

**UNIVERSIDADE VIRTUAL DO ESTADO DE SÃO PAULO**  
**CURSO DE BACHARELADO EM CIÊNCIA DE DADOS**

**CARLOS GOMES BARBOZA FILHO, 2002910**

**ERIK DANIEL FONSECA DE JESUS, 2003945**

**JOÃO ROBSON RAMALHO, 2011198**

**PATRÍCIA RODRIGUES DA SILVA, 2009767**

**RENATO APARECIDO TEREZAN DE MOURA, 2000514**

**ANÁLISE DE DESCRITORES DE PRODUTOS**  
**COMERCIALIZADOS POR *E-COMMERCE***

INSERIR O LINK DO VÍDEO AQUI

**SÃO PAULO**  
**NOVEMBRO / 2023**

**UNIVERSIDADE VIRTUAL DO ESTADO DE SÃO PAULO**  
**CURSO DE BACHARELADO EM CIÊNCIA DE DADOS**

**CARLOS GOMES BARBOZA FILHO, 2002910**

**ERIK DANIEL FONSECA DE JESUS, 2003945**

**JOÃO ROBSON RAMALHO, 2011198**

**PATRÍCIA RODRIGUES DA SILVA, 2009767**

**RENATO APARECIDO TEREZAN DE MOURA, 2000514**

## **ANÁLISE DE DESCRITORES DE PRODUTOS COMERCIALIZADOS POR *E-COMMERCE***

Pré-projeto de pesquisa submetido como requisito parcial para aprovação na disciplina Trabalho de Conclusão de Curso do Curso de Bacharelado em Ciência de Dados da Universidade Virtual do Estado de São Paulo.

**SÃO PAULO**

**NOVEMBRO / 2023**

**UNIVERSIDADE VIRTUAL DO ESTADO DE SÃO PAULO**  
**CURSO DE BACHARELADO EM CIÊNCIA DE DADOS**

**CARLOS GOMES BARBOZA FILHO, 2002910**

**ERIK DANIEL FONSECA DE JESUS, 2003945**

**JOÃO ROBSON RAMALHO, 2011198**

**PATRÍCIA RODRIGUES DA SILVA, 2009767**

**RENATO APARECIDO TEREZAN DE MOURA, 2000514**

**ANÁLISE DE DESCRITORES DE PRODUTOS**  
**COMERCIALIZADOS POR *E-COMMERCE***

Trabalho de Conclusão de Curso aprovada, em \_\_\_\_/\_\_\_\_/2023 pela banca formada:

---

Helton Silva Bernardo  
Pós-Graduação em Engenharia Elétrica,  
Universidade Estadual Paulista “Júlio de Mesquita Filho” UNESP  
(Orientador Acadêmico)

---

Leo Natan Paschoal  
Mestre em Ciências da Computação e Matemática Computacional,  
Universidade de São Paulo, USP  
(Avaliador Externo)

2023

# **Sumário**

Sumário

**1 Introdução1**

**2 Objetivos3**

2.1 Objetivos gerais3

2.2 Objetivos específicos3

**3 Justificativas4**

**4 Revisão bibliográfica5**

**5 Metodologia12**

**6 Cronograma20**

**7 Resultados e discussão21**

**8 Conclusão22**

**9 Referências23**

**10 Anexos4**

## 1 Introdução

A expressão “compras online”, segundo SCHLEMMER & PADOVANI (2013), significa o processo que permite aos clientes adquirirem produtos e serviços através de uma plataforma eletrônica. Até a década de 1970, o comércio era realizado principalmente em ambientes físicos, como compras presenciais em lojas, acesso porta a porta e ligações telefônicas. Porém, é importante lembrar que o comércio eletrônico, ou o conceito de comércio eletrônico, foi criado antes da popularização da *World Wide Web* (WWW).

O comércio eletrônico foi inventado em 1979 por Michael Aldrich, um inglês que era funcionário da empresa britânica *Rediffusion Computers*. Aldrich desenvolveu um sistema inovador de compras online utilizando uma televisão modificada conectada a uma linha telefônica e um computador adaptado para transações em tempo real. O método de interação utilizado na época era o chamado de teletexto.

O termo “e-commerce” é uma abreviatura de “*electronic commerce*”, que se traduz como “comércio eletrônico”, sendo um conceito amplamente aplicado com a popularização da internet, no qual todo o processo de compra e venda foi bastante simplificado. No início, o comércio eletrônico era usado principalmente para vender produtos menores, como CDs, DVDs e livros. No entanto, ao longo do tempo, o seu âmbito expandiu-se significativamente para incluir a venda de bens de elevado valor, tais como carros, casas, iates, aviões, obras de arte e vários artigos de luxo.

Segundo TURAN (2012), o e-commerce engloba uma ampla variedade de modelos de negócios, abrangendo desde sites voltados diretamente para consumidores finais até plataformas de leilões, serviços e organizações. Uma das características mais marcantes do e-commerce é a capacidade de permitir que os consumidores realizem transações eletronicamente a qualquer momento, independentemente do dia, horário ou localização geográfica. Essa flexibilidade horária e geográfica tem sido fundamental para o rápido crescimento desse setor desde a sua criação e sugere um potencial futuro para superar as vendas convencionais (MENDONÇA, 2016).

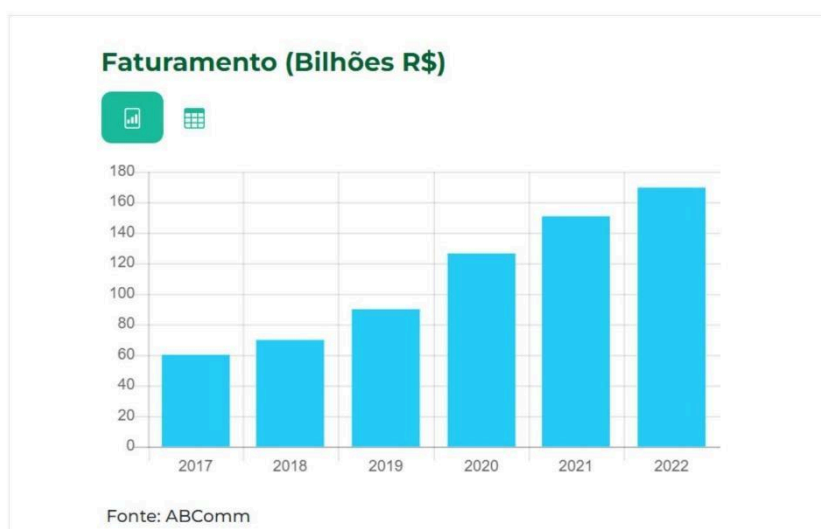
Uma notável vantagem desse modelo de negócios é que ele capacita pequenos comerciantes a alcançar uma vasta audiência de consumidores, independentemente de estarem localizados próximos ou distantes geograficamente.

Essa capacidade de alcance é algo que uma loja física tradicional dificilmente conseguiria igualar, tornando o comércio eletrônico uma opção altamente atrativa para empreendedores de pequeno porte. Simplificando, o comércio eletrônico representa uma grande mudança nas práticas comerciais, permitindo transações eletrônicas em cenários altamente dinâmicos e flexíveis (MENDEZ, 2013). A sua trajetória de crescimento constante desde a sua criação demonstra que o comércio eletrônico tem o potencial para redefinir o panorama das vendas e do comércio e moldar de forma fundamental e sustentável as interações comerciais no século XXI.

