se faz presente para identificar um padrão de perfil dos possíveis eleitores de cada candidato com base em suas informações na rede social, como quais outros perfis o usuário segue, que tipo de conteúdo que ele compartilha entre seus seguidores, em quais contas e páginas o usuário é mais ativo, entre outros. Soma-se tal característica à análise do comportamento que o candidato possui na rede social, como frequência de postagens e teor apelativo das postagens. Como resultado obtemos um bot para disparo de mensagens muito específico e eficaz em sua tarefa, podendo ser facilmente confundido com uma pessoa na rede, assim como ocorreu no caso de exemplo supracitado.  
  
 O fato é que, após as eleições mencionadas anteriormente, o uso de da ciência de dados para construção de bots se tornou peça-chave de praticamente toda eleição, seja ela em âmbito nacional ou internacional. Com isso, o tema vem tomando cada vez mais espaço nos debates a respeito do poder de influência direta e indireta que a internet, e, sobretudo, as redes sociais têm sobre a opinião de seus usuários e como isso têm afetado perigosamente o resultado final das urnas.

**2 CONCEITOS GERAIS**

Esta seção apresenta os principais conceitos cujo entendimento é de suma importância para que se compreenda o presente trabalho e o algoritmo em si. Tais conceitos são relacionados à navegação por páginas da internet de maneira automatizada e extração e mineração de textos pertencentes às mesmas.

**2.1 Crawling**

O crawling, também conhecido como spidering, é uma técnica muito comum dentro de testes automatizados cujo principal objetivo é capturar e navegar pelos links existentes em determinada página web. Seu uso é de vital importância quando pretende-se extrair um link, navegar por uma cascata de páginas similares entre si, ou apenas indexá-las, como os sistemas de busca Google, Bing, Yahoo fazem.

**2.2 Scraping**

O scraping é uma prática que combina uma gama de técnicas de programação e tecnologias a fim de extrair dados e informações geralmente encontradas em páginas web. Seu uso é extremamente comum quando se pretende realizar uma análise de dados posterior à captura dessas informações, ou ainda, quando se deseja construir uma base de dados grande o suficiente para aplicação de conceitos de ciência de dados, como aprendizado de máquina, treinamento de inteligências artificiais, entre outros.   
 Apesar de ser relativamente popular e difundida na comunidade de data science, essa prática deve ser realizada com cautela, pois muitos sites proíbem ou restringem esse uso em seus domínios, o que gerar penas online ou até processos judiciais [[[3]](#footnote-2)].

**3 OBJETIVO GERAL**

Dada tal premissa, a presente pesquisa possui o afinco de identificar o uso de Inteligência Artificial como parte da estratégia das campanhas eleitorais dos candidatos a governador do Estado de São Paulo, no segundo turno das eleições ocorridas em 2018, João Dória (pelo PSDM) e Márcio França (pelo PSB).

**4 OBJETIVOS ESPECÍFICOS DO ALGORITMO**

Com o algoritmo, um dos objetivos consiste em capturar todos os tweets dos candidatos supracitados nos períodos oficiais de campanha eleitoral, acrescentando-se uma pequena margem temporal antes e depois da data inicial e final, respectivamente.  
 Além disso, o referido algoritmo também possui uma segunda função de coletar notícias relevantes para a pesquisa nos principais veículos de comunicação do estado de São Paulo e do Brasil relacionadas às seguintes palavras-chave: “Inteligência Artificial”, “Fake News”, “Eleições SP 2018”, além dos próprios nomes dos candidatos. O intervalo de datas adotado na captura das notícias seguiu os mesmos critérios adotados no scraping dos tweets, dessa forma, o período permaneceu o mesmo.  
  
 Por fim, com a obtenção do conteúdo dos tweets de ambos candidatos para a análise de temáticas mais abordadas e apelo emocional presente no texto, juntamente com as notícias capturadas nos principais veículos de comunicação do país relacionadas às estratégias eleitorais de cada político mencionado anteriormente, pretende-se chegar à uma conclusão razoável que evidencie o uso, ou a ausência dele, de tecnologias relacionadas à inteligência artificial, machine learning, big data, bots, entre outras vertentes incluídas no ramo da ciência de dados como parte da estratégia da campanha eleitoral de João Dória Jr. e Márcio França.

