

## API DE NOTÍCIAS FINANCEIRAS PARA EMPRESAS BRASILEIRAS

Lucas Bolela Silveira  
Graduado em Sistemas de Informação – Uni-FACEF  
lucasbolelasilveira@gmail.com

Vinícius da Costa Junqueira  
Graduado em Sistemas de Informação – Uni-FACEF  
viniciusjunqueira@gmail.com

Leandro Borges  
Docente do Uni-FACEF  
leandro.borges@facef.br

### RESUMO

O presente trabalho irá desenvolver uma API de Notícias Financeiras para Empresas Brasileiras, com o objetivo de fornecer informações financeiras relevantes e oportunas para organizações brasileiras. A metodologia envolveu o uso das linguagens de programação Python e o framework Django, juntamente com a implementação de processos ágeis baseados na metodologia Scrum. A API foi desenvolvida com a capacidade de realizar raspagem de dados em fontes de notícias financeiras, padronizando essas informações em formato JSON e armazenando-as em um banco de dados PostgreSQL. Os resultados mais significativos incluem a criação de uma API escalável que permite aos usuários acessar notícias financeiras de várias fontes, a integração com sistemas de comunicação interna, como o Slack.

**Palavras-chave:** API. Notícias Financeiras. Raspagem de Dados.

### ABSTRACT

*This work presented the development of a Financial News API for Brazilian Companies, with the aim of providing relevant and timely financial information for Brazilian organizations. The methodology involved the use of Python programming languages and the Django framework, together with the implementation of agile processes based on the Scrum methodology. The API was developed with the ability to scrape data from financial news sources, standardizing this information in JSON format and storing it in a PostgreSQL database. The most significant results include the creation of a scalable API that allows users to access financial news from multiple sources, integration with internal communication systems such as Slack.*

**Keywords:** API. Financial News. Data Scraping.

## 1. INTRODUÇÃO

No atual cenário econômico, as organizações enfrentam constantes desafios para se manterem atualizadas e competitivas. A capacidade de acessar informações financeiras relevantes e precisas é um fator determinante para o sucesso empresarial. Nesse contexto, os sistemas de informação desempenham um papel crucial ao fornecerem ferramentas e recursos que auxiliam na tomada de decisões estratégicas.

Este trabalho de conclusão de curso tem como objetivo desenvolver um sistema de informação para empresas brasileiras, especificamente uma API (Interface de Programação de Aplicação) adaptável de notícias financeiras. O sistema utilizará *web scrapers* para realizar a raspagem em sites de notícias especializados, buscando as principais notícias do dia e compilando-as por categoria, de forma a facilitar o acesso rápido e eficiente às informações mais relevantes.

A importância de um portal de notícias financeiras para as organizações é inegável. A tomada de decisões empresariais é influenciada por uma série de fatores externos, como mudanças nas políticas econômicas, flutuações nos mercados financeiros e eventos globais que afetam diretamente as empresas. Ter acesso a um compilado de notícias financeiras atualizadas e organizadas por categorias permite que os gestores e profissionais das empresas estejam sempre informados e possam reagir de forma proativa a essas mudanças.

A utilização de *web scrapers* como ferramenta de coleta de dados é uma abordagem eficiente e escalável. Essa tecnologia permite automatizar o processo de obtenção de informações, eliminando a necessidade de pesquisa manual e reduzindo o tempo gasto na busca por notícias relevantes. Além disso, o sistema proposto será um compilador do conteúdo das notícias e categorizando-as de acordo com critérios predefinidos. Deve ser decidido adequadamente as linguagens de programação e a implementação de processos eficazes são cruciais para o sucesso do projeto, incluindo as linguagens de programação escolhidas e as práticas ágeis de gestão de projeto, como a metodologia Scrum.

O resultado final será uma API adaptável, capaz de ser perfeitamente integrada com uma ampla gama de ferramentas e sistemas utilizados internamente por uma organização. Essa API se tornará uma peça fundamental no ecossistema da empresa, permitindo o acesso simplificado e a utilização eficaz das informações financeiras compiladas e categorizadas pelo sistema, capacitando gestores e profissionais a estarem sempre atentos às principais e mais recentes notícias do mercado. Os usuários poderão acessar as informações de forma rápida e eficiente, utilizando nossa base como referência para obter dados essenciais da economia, como mercado de ações, câmbio, finanças corporativas, entre outras.

Espera-se que este trabalho contribua para a área de sistemas de informação, fornecendo uma solução inovadora e eficiente para o acesso a notícias financeiras relevantes. A disponibilidade de informações precisas e atualizadas é um fator-chave para o sucesso das organizações no mercado atual. Portanto, o desenvolvimento dessa API de notícias financeiras será um diferencial para as empresas brasileiras, capacitando-as a tomar decisões estratégicas embasadas em dados confiáveis e atualizados.