A revolução digital mudou profundamente a forma como o comércio é conduzido, com o comércio eletrônico emergindo como um dos setores mais dinâmicos e influentes. Neste contexto, a personalização é fundamental para satisfazer as expectativas dos consumidores e otimizar os resultados dos negócios do comércio eletrônico. Um dos mecanismos básicos para essa personalização é um algoritmo baseado em linguagem natural, que é alimentado por descritores de produtos. Este estudo propõe analisar o papel dos descritores de produtos no comércio eletrônico e os desafios e implicações associados à sua utilização em termos de personalização da experiência do usuário.

#### Faturamento do E-commerce no Brasil

Dados da **ABComm Forecast**, uma pesquisa realizada pela Associação Brasileira de Comércio Eletrônico, também relatam um **faturamento de R\$169,59 bilhões do e-commerce no Brasil em 2022**, uma diferença de mais de 18 bilhões em relação ao faturamento de 2021 (150.82).



Para 2023 a expectativa é ainda maior, com **faturamento estimado de R\$ 185,7 bilhões** podendo chegar até R\$ 232.51 bilhões, em 2026, uma diferença de mais de R\$ 100 bilhões se comparado ao faturamento de 2020 (126.45).

Figura 1 - Crescimento do faturamento de e-commerce no Brasil nos últimos 6 anos

Fonte: <https://edrone.me/pt/blog/dados-ecommerce-brasil>

## **2 Objetivos**

### **2.1 Objetivo geral**

O objetivo principal é demonstrar a importância de manter os descritivos de produtos no *e-commerce* estrategicamente atualizados de maneira que possam ser facilmente identificados por suas principais características e diferenciais, buscando melhorar a experiência e o engajamento do consumidor.

### **2.2 Objetivos específicos**

Na primeira etapa da pesquisa os objetivos específicos serão:

1. Identificar situações reais onde a falta de um descritivo estratégico contribui para uma má experiência por parte dos consumidores.
2. Explorar os varejos mais relevantes e ferramentas de SAC (Serviço de Atendimento ao Consumidor) de terceiros como o site “Reclame aqui” para localizar as reclamações que poderiam ter sido evitadas
3. Explicar como um descritivo mal executado pode atrapalhar,
4. Mostrar como podemos construir um descritivo mais estratégico e mais abrangente a respeito de como os clientes das indústrias possam encontrar seus produtos.
5. Utilizaremos a ferramenta digital Google Trends para demonstrar como as pessoas pesquisam por um produto e relacionar com como as marcas pensam que os clientes buscam seus produtos.



### 3 Justificativa

Nos últimos anos, temos visto um progresso significativo nos cenários de vendas eletrônicas impulsionados principalmente por algoritmos que orientam os usuários através do conteúdo apresentado. Este estudo analisa o papel fundamental dos descritores de produtos na experiência do usuário no comércio eletrônico.

O motivo do nosso estudo é para mostrar como as marcas podem atender melhor seus clientes finais utilizando recursos avançados de análise de dados em larga escala. A ideia principal é mostrar o valor que pode ser gerado ao mapear e ajustar oportunidades de descritivos incorretos no *e-commerce* para evitar que a indústria e o consumidor final sejam prejudicados por um problema de comunicação ou até mesmo que o consumidor não encontre o produto desejado em seu *e-commerce* de confiança por falta de padronização estratégica no cadastro dos produtos.

Um famoso caso de problemas em descritores de produtos comercializados eletronicamente é o da empresa Samsung lançou o Galaxy Note 7 que, em 2016, sem fornecer informações claras sobre os problemas de superaquecimento da bateria vendeu milhares de unidades do seu produto. Os consumidores compraram sem saber dos riscos, resultando em uma série de incêndios e recalls massivos, causando danos à reputação da empresa e milhares de dólares em prejuízos aos



consumidores.

Figura 2 - Esclarecimentos sobre problemas crônicos na bateria do Note 7

Fonte: <https://arstechnica.com/gadgets/2017/01/galaxy-note-7-investigation-blames-small-battery-cases-poor-welding/>

## 4 Revisão Bibliográfica

A constatação de que as descrições dos produtos afetam diretamente as vendas e os lucros do vendedor no comércio eletrônico, remonta aos primórdios do próprio comércio online. Desde o surgimento da Internet comercial nas décadas de 1990 e 2000, tornou-se claro que as descrições dos produtos desempenham um papel importante na persuasão dos consumidores e na tomada de decisões. Segundo STEININGER (2019) essas evoluções ao longo do tempo dividiram o comércio eletrônico em três períodos criticamente diferentes que apresentamos a seguir:

Nos estágios iniciais do comércio eletrônico (entre os anos 1990 e 2000), as descrições dos produtos costumam ser curtas e sem detalhes. Os profissionais de *marketing* estão mais interessados em construir presença online e disponibilidade de produtos do que em melhorar a descrição. No entanto, rapidamente se tornou claro que a descrição era um fator importante para atrair e persuadir os consumidores a fazer compras online.

De acordo com MENDONÇA (2016), os primeiros casos de sucesso de *e-commerce* como o site da Amazon passaram a investir em descrições mais detalhadas e precisas. Eles sabem que, ao fornecer informações completas sobre um produto, incluindo características, benefícios e especificações técnicas, podem aumentar a confiança do consumidor e aumentar as taxas de conversão. A Amazon implementou um sofisticado sistema de recomendação chamado “*Amazon Recommends*” desde o seu início. O sistema usa algoritmos sofisticados para analisar o histórico de compras de um cliente, visualizações de produtos e até mesmo interações com avaliações de outros compradores.

A segunda era trata-se dos anos 2000 até 2010, quando houve uma evolução significativa da estratégia de descrição dos produtos. À medida que o comércio eletrônico cresceu e amadureceu, as estratégias de descrição de produtos também evoluíram. Neste período, os profissionais de *marketing* estão começando a adotar práticas como uma descrição cativante (onde passa a enfatizar não apenas as características técnicas, mas também os benefícios emocionais que o produto proporciona ao consumidor).

A ilustração começa contando uma história e criando uma conexão emocional. Neste tempo, também a incorporação de imagens de alta qualidade

tornou-se a norma. As fotos não só mostram o produto, mas também como ele é utilizado no dia a dia.

Por último, e não menos importante, temos o surgimento das avaliações e feedback dos clientes. Com a introdução de um sistema de avaliação, os usuários podem compartilhar suas experiências sobre os produtos e influenciar nas decisões de compra.

Com os avanços da tecnologia, as estratégias de descrição de produtos no comércio eletrônico evoluíram um passo adiante. Começa então, a partir de 2010, a era da personalização e da análise. A personalização se tornou a chave para otimizar vendas e lucros. Agora alimentados por algoritmos de aprendizado de máquina, os sistemas de recomendação analisam o comportamento do usuário e as interações anteriores para fornecer descrições e produtos altamente relevantes. Segundo dados do Relatório da Comissão do Parlamento Europeu (vide referência 2015), essa personalização personalizada não apenas aumenta as taxas de conversão, mas também aumenta a receita, pois os consumidores tendem a gastar mais quando recebem recomendações precisas e envolventes.