**5 MATERIAIS E FERRAMENTAS**

Nesta seção, estão contidos as principais ferramentas e softwares utilizados em maior ou menor grau para construção do algoritmo e, consequentemente, para a seleção e mineração de textos. Bem como os métodos e critérios de seleção e extração de tweets e notícias.

**5.1 Python 4.6**

O Python é uma linguagem de programação criada em 1990 multi-paradigma cujo uso abrange desde pequenos projetos e scripts para uso pessoal até softwares grandes e complexos, como parte dos serviços do Youtube, Instagram, Spotify, entre muitos outros. Devido sua rapidez e simplicidade, além do seu vasto acervo de bibliotecas, API’s e frameworks de código aberto e gratuitos para uso, atualmente essa linguagem é uma das principais na comunidade científica da computação, isto é, ela é comumente usada para algoritmos e aplicações que envolvem aprendizagem de máquina, inteligência artificial, bots, big data, deep learning e outras vertentes correlatas.

**5.2 SQL 2016**

O SQL, acrônimo de “Structured Query Language”, ou ainda, do português, “Linguagem de Consulta Estruturada”, é uma linguagem de programação voltada para a criação, manipulação e exclusão de banco de dados relacionais, modelo extremamente difundido na academia e no mercado.  
 Embora que nos últimos anos esse modelo tenha perdido espaço mercadológico para os chamados bancos “NoSQL” (bancos não relacionais) e modelos orientados a objetos, ele ainda se apresenta muito perspicaz na facilidade e dinamismo na criação, manipulação e alocação de tabelas (entidades), atributos e bases de dados.

**5.3 MySQL 8**

O MySQL é um SGBD (Sistema de Gerenciamento de Banco de Dados) de código aberto criado em 1995 e atualmente mantido pela empresa Oracle Corporation. Graças à sua versão gratuita denominada Community e seu foco no público de programadores iniciantes e estudantes na área de tecnologia, segundo a pesquisa de 2019 realizada pelo site StackOverflow - maior fórum de programadores do mundo – o MySQL lidera o ranking de popularidade com mais da metade de todos os usuários que utilizam banco de dados para quaisquer fins, superando concorrentes de peso como o SQL Server da Microsoft e o Firebase, mantido atualmente pelo Google [[[4]](#footnote-3)].

**5.4 Anaconda 3**

A plataforma open source Anaconda 3 reúne uma série de softwares, bibliotecas e ferramentas úteis para a programação científica a fim de facilitar a integração de diferentes aplicativos e oferecer recursos avançados que não se encontrariam em programas convencionais de programação. No caso do algoritmo dessa pesquisa, foi utilizada a IDE (Integrated Development Environment) Spyder, responsável por reunir uma série de pacotes e módulos úteis para o processo de crawling e scraping de dados, assim como facilitar a visualização dos resultados obtidos e reduzir a complexidade dos códigos.

**5.5 Selenium WebDriver 3.141.39**

O Selenium é uma ferramenta de automatização de testes funcionais em navegadores de internet, como o Google Chrome e Mozilla Firefox. Apesar de ser escrito em Java, ele suporta a integração com, além de sua própria linguagem mãe, o JavaScript, Ruby, C#, PHP, Perl e o Python, que é o que se utilizou para o desenvolvimento do crawler da pesquisa. Em síntese, ele é responsável por navegar por links de forma automática sempre executando as funções e comportamentos pré-programados. Para a pesquisa, o Selenium WebDriver serviu apenas para o crawling, ou seja, tanto para entrar nos sites dos veículos de comunicação definidos, como para capturar o título e link das notícias que atendiam todos os parâmetros e pré-requisitos.

**5.6 Tweepy 3.8.0**

Por questões legais que se adequam aos termos de uso da rede social do Twitter, o único meio que se pode realizar o scraping de dados dos usuários de sua plataforma, como os tweets dos candidatos já citados durante toda a pesquisa, é utilizando a própria API (Application Program Interface) da rede, chamada de Tweepy. Basicamente, essa ferramenta fornece aos programadores da linguagem Python devidamente licenciados no programa Twitter Developer uma série de funcionalidades para a captura dos dados pretendidos, de forma mais intuitiva e simples do que seria feito – por meios ilegais ou demasiadamente restritos – com scrapers de terceiros não autorizados.