## 2. REFERENCIAL TEÓRICO

Dentro do contexto de API de notícias financeiras para empresas brasileiras, exploraremos o referencial teórico que sustenta a pesquisa e desenvolvimento de uma API de notícias financeiras. Este referencial teórico é crucial para compreender os conceitos-chave subjacentes à criação dessa ferramenta tecnológica inovadora. As principais áreas que serão abordadas neste contexto são o Web Scraper, o Sistema de Notificação Interna e as Fontes de Notícias.

### 2.1. WEB SCRAPER

Uma das técnicas de extração de informações e notícias que será utilizado é o *Web Scraper*, segundo Zhao (2017) é uma técnica para extrair dados da *World Wide Web* (WWW) e salvá-los em um sistema de arquivos ou banco de dados para posterior recuperação ou análise. Entre as razões de se utilizar esse método, estão tais como previsão de tendências, análise de concorrência e comportamento do consumidor, é uma prática comum em diversas áreas, como pesquisa acadêmica, análise de dados, marketing digital, entre outras. No projeto, por meio de uma rotina, fará a captura de notícias dos principais meios de informações relacionadas à área de finanças.

Existem duas abordagens principais para o desenvolvimento de um *web scraper*: a primeira é a abordagem baseada em bibliotecas, em que se utiliza bibliotecas específicas de uma linguagem de programação, como Python unido ao BeautifulSoup, que possibilita leitura de estrutura HTML (Linguagem de Marcação de HiperTexto), navegando e separando as informações contidas no site, se tornando um poderoso artefato na extração de Dados. Já a segunda é a abordagem baseada em ferramentas, em que se utiliza uma ferramenta específica, como o ParseHub ou o Octoparse, para extrair os dados. Estes fornecem uma interface visual, simplificando a forma como os dados são coletados e padronizados, de grande ajuda para leigos em programação e estrutura HTML.

NIU et al (2022), desenvolveu um facilitador de extração de dados de sites mais populares, como *blogs* e sites de notícias (jornais) com auxílio para trazer as imagens junto ao conteúdo. A proposta dessa pesquisa consiste em utilizar *web crawlers*, um tipo de robô que opera por mecanismos de buscas (Google, Bing), transformando links em blocos visuais, então o *framework* consegue detectar a estrutura da página. Seu funcionamento segue os passos: Primeiramente coloque a URL (*Uniform Resource Locator*) do site que deseja fazer a raspagem dos dados no *scraper tool* e clique no botão para iniciar a raspagem. Em seguida a ferramenta faz uma coleta do conteúdo do site e iniciará uma montagem com HTML, dados em JSON (*JavaScript Object Notation*) e arquivos multimídia (Vídeos, áudios e imagens), e são inseridos todos agora como HTML, formando um padrão de tags. Por fim, esses dados são desestruturados e transformados no formato de saída desejado (pdf, csv, entre outros). Com o processo utilizado no artigo citado acima, é possível encontrar caminhos já testados e comprovados que serão de grande utilidade no momento do desenvolvimento de um *web scraper* para os jornais selecionados.

Por mais que esta tecnologia e este meio de extração de dados, nos quais o *web-scraping* trata, ainda necessita de normas e validações para a veracidade dos dados, isso por diversos processos que se tornaram padrões pelos motores de busca (Google, Safari). Cada país segue com suas leis regionais, contudo de acordo

com o CFAA (Lei Fraude e Abuso de Computador), algumas leis são gerais como *copyright* e direitos sobre o conteúdo, como retratado por KROTOV (2017), mostrando que é importante colocar em prática o uso da ética quando disponibilizar essas informações coletadas.

## 2.2. SISTEMA DE NOTIFICAÇÃO INTERNA

Sistemas de notificação interna são ferramentas utilizadas por empresas e organizações para enviar mensagens e alertas importantes para seus funcionários, seja por meio de e-mails, mensagens instantâneas, ou outros meios de comunicação. Idealmente, uma boa comunicação interna é um processo de mão dupla. De acordo com Men e Stacks (2014), houve uma crescente nas evidências, que apoiar a comunicação interna, tornando-a eficaz, desempenha um papel fundamental no desenvolvimento de atitudes positivas, como confiança, comprometimento organizacional, satisfação no trabalho, identificação organizacional e relações positivas entre funcionário e organização.

Os sistemas de notificação interna são importantes para manter os funcionários informados sobre atualizações, eventos, notícias e outros assuntos relacionados ao trabalho. Além disso, esses sistemas também podem ser usados para alertar os funcionários em casos de emergência, como incêndios, evacuações e outras situações de risco.