Em resumo, a noção de que as descrições dos produtos afetam as vendas e os lucros do comércio eletrônico existe desde o advento da Internet comercial. Com o tempo, essa percepção evoluiu de descrições simples para estratégias mais detalhadas com elementos visuais e, mais recentemente, para personalização avançada baseada em dados e algoritmos. O poder das descrições dos produtos no comércio eletrônico hoje é inquestionável e a sua influência continua a moldar o sucesso dos negócios online (BOYINBODE & AFOLABI, 2018).

#### **4.1 - Conceitos básicos sobre descritores de produtos**

No comércio eletrônico, o termo “descriptor de produto” refere-se a um elemento de informação associado a um produto numa plataforma de negociação online, que desempenha um papel importante na apresentação e compreensão desse produto aos usuários. Segundo KHANAL & MISHRA (2016), esses elementos incluem informações textuais, visuais e interativas destinadas a fornecer informações importantes sobre recursos, especificações, preços e condições de compra do produto.

As descrições dos produtos se destacam como uma ferramenta versátil de comunicação entre comerciantes e consumidores no ambiente digital. Essas

informações representam a primeira impressão que o consumidor tem de um produto e podem incluir elementos como: (i) o título do produto; (ii) a descrição do produto; (iii) palavras-chaves e *tags* e, (iv) depoimentos/opiniões e classificações de clientes.

O título do produto é uma parte importante que chama a atenção do consumidor. Ele deve ser estruturado de forma a resumir em poucas palavras as informações mais relevantes sobre o seu produto para despertar o interesse do comprador. Já a descrição do produto fornece uma lista detalhada de definições do produto e descreve seus recursos, benefícios e características técnicas. A precisão e clareza desta descrição são importantes para a tomada de decisão do consumidor (GÖKMEN, 2012).

As palavras-chaves e *tags* são os elementos invisíveis, mas muito importantes, que facilitam a indexação dos seus produtos pelos motores de busca e a categorização eficaz da sua plataforma de *e-commerce*. Enquanto os depoimentos/opiniões e classificações de clientes fornece um *feedback* valioso de clientes em potencial e compradores anteriores que influenciam suas decisões de compra (KHANAL & MISHRA, 2016).

#### **4.2 - Personalização de comércio eletrônico usando algoritmos**

A personalização é um dos avanços mais importantes no comércio eletrônico atualmente e é conseguida através da utilização de algoritmos sofisticados e de linguagem natural para compreender e responder às necessidades do consumidor.

Os algoritmos de recomendação são uma das bases da personalização do comércio eletrônico. Segundo SONG (2019), esses algoritmos analisam o comportamento do usuário e o histórico de compras usando uma variedade de técnicas, incluindo filtragem colaborativa, filtragem baseada em conteúdo, sistemas de recomendação híbridos e aprendizado de máquina. Ao identificar padrões de preferência, comportamento de navegação e interações anteriores do usuário, transformamos esses dados em recomendações de produtos altamente personalizadas.

A filtragem colaborativa é uma das abordagens mais comuns para a construção de sistemas de recomendação. Analise o comportamento e as escolhas do usuário em comparação com outros usuários semelhantes (FERREIRA, 2020). A tecnologia utiliza algoritmos de aprendizado de máquina para identificar relações

entre usuários e produtos e gerar recomendações baseadas no comportamento coletivo.

Por outro lado, a filtragem baseada em conteúdo avalia as características do produto e o histórico de interação do usuário. Encontre correlações entre atributos do produto e preferências do consumidor. Essa abordagem funciona particularmente bem para recomendar produtos que compartilham certas características ou recursos com produtos nos quais os usuários já demonstraram interesse.

O aprendizado de máquina desempenha um papel importante na personalização por meio de algoritmos de recomendação. As técnicas de modelagem preditiva, como modelos de fatoração matricial, permitem que os algoritmos identifiquem padrões complexos nos dados do usuário e gerem recomendações altamente precisas. Esses modelos evoluíram consideravelmente, incorporando recursos como aprendizado profundo para melhorar ainda mais a precisão das recomendações.

#### **4.3 - Linguagem Natural como Facilitadora da Personalização**

A linguagem natural emerge como uma ferramenta essencial para melhorar a personalização em *e-commerce*. Ela possibilita compreender a linguagem dos consumidores e permite a criação de interfaces mais naturais e eficazes, bem como a oferta de suporte personalizado.

O Processamento de Linguagem Natural (PLN) é uma disciplina que permite que os sistemas de *e-commerce* compreendam, interpretem e gerem respostas em linguagem humana. Os avanços na PNL permitem que *chatbots* e assistentes virtuais entendam as dúvidas dos usuários e forneçam respostas relevantes e personalizadas. Segundo ADRIANO DE SOUSA & KLEIN, BERNARDES VOESE (2022), a análise de sentimento é outra aplicação da PNL que auxilia na personalização. Ele mede o tom e a emoção por trás das interações com os clientes, permitindo que as empresas personalizem respostas e ofertas com base no sentimento do cliente. É muito útil para gerenciar reclamações de clientes e oferecer ofertas personalizadas para aumentar a satisfação do cliente.

Segundo VELUDO-DE-OLIVEIRA & HUERTAS (2018), a personalização tem um enorme impacto na experiência do usuário de comércio eletrônico e oferece vários benefícios importantes. Dentre eles, listamos seis impactos mais significativos:

O primeiro deles é a maior relevância: A personalização torna as recomendações e o conteúdo mais relevantes para os usuários, aumentando a probabilidade de eles encontrarem produtos de seu interesse. Isso economiza tempo e melhora a satisfação do cliente.

O segundo impacto é o aumento da fidelização do cliente. Os consumidores que experimentam a personalização têm maior probabilidade de se tornarem clientes fiéis porque se sentem compreendidos e valorizados. Como resultado, as taxas de retenção e a fidelidade à marca aumentam.

Como impacto subsequente, temos o aumento das taxas de conversão: a personalização aumenta as taxas de conversão porque recomendações direcionadas incentivam os usuários a realizar ações como comprar ou assinar serviços.

Outro item dessa melhoria é o maior envolvimento: os usuários tendem a se envolver mais com plataformas que oferecem uma experiência personalizada, o que significa que passam mais tempo no site e provavelmente terão mais interações no futuro.

O quinto impacto relevante é a redução do abandono do carrinho: a personalização pode ajudar a prevenir o abandono do carrinho, oferecendo incentivos especiais e recomendações direcionadas aos compradores que indicaram intenção de compra, mas ainda não concluíram a transação.

E por último, obtém-se a maior satisfação do cliente através de experiências personalizadas quando os usuários obtêm exatamente o que desejam e precisam, levando a comentários e recomendações mais positivas.

#### **4.4 - A importância do atendimento ao cliente como determinante de vendas no e-commerce**

O comércio eletrônico é uma indústria que continua a crescer e se desenvolver à medida que a concorrência entre as lojas virtuais se intensifica. Nesse cenário, a qualidade do atendimento, principalmente por meio do serviço de atendimento ao cliente (SAC), torna-se um fator importante que pode afetar significativamente o sucesso ou o fracasso da sua loja online.