**5.7 Outros pacotes Python: NLTK, RE e PYMYSQL**

O pacote NLTK (versão 3.4.5), acrônimo de “Natural Language Toolkit”, é escrito em Python e possui o propósito de realizar processamentos qualitativos e quantitativos de textos submetidos a ele pelo programador. Com essa ferramenta, é possível efetuar-se diversos tratamentos em palavras, desde a retirada de termos que não possuem sentido sozinhos (stopwords) do corpo do texto, até uma análise sintética de emoções presentes no discurso do seu locutor. Essa biblioteca foi de suma importância para a obtenção de resultados advindos dos tweets dos candidatos, facilitando a análise para a posterior conclusão.

O pacote RE (nativo do Python) é destinado para se trabalhar com as “Regular Expressions”, isto é, ele é o responsável por identificar e separar as palavras considerando apenas os valores alfanuméricos do texto. Na realidade, esse pacote é extremamente mais amplo e complexo do que o descrito acima, existindo em praticamente todas as linguagens de programação contemporâneas, mas, para a construção do algoritmo deste trabalho, apenas este recurso fora usado.

Por fim, o PYMYSQL (versão 0.7.2) é uma das bibliotecas responsáveis por conectar a aplicação escrita em Python com o banco de dados do MySQL, permitindo que os dados obtidos tanto pelo crawler quanto pelo scraper sejam salvos e passíveis à visualização mesmo após o encerramento da aplicação.

**6 METODOLOGIA, CRITÉRIOS, PRÉ-REQUISITOS E LIMITES DE PESQUISA**

**6.1 Critérios de busca de notícias**

Conforme já mencionado na seção de objetivos específicos, objetivou-se buscar e capturar o link de notícias relevantes para a pesquisa através das palavras-chave “Inteligência Artificial”, “Fake News”, “Eleições SP 2018”, “João Dória” e “Márcio França”, no veículos de comunicação G1, Folha de São Paulo, Estadão e O Globo, levando em consideração o período que abrange o dia 01 de agosto de 2018 ao dia 01 de novembro de 2018. O período foi escolhido pelo fato de abranger o período oficial das campanhas eleitorais (15/08/2018, data máxima para registro da candidatura, até 29/10/2018, resultado das eleições de 2º turno) com uma pequena margem pré e pós eleições. Para evitar dados muito difusos ao objetivo da pesquisa, o crawler limitou-se à captura das notícias presentes até a 5ª página de resultados de cada site. Dessa forma, quanto mais próximo da primeira paǵina, mais relevante é o conteúdo para a combinação de palavras-chave impostas na busca.  
 Por fim, é importante evidenciar que em determinados veículos a busca por notícia adquiriu um caráter mais generalista por não retornar nenhuma notícia ou matéria quando todas as palavras-chave eram consideradas, dessa forma, os sites e suas respectivas buscas ficaram como mostra a tabela a seguir:

Tabela 1 - Correlação de sites e termos buscados.

|  |  |
| --- | --- |
| **Sites** | **Termos buscados** |
| Estadão | “João Dória” e “Márcio França” |
| Folha de São Paulo | “João Dória”, “Márcio França”, “Fake News” e “Eleições SP 2018” |
| G1 | “João Dória”, “Márcio França”, “Fake News”, “Eleições SP 2018” e “Inteligência Artificial” |
| O Globo | “João Dória”, “Márcio França” e “Fake News” |

Fonte: Elaboração do autor.

**6.2 Critérios de captura dos tweets**

A captura de tweets só se limitou ao período de datas que abrange o dia 01 de agosto de 2018 ao dia 01 de novembro do mesmo ano. Ademais, tendo em vista que o uso desse tipo de conteúdo só se justifica para quantificar a frequência das temáticas mais abordadas por cada candidato e analisar a presença de emoção nos seus textos, somente a extração do conteúdo textual dos tweets se fez necessária, sem levar em conta imagens, gifs e ou vídeos anexados no mesmo.