A satisfação do cliente é amplamente reconhecida como a base para o sucesso de qualquer negócio. Quando os consumidores têm uma experiência positiva com o SAC de uma loja online, eles ficam muito mais propensos a repetir

compras e recomendar a empresa para amigos e familiares. Por outro lado, um mau atendimento ao cliente pode ter consequências desastrosas, incluindo danos à reputação de uma marca e a perda de clientes insatisfeitos. Para ilustrar isso, vejamos o exemplo da Amazon, que cresceu seu ativo absurdamente quando passou a usar dados de satisfação do cliente e continuou investindo no atendimento ao cliente para manter uma base de clientes fiéis e crescentes (CONSOLI, 2014).

Um SAC eficaz desempenha um papel importante na fidelização do cliente. É mais provável que os consumidores permaneçam fiéis à sua loja online se sentirem ouvidos, compreendidos e bem tratados. Isto não só contribui para a retenção de receitas, mas também poupa recursos que poderiam ser usados para adquirir novos clientes. A empresa americana de comércio eletrônico Zappos, especializada em calçados e vestimentas, é um excelente exemplo desse conceito, que construiu uma base de clientes fiéis por meio de um excelente atendimento ao cliente (IKEDA, MARTINS e CAMPOMAR, 2013), sendo denominada por alguns especialistas como *online category killer*.

O SAC é a primeira linha de defesa para resolução de problemas e conflitos com clientes. Se for eficiente e ágil, pode resolver problemas rapidamente, evitando o desagrado público e potenciais litígios. Esse aspecto fica evidente na experiência de comércio eletrônico da Etsy, segundo PEREIRA (2010). A Etsy (empresa americana especializada em produtos feitos a mão e materiais para artesanato) é excelente em lidar com problemas de forma eficaz e minimizar qualquer impacto negativo em sua reputação.

O SAC não é apenas uma função de atendimento ao cliente. Também serve como uma rica fonte de *feedback* do cliente. A seção de “Perguntas Frequentes, Sugestões e Reclamações” fornece informações valiosas que podem ser usadas para melhorar nossos produtos, serviços e processos. Segundo LAGUNA & GUAREZI (2021), empresas como a Shopify usam ativamente o *feedback* dos clientes para melhorar suas plataformas e fornecer uma experiência de compra mais satisfatória. Sua experiência é tamanha que deixou de ser uma loja online especializada em equipamentos de *snowboarding* para desenvolver *softwares* para atender outras lojas online.

A reputação é um ativo importante no comércio eletrônico. Um bom sistema de atendimento ao cliente desempenha um papel importante na construção de uma reputação positiva e confiável. A fidelidade e a imagem da marca são fortalecidas

quando os clientes têm certeza de que sua empresa pode resolver seus problemas e atender às suas necessidades. Como se refere FELIX *et al.*, em 2023, sobre a gigante chinesa do *e-commerce*, Alibaba, ilustra esse princípio ao investir consideravelmente em seu SAC para sustentar sua imagem de confiabilidade global.

Em um mercado saturado, a excelência no atendimento ao cliente pode ser um diferencial crucial. Os consumidores estão dispostos a pagar mais e permanecer leais a uma empresa que oferece um atendimento ao cliente superior em comparação com seus concorrentes. Segundo PINTO (2023), a empresa Apple é um exemplo marcante desse fenômeno, com sua abordagem centrada no cliente sendo um pilar central de seu sucesso no comércio eletrônico.



## 5 Metodologia

A metodologia adotada é a da pesquisa-ação (LEWIN, 1947), que tem como uma de suas principais características, a ausência inicial de todos os elementos necessários para a delimitação do objetivo e gradual formulação dos meios adequados para obter resultados satisfatórios. Associando os dados coletados aos conceitos articulados na observação do modelo experimental em suas dinâmicas, busca-se articular as seis fases específicas da metodologia (CHIZZOTTI, 2018), visando a superação gradativa e sequencial daquilo que se pretende resolver.

Referenciando a escolha, definimos as fases metodológicas, que segundo CHIZZOTTI, são:

<b>A – Definição do problema:</b>	Determina-se o problema a resolver ou superar, circunscrevendo suas causas e consequências através da revisão bibliográfica.
<b>B – Coleta e análise de dados</b>	Coleta e análise de dados e informações levantadas na fase anterior, tabulação, análise quantitativa e Inferências estatísticas
<b>C – Planejamento da ação:</b>	Elencar e articular estrategicamente os elementos, objetivos, recursos, pessoas, lugares, tempos e meios, gerando um plano de execução.
<b>D - Execução da ação:</b>	As ações executadas são avaliadas com a coleta de resultados obtidos, realimentado a pesquisa e corrigindo possíveis equívocos.
<b>E - Avaliação da ação:</b>	Os resultados são novamente avaliados para a redefinição do problema com toda a complexidade observada nas ações concretas da execução.
<b>F - Continuidade da ação:</b>	Os impasses e soluções são mapeados para ampliar a compreensão do problema e propor a continuidade da pesquisa até a conclusão da tese.

Na metodologia complementar de Ciência de Dados, são mapeados os problemas que se deseja resolver ou apoiar a solução, analisados os dados e propondo um modelo de compreensão que possibilita observar padrões e até mesmo prever resultados, numa abordagem chamada *top-down* (cima para baixo). Essa metodologia deve ser eficaz em atender as demandas apresentadas sem limitações ou dependência direta de qualquer tecnologia, adequada ao volume de dados que serão processados e gerar uma resposta confiável que represente a análise quantitativa e/ou qualitativa dos dados (AMARAL, 2016).

Este projeto usará uma metodologia básica que consiste em cinco etapas bem definidas, com três delas apresentando realimentação ou avaliação de consistência de dados, objetivando que o modelo evolua constantemente através da

iteração, do refinamento e da análise contínua dos mesmos. A metodologia também deve permitir a análise contínua dos dados em qualquer período, não se limitando à sazonalidade de demandas (CASTRO e FERRARI, 2016).

As cinco etapas oferecem um acompanhamento constante do desenvolvimento do modelo e possibilitam que ajustes e correções de premissas ou objetivos possam ser realizados durante as dinâmicas de preparação e limpeza na mineração de dados brutos e possíveis transformações necessárias nos mesmos, otimizando e viabilizando as etapas posteriores. As fases da metodologia de coleta e análise de dados estão descritas:

**Definição do problema:** Nesta etapa estudamos o problema e definimos o escopo e objetivos do projeto. O objetivo da pesquisa é analisar os dados sobre alcance, engajamento, interações e outras métricas sobre a comunicação e difusão científica realizada pelas redes sociais oficiais da instituição e seus resultados na dinâmica de comunicação com as comunidades interna e externa e seu letramento científico.

**Coleta de Dados:** Os dados serão coletados em bancos de produtos de empresas de e-commerce, visando buscar diferentes versões das descrições dos produtos em diferentes situações

**Limpeza e preparação dos dados:** Os dados coletados serão limpos de inconsistências nos registros, removidos os dados ausentes e após a correção de erros, preparados para padronização de formato e transformados em dados adequados para análise.

**Análise exploratória dos dados:** A análise exploratória possibilitou compreender a estrutura dos dados e identificar padrões e tendências que geram inferências estatísticas e visuais.

**Modelagem e avaliação:** Análises descritivas sobre os dados quantitativos mais abrangentes a respeito de como os clientes das indústrias podem encontrar seus produtos e relacionar como as marcas pensam que os clientes buscam seus produtos.

**Interpretação e comunicação dos resultados:** Análises estatísticas sobre variáveis quantitativas em períodos cronológicos serão disponibilizadas através de painel eletrônico (*dashboard*) de gráficos interativos na plataforma

## 5.1 Ciclos de desenvolvimento em Ciência de Dados

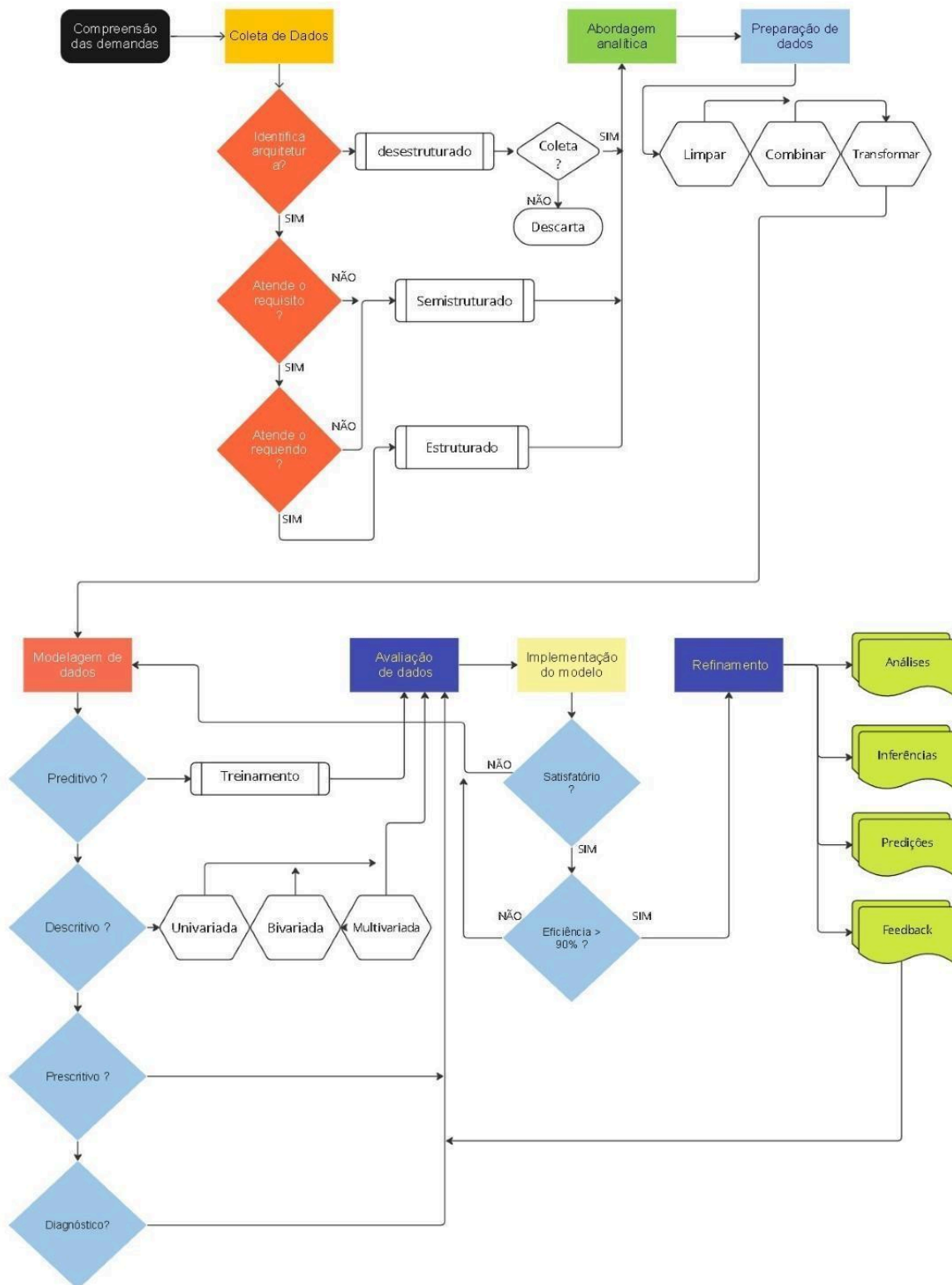


Figura 3 As fases da metodologia de coleta e análise de dados

## 5.2. Webscraping - Funcionamento do monitoramento RA

O sistema de monitoramento do RA é um robô (web crawler) que foi desenvolvido para monitorar sites de e-commerce com o objetivo de capturar dados de determinada(s) página(s) dentro do site. Este robô de monitoramento (web crawler) pode capturar qualquer informação (pública) disponível na página e que seja acessível através da internet.

Os principais dados capturados pelo robô de monitoramento dentro dos sites configurados são (mas não limitado a):

1. Título de um produto
2. Descrição de um produto
3. Unit Price (Preço padrão)
4. From Price (Em promoção)
5. Disponibilidade do produto (Can buy)
6. ProductCode (SKU do Varejo)
7. Brand (Marca do Produto)
8. Seller
9. Rating
10. Review

## 5.3. Especificação Funcional do Monitoramento

O robô de monitoramento (web crawler) funciona de acordo com o fluxo exemplificado no fluxograma abaixo.

## 5.4. Bibliotecas e suas funcionalidades:

- **Requests:**

Utilizada para viabilizar protocolos de requisição entre nosso programa de computador e sites externos, onde utilizaremos o comando “requests.get({URL do site}, {especificações do navegador})” para receber a página contida em {URL do site};

- **Beautifulsoup4:**

Usando o método BeautifulSoup({arquivo obtido pelo Requests}, “html.parser”) conseguimos aplicar a estrutura HTML ao arquivo e realizar interações com a página caso seja necessárias para obter mais informações para o request.get da biblioteca requests;

- **Boto3:**

Biblioteca utilizada para se comunicar com o S3 Bucket, que é onde armazenaremos todos os dados coletados, para posterior análise utilizando Lambda. Aqui está um exemplo de uma primeira configuração do S3 Bucket.