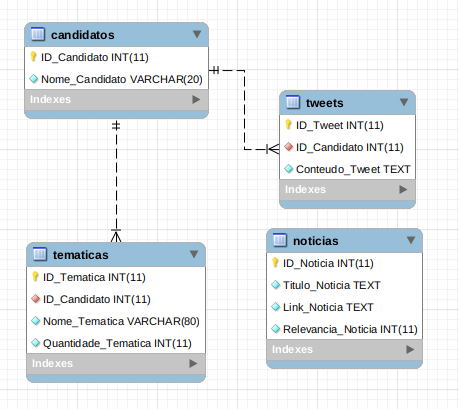
**6.3 Limites da pesquisa**

Por via de regra, é importante salientar todas as limitações que o algoritmo possui ao realizar a busca de notícias e a captura do texto dos tweets de ambos os candidatos.  
 A priori, é de suma importância que o leitor entenda que a responsabilidade e o domínio sobre o conteúdo postado na rede social do Twitter fica a cargo unica e exclusivamente do autor dos tweets (exceto quando o mesmo infringe alguma regra da rede social), logo, da mesma forma que o algoritmo não foi capaz de recuperar tweets potencialmente excluídos ou editados no período das campanhas eleitorais, há a possibilidade de o leitor realizar o mesmo processo de crawling e scraping e obter resultados que não se assemelham de maneira integral aos obtidos e analisados nesta pesquisa.  
 Além disso, o mesmo se aplica às notícias capturadas para análise. É de total responsabilidade dos veículos de comunicação manter todas as notícias, matérias, artigos e textos livres e disponíveis para visualização de qualquer usuário da internet.  
 Também se faz necessário explanar que as técnicas e métodos trabalhados para a obtenção, mineração, limpeza e análise dos dados foram escolhidos para que o prazo de 12 (doze) meses de duração de bolsa fosse respeitado. Notavelmente há diversas outras técnicas e métodos mais eficazes que os escolhidos, no entanto, devido sua complexidade, são implementações que certamente excederiam o prazo dado para o bolsista e orientadora.

**7 SOLUÇÃO DESENVOLVIDA E RESULTADOS**

Após uma revisão de pesquisas correlatas e sugestões por parte dos próprios docentes do curso de Sistemas de Informação da Escola de Artes, Ciências e Humanidades da Universidade de São Paulo, foi decidido que a melhor solução para que os objetivos da presente pesquisa fossem alcançados, levando em conta as limitações já expostas no documento, seria a construção de um banco de dados relacional no MySQL para a indexação das notícias obtidas pelo crawler e para a alocação dos tweets dos dois candidatos capturadas pela API do Tweepy. As entidades, relacionamentos e atributos da base de dados estão expressas no MER (Modelo Entidade Relacionamento) a seguir, seguindo todos os padrões de modelagem impostos pela UML (Unified Modeling Language).

Imagem 1 - MER do banco de dados construído na pesquisa

  
Fonte: Elaboração do autor.

Apenas para caráter expositivo e didático, temos que as entidades são estruturas onde guardamos os atributos de determinado ator. Esses atributos, por sua vez, são características que serão armazenadas no banco de dados para posterior análise ou manipulação, como a “Titulo\_Noticia” da entidade “noticias”. Ao lado dos atributos das entidades, temos o que chamamos de “tipo” do atributo, onde se explicita que tipo de dado aquela variável espera receber. Como na entidade tweets, na qual seu atributo “ID\_Tweet” mostra que só permitirá a inserção de valores do tipo inteiro, isto é, só armazenará números inteiros. Por fim, temos as chaves estrangeiras e primárias, que possuem o objetivo de interligar e criar dependências entre entidades. As chaves primárias são acompanhas de um ícone de chave amarela e as estrangeiras de um ícone de losango vermelho. Os traços observados entre as diferentes entidades representam que tipo de relacionamento existe entre as mesmas, de forma que a ponta acompanhada por dois riscos paralelos representa o número 1 (um) e a ponta acompanhada por três riscos formando uma pirâmide representa o número n. Isso significa que, por exemplo, a entidade “candidato” possui uma relação de 1 para n (1:n) com a entidade tweets, ou seja, para cada candidato, podem existir vários tweets, mas que o contrário não é válido [[[5]](#footnote-4)][[[6]](#footnote-5)].

O código python que foi desenvolvido para esse trabalho encontra-se hospedado na plataforma do GitHub [[[7]](#footnote-6)], visando auxiliar o leitor da presente pesquisa a entender todo o processo de desenvolvimento, além de contribuir para a expansão da comunidade open source. Todos os direitos e termos de uso seguem a licença MIT de open source [[[8]](#footnote-7)].

Realizado o processo de crawling das notícias presentes nos veículos de comunicação citados anteriormente atendendo os pré requisitos e critérios também respaldados por esse texto, obtivemos os seguintes resultados:

Tabela 2 - Resultados do processo de crawling de notícias

|  |  |
| --- | --- |
| **Site de notícias** | **Quantidade de notícias correlatas com a busca\*** |
| Estadão | 60 |
| Folha de São Paulo | 35 |
| G1 | 90 |
| O Globo | 39 |

Fonte: Elaboração do autor com base nos resultados retornados do crawler  
**\* No período entre 01/08/2018 e 01/11/2018**

No total, foram retornadas 224 notícias que associam-se, em maior ou menor grau, com as palavras chave dispostas na seção 6.1 deste trabalho. É importante reiterar que a quantidade retornada foi a soma de todas as notícias, matérias e artigos presentes até a 5ª página de resultados de cada site, evitando o acúmulo de textos muito destoantes aos objetivos principais da pesquisa. A página na qual determinada notícia foi encontrada é representada pelo atributo “Relevancia\_Noticia” da entidade “noticias” disposta em nossa base de dados (vide Imagem 1). Podendo variar de 1 a 5, o quão mais próxima a relevância está do número 1, maior a chance da notícia vinculada a ela ser relevante para a combinação de palavras chave buscadas no site.

Já o processo de scraping dos perfis de João Dória e Márcio França na rede social do Twitter feito pela API Tweepy, nos retornou os seguintes resultados:

Tabela 3 - Resultados retornados pelo processo de scraping de tweets

|  |  |
| --- | --- |
| **Candidato** | **Quantidade de tweets feitos\*** |
| João Dória Jr. (PSDM) | 636 |
| Márcio França (PSB) | 1071 |

Fonte: Elaboração do autor baseado em resultados obtidos com o scraper do Tweepy  
**\* No período entre 01/08/2018 e 01/11/2018**

Em suma, o algoritmo desenvolvido conseguiu capturar e alocar em nossa base de dados um total de 1707 tweets. Vale lembrar que tweets em resposta a outros usuários da rede, em resposta à menção do user (nome que acompanha o caractere @) e retweets também foram inclusos nessa conta.  
 Além do conteúdo textual em si, nosso algoritmo também capturou a quantidade de respostas, curtidas e retweets que cada tweet recebeu com o afinco de obter alguma referência de engajamento existente entre os candidatos e os outros usuários da rede. Considerando que as curtidas, retweets e respostas possuem o mesmo peso de engajamento, após o cálculo da média simples de tais atributos de todos os tweets de cada candidato, obtivemos os seguintes números:

Tabela 4 - Média simples de engajamento do público com as postagens de cada candidato

|  |  |
| --- | --- |
| **Candidato** | **Quantidade média de interações por tweet** |
| João Dória Jr. (PSDM) | 73.9 |
| Márcio França (PSB) | 75.3 |

Fonte: Elaboração do autor com base nos resultados obtidos pelo scraper

Ademais, o algoritmo desenvolvido também foi capaz de capturar as palavras mais repetidas durante o período eleitoral no perfil de cada um dos políticos, excluindo as chamadas stopwords (vide sessão 5.7) e outras palavras semelhantes que não são inclusas nativamente nesse pacote (a lista de palavras inclusas pelo algoritmo pode ser vista no link do GitHub já disponibilizado anteriormente). Com a lista dessas palavras, podemos chegar a uma conclusão mais sucinta de temáticas mais abordadas por candidato e se as mesmas tiveram alguma relação com o uso de bots para disparo de mensagens automáticas. Os termos mais mencionados por cada candidato estão dispostos em ordem decrescente na tabela a seguir.