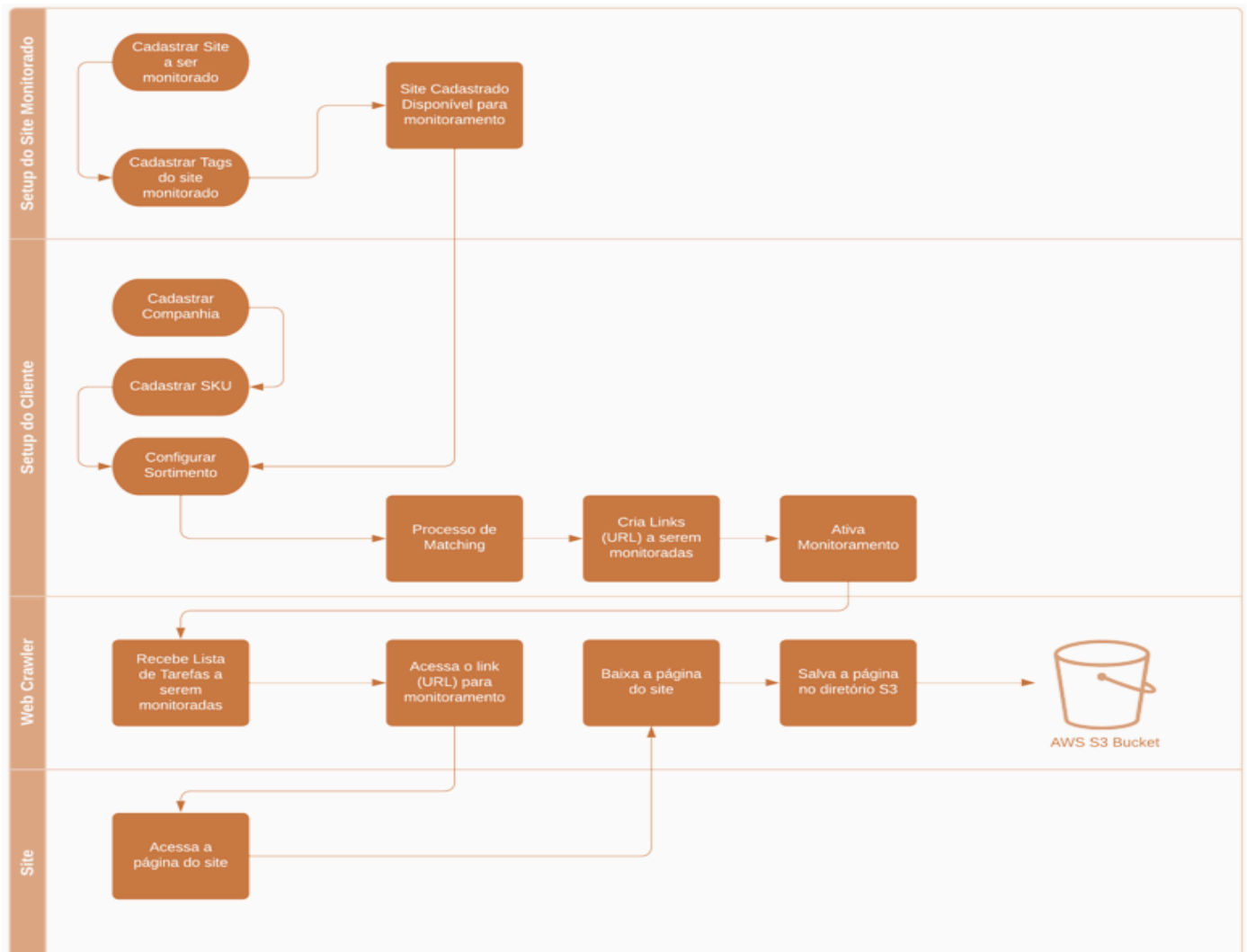



Figura 4 Configuração do S3 Bucket

Tabela descritiva das etapas referente a figura 4

<b>Etapas do Fluxo</b>	<b>Processo</b>	<b>Descrição</b>
Setup do site monitorado	Cadastrar site a ser monitorado	A atividade consiste em cadastrar o nome do site e a sua URL no banco de dados da plataforma. Exemplo: Site: Amazon URL: <a href="https://www.amazon.com">https://www.amazon.com</a>
Setup do site monitorado	Cadastrar tags do site monitorado	A atividade consiste em configurar as tags <sup>1</sup> do site monitorado.
Setup do cliente	Cadastrar uma companhia / empresa	A atividade consiste em cadastrar o nome da empresa do cliente na plataforma. Exemplo: Unilever A tabela utilizada para o cadastro do nome da empresa na plataforma é a [Company].

Setup cliente	do	Cadastrar SKU da empresa	<p>A atividade consiste em cadastrar o produto (SKU) do cliente na plataforma.</p> <p>Exemplo: Açúcar Demerara Mãe Terra 400g</p> <p>Products</p>  <p>Açúcar Demerara Mãe Terra 400g EAN: 7896496940158 - SKU: 67620251 Acreditamos que mudança pode ser feita com pequenos passos, como substituir o açúcar branco pelo nosso Açúcar Demerara O Açúcar Demerara é...</p> <p>Açúcar Mãe Terra Mascavo Orgânico 1kg EAN: 7896496941810 - SKU: 67893165 • O nosso açúcar mascavo orgânico é obtido da secagem do melado da cana, sem uso de aditivos químicos! • O açúcar mascavo é a versão sólida...</p> <p>A tabela utilizada para o cadastro do nome da empresa na plataforma é a [Product].</p>
Setup cliente	do	Processo de matching	<p>O processo de matching é uma atividade de setup / on-boarding do cliente na qual a URL do produto encontrado no site de e-commerce é cadastrado na plataforma.</p> <p>Este é a URL que o robô de monitoramento utiliza para acessar a página e localizar os dados que serão armazenados e processados na plataforma.</p>
Web Crawler		Recebe lista de tarefas para serem monitoradas	<p>A etapa de monitoramento dos dados nos sites se inicia diariamente (ou dentro da agenda programada) quando o robô de monitoramento recebe uma "tarefa" (task).</p> <p>A tarefa é um comando para o robô de monitoramento acessar, através da URL do site cadastrado, a página a ser monitorada.</p>

Web Crawler	Download da página do site <sup>2</sup>	A etapa de download consiste em armazenar a página do site monitorado no diretório da plataforma (AWS S3) para posterior processamento dos dados.
-------------	---	---

Os elementos HTML ou também chamados de tags HTML, são utilizados para informar ao navegador que tipo de estrutura é essa que está sendo construída, podendo ser títulos, parágrafos, imagens, links, preço, entre outros.

### **Pré-requisitos para o funcionamento do robô de monitoramento:**

1. Cadastro do site (URL) a ser monitorado pelo robô de monitoramento
2. Configuração das tags HTML do site de acordo com as informações que devem ser capturadas dentro da página monitorada
3. Cadastro do cliente na plataforma (Company)
4. Cadastro de SKU
6. Cadastro do sortimento (SKU x Lojas)
7. Processo de Matching - Cadastro dos links das páginas (URL) a serem monitoradas

### ***Print Screen* (Captura da Página para Evidência do Monitoramento)**

Todos os monitoramentos que forem realizados pelo robô possuem um “*print screen*” com os resultados exibidos na página que também é armazenado no diretório AWS S3 para ser utilizado como evidência na plataforma.