Tabela 5 - Termos mais abordados por cada candidato no período analisado da pesquisa

|  |  |
| --- | --- |
| **João Dória** | **Márcio França** |
| “Estado” (123 vezes mencionado) | “Estado” (152 vezes mencionado) |
| “População” (111 vezes mencionado) | “Apoio” (139 vezes mencionado) |
| “SP” (107 vezes mencionado) | “Juntos” (117 vezes mencionado) |
| “Gestão” (88 vezes mencionado) | “Certo” (87 vezes mencionado) |
| “DefendaSP” (75 vezes mencionado) | “Vida” (77 vezes mencionado) |

Fonte: Elaboração do autor com base nos dados obtidos pelo scraper

Foi decidido que o algoritmo fosse capaz de quantificar todas as palavras de todos os tweets capturados, dessa forma, o script descobriu por si só as palavras e temáticas mais abordadas por cada candidato, visando não se limitar a uma série de termos que poderiam ser pré estabelecidos e provocassem uma potencial parcialidade nos resultados finais. Apenas para caráter expositivo, a seguir temos temáticas que são normalmente associadas às campanhas de cada político e são de interesse popular.

Tabela 6 - Temáticas populares mais abordadas por cada candidato durante as campanhas eleitorais

|  |  |
| --- | --- |
| **João Dória** | **Márcio França** |
| “Saúde” (36 vezes mencionada) | “Interior” (44 vezes mencionada) |
| “Segurança” (34 vezes mencionada) | “Segurança” (39 vezes mencionada) |
| “Empregos” (29 vezes mencionada) | “Saúde” (37 vezes mencionada) |
| “Investimentos” (19 vezes mencionada) | “Educação” (33 vezes mencionada) |
| “Interior” (18 vezes mencionada) | “Jovens” (26 vezes mencionada) |

Fonte: Elaboração do autor baseada nos dados obtidos pelo scraper

Por fim, usando a biblioteca Natural Language Toolkit (NLTK), o algoritmo realizou uma análise sintética de emoções presente nos tweets, com isso, pretendeu-se identificar se o conteúdo postado na rede social do Twitter de cada político é dotado de viés apelativo para convencimento, aproximação do eleitor ou afins, pois se caso essa última característica se mostrar presente, a tendência de que o conteúdo em questão foi gerado por um bot aumenta significativamente.

**8 CONCLUSÕES**

1. O’Sullivan, Donie. Russian bots retweeted Trump nearly 500,000 times in final weeks of 2016 campaign. **CNN**, Nova York, 27 de jan. de 2018. Disponível em: <https://money.cnn.com/2018/01/27/technology/business/russian-twitter-bots-election-2016/>. Acesso em 25 de jan. de 2020. [↑](#footnote-ref-0)
2. Mayer, Jane. How Russia helped swing the election for Trump. **The New Yorker**, Nova York, 24 de set. de 2018. DIsponível em: <https://www.newyorker.com/magazine/2018/10/01/how-russia-helped-to-swing-the-election-for-trump/>. Acesso em 25 de jan. 2020. [↑](#footnote-ref-1)
3. WEB SCRAPING e Web Crawling são legais ou ilegais? **Data Science Academy**, 01 de jul de 2019. Disponível em <http://datascienceacademy.com.br/blog/web-scraping-e-web-crawling-sao-legais-ou-ilegais/>. Acesso em 07 de mar. de 2020. [↑](#footnote-ref-2)
4. DEVELOPER Survey Results. **StackOverflow**, 2019. Disponível em <https://insights.stackoverflow.com/survey/2019#technology/>. Acesso em 20 de jan. de 2020. [↑](#footnote-ref-3)
5. Joel. MER e DER: Modelagem de Banco de Dados. **DevMedia**, 2014. Disponível em <https://www.devmedia.com.br/modelo-entidade-relacionamento-mer-e-diagrama-entidade-relacionamento-der/14332>. Acesso em 14 de mar. de 2020. [↑](#footnote-ref-4)
6. ELMASRI, R.; NAVATHE, S.B. Sistemas de Bancos de Dados, 6ta. Edição. Editora Pearson, 2011. [↑](#footnote-ref-5)
7. Disponível em <https://github.com/dan-santos/crawler> [↑](#footnote-ref-6)
8. Disponível em <https://mit-license.org/> [↑](#footnote-ref-7)