Exemplo de um *print screen* a partir do monitoramento realizado pelo navegador (browser):



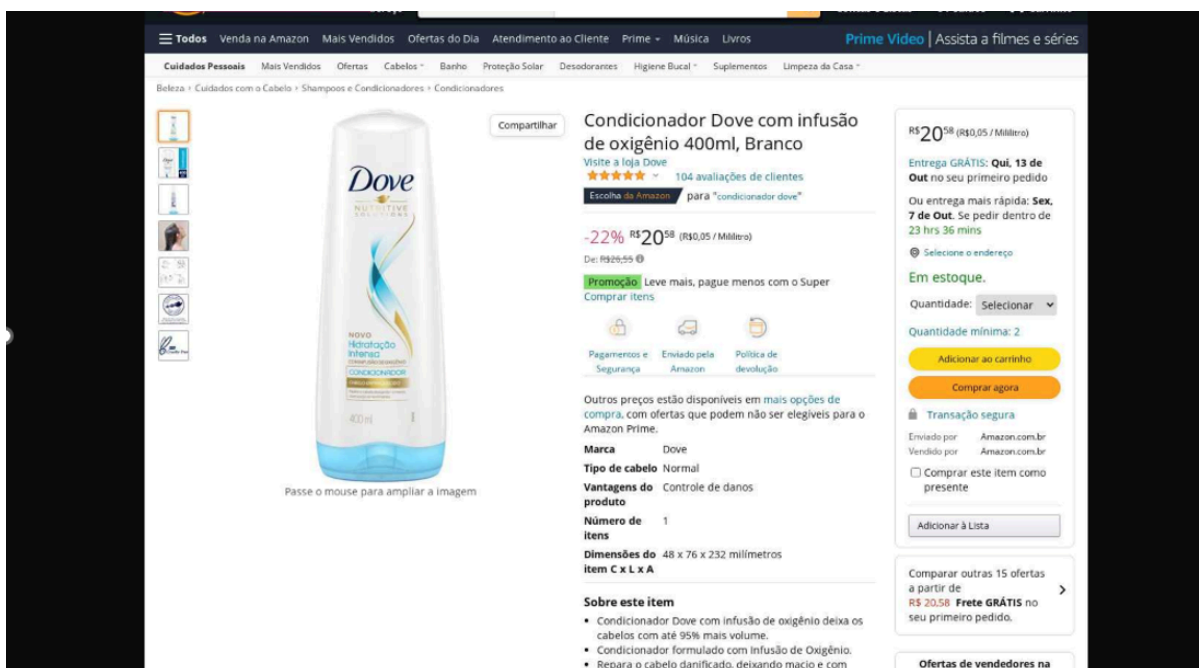


Figura 5 Imagem de acesso dos dados do site- através de uma API ou do navegador

## Armazenamento dos dados monitorados / coletados

Todos os dados que foram acessados com sucesso e coletados pelos robôs de monitoramento são organizados e transferidos para o armazenamento no diretório da aplicação (AWS S3).

Na etapa de armazenamento dos dados coletados temos o seguinte processo:

1. Se o tipo de coleta não é usando a API do site, os dados brutos são armazenados em formato de arquivo HTML.

Uma vez os arquivos armazenados no repositório da AWS S3, os dados ficam disponíveis para serem processados pelo sistema.

## Processamento e Agregação utilizando Lambda

Uma vez que todos os dados brutos estão dentro do S3 Bucket, podemos utilizar o serviço Lambda também da Amazon que nos permite analisar estes arquivos utilizando computação em nuvem.

Com o ambiente de agregação pronto, voltamos a utilizar o BeautifulSoup4 mas agora em conjunto com o Pandas para coletar e estruturar os dados em arquivos .csv.

Com os dados coletados, faremos comparações com o Setup do cliente para identificar oportunidades de melhorias ou inconsistências.

## 6 Cronograma

A execução de projeto de pesquisa segue proposto no cronograma abaixo:

ETAPAS/MÊS	AGO	SET	OUT	NOV/DEZ
Escolha do tema	A			
Levantamento bibliográfico	A D	B		
Desenvolvimento do projeto	C	D	D	
Coleta de dados	C	D E	D	
Análise dos dados		C	D E	E F
Conclusão e apresentação			D D	E

Figura 6 Cronograma de projeto de pesquisa

### Etapas da pesquisa-ação

<b>A – Definição do problema:</b>	Determina-se o problema a resolver ou superar, circunscrevendo suas causas e consequências através da revisão bibliográfica.
<b>B – Coleta e análise de dados</b>	Coleta e análise de dados e informações levantadas na fase anterior, tabulação, análise quantitativa e Inferências estatísticas
<b>C – Planejamento da ação:</b>	Elencar e articular estrategicamente os elementos, objetivos, recursos, pessoas, lugares, tempos e meios, gerando um plano de execução.
<b>D - Execução da ação:</b>	As ações executadas são avaliadas com a coleta de resultados obtidos, realimentando a pesquisa e corrigindo possíveis equívocos.
<b>E - Avaliação da ação:</b>	Os resultados são novamente avaliados para a redefinição do problema com toda a complexidade observada nas ações concretas da execução.
<b>F - Continuidade da ação:</b>	Os impasses e soluções são mapeados para ampliar a compreensão do problema e propor a continuidade da pesquisa até a conclusão da tese.

Figura 7 Cronograma - Pesquisa-Ação

## 7 Resultados e discussão

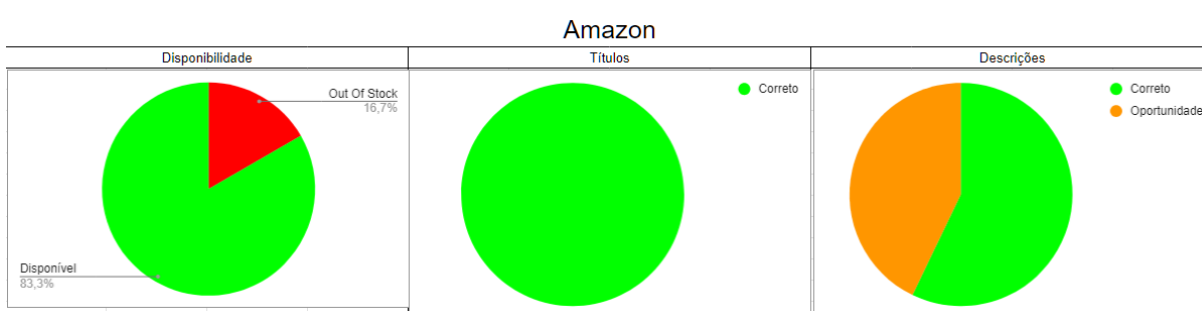
Como resultado deste trabalho de conclusão de curso (TCC) gostaríamos de mostrar para as indústrias que é possível melhorar a execução digital de seus produtos para evitar problemas de SAC, ter títulos e descrições mais atraentes e que ajudem o consumidor final a efetuar a compra e principalmente, manter seus produtos disponíveis.

Não é necessário ligar para cada varejista ou agregar várias fontes de dados desestruturados para tentar entender as oportunidades disponíveis para o e commerce de uma indústria, nosso projeto possibilita monitorar todos os varejistas que a indústria julgue relevante e levar insights em apenas um clique. Demonstrando com dados quais as oportunidades mais latentes para atuação.

Ao coletar os **títulos, descrições e a disponibilidade** de conseguimos metrificar o quanto destes KPI está de acordo com o que a indústria espera, e transformar estes dados coletados em insights a partir de gráficos atualizados diariamente para refletir a realidade de cada ecommerce e possibilitar a indústria a visualização de milhares de produtos em centenas ou dezenas de varejistas em apenas um clique.

A seguir, mostramos um exemplo da coleta de **títulos, descrições e a disponibilidade** realizada na Amazon para a Indústria Elgin no dia 07/11/2023. Dados demonstrados na planilha: [Clique aqui para verificar a planilha](#)

Legenda prodriguesdasilva5@gmail.com



Como objetivo final, desejamos desempenhar papel fundamental nas decisões estratégicas de nossos clientes, para auxiliá-los com dados a entender e atuar nas oportunidades disponíveis. Com o intuito principal de melhorar a qualidade do ecommerce e melhor atender seus consumidores.

## 9 Referências

- ADRIANO DE SOUSA, H.; KLEIN, L. BERNARDES VOESE, S. **The Impact of Institutional Pressures on the Use and Maintenance of E-Commerce in Brazilian - Micro and Small Enterprises (MSEs)**. Revista Brasileira de Gestão de Negócios, v. 24, n. 2, p. 366-382, 2022.
- AMARAL, Fernando. **Introdução à Ciência de Dados: Mineração de Dados e Big Data** - ISBN: 9788576089346, 1a. Edição, Editora Alta Books, 2016
- BOYINBODE, O. K.; AFOLABI, M. O. **Endorsed Transactions e-Commerce in Africa: The Case of Nigeria**. EAI Endorsed Transactions on Serious Games, v. 4, nº 15, p. 3 – 8, 2018.
- CONSONI, G. B. **Recuperação de informação em sistemas de recomendação: análise da interação mediada por computador e dos efeitos da filtragem colaborativa na seleção de itens no website da Amazon.com**, 2014. Disponível em: <https://www.lume.ufrgs.br/bitstream/handle/10183/97844/000920505.pdf>. Acesso em 30 ago 2023.
- DE CASTRO, L. N.; FERRARI, D. G. **Introdução à Mineração de Dados: Conceitos Básicos. Algoritmos e Aplicações**, Saraiva, 2016.
- FELIX, A. R., CRUZ, C. B., ALVES, G., XAVIER, D. R., & DE ANDRADE BAPTISTA, J. A. **Análise das operações comerciais realizadas em plataformas e-commerce**. Advances in Global Innovation & Technology, v. 1, n. 2, 2023.
- FERREIRA, M. **Conhecimento do consumidor, tendência à maximização e a busca de informação após decisões**. RAE, vol. 60, nº 1, 2020, p. 20 +. Gale Academic OneFile. Disponível em: <https://link.gale.com/apps/doc/A625710886/AONE?u=capes&sid=bookmark-AONE&xid=827bc428>. Acesso em 29 ago 2023.
- GÖKMEN, A. **Virtual business operations, e-commerce & its significance and the case of Turkey: Current situation and its potential**. Electronic Commerce Research, v. 12, nº 1, p. 31 – 51, 2012. Disponível em: <https://doi.org/10.1007/s10660-011-9084-2IBAM>. Acesso em 29 ago 2023.
- IKEDA, A. A.; MARTINS, C. A.; CAMPOMAR, M. C. **Sinalização de qualidade em e-commerce no momento pré-compra: diferenças entre websites de varejistas de produtos de luxo e de massa**. PMKT: Revista Brasileira de Pesquisas de Marketing, Opinião e Mídia, São Paulo, n. 13, p. 15-33, 2013.
- KHANAL, A. R.; MISHRA, A. K. **Financial performance of small farm business households: the role of internet**. China Agricultural Economic Review, v. 8, n. 4, p. 553-571, nov. 2016.
- LAGUNA, A., GUAREZI, E. Z. M. **O Modelo de negócios dropshipping: um estudo de caso da Shopify, plataforma de e-commerce**. 2021. Disponível em: <https://repositorio.ufsc.br/handle/123456789/223149>. Acesso em 30 ago 2023.
- MENDEZ, L. Z. R. **E-commerce: origem, desenvolvimento e perspectivas**. Monografia de Bacharelado, Universidade Federal do Rio Grande do Sul. Disponível em: <http://hdl.handle.net/10183/78391>, 2013.

MENDONÇA, H. G. de. **E-Commerce**. Revista Inovação, Projetos e Tecnologias, v. 4, n. 2, p. 240–251, 2016. Disponível em: <https://doi.org/10.5585/iptec.v4i2.68>

PARLAMENTO EUROPEU. **Transformação digital da indústria e empresas europeias: Relatório do Fórum de Políticas Estratégicas em Empreendedorismo**. 2015. Disponível em: [https://www.europarl.europa.eu/doceo/document/A-8-2017-0183\\_PT.html](https://www.europarl.europa.eu/doceo/document/A-8-2017-0183_PT.html)  
Acesso em: 01 set 2023

PEREIRA, A. de M. **Espaços e formas de consumo em uma loja virtual: um estudo da Etsy.com**. 2010. Disponível em: <http://hdl.handle.net/10183/25560>. Acesso em 30 ago 2023.

PINTO, M. **Estratégias e Tendências de social media no e-commerce**. The Trends Hub, n. 3, 2023.

SCHLEMMER, A.; PADOVANI, S. **Analytical study of stages and information that constitute online shopping**. Brazilian Journal of Information Design, v. 10, p. 224 +, 2013. Disponível em: <https://link.gale.com/apps/doc/A465696154/AONE?u=capes&sid=bookmark-AONE&xid=2bb1a7bd> Acesso em: 01 set 2023.

SONG, A. K. **The digital entrepreneurial ecosystem: a critique and reconfiguration**. Small Business Economics. Forthcoming. 2019. Disponível em: [https://econpapers.repec.org/article/kapsbusec/v\\_3a53\\_3ay\\_3a2019\\_3ai\\_3a3\\_3ad\\_3a10.1007\\_5fs11187-019-00232-y.htm](https://econpapers.repec.org/article/kapsbusec/v_3a53_3ay_3a2019_3ai_3a3_3ad_3a10.1007_5fs11187-019-00232-y.htm) Acesso em: 25 ago 2023.

STEININGER, D. M. **Linking information systems and entrepreneurship: A review and agenda for IT associated and digital entrepreneurship research**. Information Systems Journal, 29, 363- 407. 2019. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/isj.12206> Acesso em: 26 ago 2023

TURAN, A. H. **Internet Shopping Behavior of Turkish Customers: Comparison of Two Competing Models**. Journal of theoretical and applied electronic commerce research, v. 7, n. 1, p. 13 – 14, 2012. Disponível em: <https://doi.org/10.4067/S0718-18762012000100007>  
Acesso em: 28 ago 2023.

VELUDO-DE-OLIVEIRA, Tânia & HUERTAS, Melby. (2018). **Satisfação, Confiança e Familiaridade no Varejo Online e o Papel Moderador da Experiência**. Revista Brasileira de Marketing. 17. 415-428, 2018. Disponível em: <https://doi.org/10.5585/remark.v17i3.3810>. Acesso em 27 ago 2023.

CHIZZOTTI, Antonio. **Pesquisa em ciências humanas e sociais**. Cortez editora, 2018. <https://scholar.google.com.br/citations?user=Y96oF2oAAAAJ&hl=pt-BR&oi=sra>