Existem diversos tipos de sistemas de notificação interna, desde sistemas simples, onde todas as atualizações sobre a empresa são repassadas inteiramente por e-mail, até plataformas mais avançadas que permitem a segmentação de mensagens e o envio em massa. Algumas plataformas também oferecem recursos adicionais, como integração com outras ferramentas de trabalho, análises de engajamento, e a possibilidade de personalização de mensagens. Uma das ferramentas que se consolidou no quesito comunicação interna foi o Slack, segundo Business of Apps (2021) já quando era testada a versão beta do aplicativo com algumas organizações, enxergaram o potencial do Slack como uma resposta às muitas formas diferentes de comunicação dos funcionários. Em vez de reuniões, e-mails e telefonemas, tudo pode ser feito por meio de um único aplicativo. Seus números são muito relevantes, visto que o Slack em 2020, teve 18 milhões de usuários ativos diariamente e 156 mil organizações utilizando o aplicativo, gerando uma receita de 902 milhões de dólares entre março de 2020 e abril de 2021, um aumento de 43% em relação ao ano anterior.

Esses sistemas também podem ser integrados com outras ferramentas utilizadas pela empresa, como o sistema de gestão de projetos e o sistema de gerenciamento de tarefas. Essa integração permite que as informações sejam compartilhadas de forma mais eficiente entre as equipes e departamentos, facilitando a comunicação e aumentando a produtividade.

No entanto, é importante destacar que o uso desses sistemas deve seguir as políticas e normas internas da empresa e respeitar a privacidade dos funcionários. Além disso, os sistemas devem ser projetados de forma a garantir a segurança das informações transmitidas e armazenadas.

Em resumo, os sistemas de notificação interna são ferramentas importantes para manter os funcionários informados sobre assuntos relacionados ao trabalho e alertá-los em casos de emergência. A escolha da plataforma mais adequada depende das necessidades e características de cada empresa, mas é importante que o sistema seja projetado de forma a respeitar as políticas e normas

internas da empresa e garantir a segurança das informações transmitidas e armazenadas. Visando todos esses fatos, têm-se que disponibilizar uma API para integração, possibilitaria empresas a desenvolver plataformas internas com dados simples e eficazes, além da facilidade do uso dos recursos em plataformas já existentes.

### 2.3. FONTE DE NOTÍCIAS

Neste estudo, foi realizada uma análise detalhada dos principais sites de notícias financeiras e de negócios no Brasil: Globo, Valor Econômico, Money Times, InfoMoney, CNN Brasil / CNN, InvestNews e Inteligência Financeira. Serão explorados diferentes aspectos, como *design*, estrutura HTML, público-alvo e características diferenciais relevantes.

O site Globo, atrai um amplo e diversificado público-alvo, abrangendo diferentes faixas etárias e interesses. A seção de notícias de economia, direciona-se a leitores interessados em acompanhar as últimas informações sobre negócios, economia e o mercado financeiro. Fornecendo notícias essenciais para empresas brasileiras, que buscam informações que irão auxiliar nas tomadas de decisões.

O Valor Econômico, um jornal especializado em notícias econômicas, financeiras e de negócios, apresenta um design sóbrio e profissional. O público-alvo do Valor Econômico é mais especializado, abrangendo profissionais do mercado financeiro, investidores, empresários e pessoas interessadas em notícias e análises aprofundadas sobre economia, finanças e negócios. Essa especialização torna o portal como uma fonte crucial para a API, pois ela atende às necessidade de um público que busca justamente essas informações. Dessa maneira temos uma garantia no fornecimento das informações com qualidade e relevantes.

O Money Times, portal de notícias financeiras e de investimentos, possui um *design* moderno e dinâmico, com uma interface atraente e de fácil utilização. O Money Times direciona-se a investidores, profissionais do mercado financeiro e pessoas interessadas em acompanhar as últimas notícias e análises sobre investimentos, economia e finanças.

O InfoMoney, também um site de notícias financeiras e de investimentos, apresenta um *design* limpo e organizado. Tendo também um público específico, é possível ter uma garantia na relevância das informações. Trazendo alguns conteúdos educativos que podem ser muito úteis para as organizações que utilizarão a API.

A versão online da CNN Brasil / CNN, um site de notícias abrangente que inclui notícias econômicas e de negócios, possui um design moderno e sofisticado, alinhado com a identidade visual da marca CNN. A CNN Brasil / CNN atrai um público amplo, interessado em notícias nacionais e internacionais, incluindo notícias econômicas e de negócios. É uma opção para aqueles que desejam uma perspectiva global sobre os acontecimentos no mundo dos negócios, bem como para profissionais do setor financeiro e investidores.

O InvestNews, um portal de notícias financeiras e de investimentos, possui um *design* limpo, moderno e elegante. O portal direciona-se a investidores, pessoas interessadas em finanças pessoais e profissionais do mercado financeiro. Seu conteúdo abrange informações sobre mercados financeiros, investimentos, economia e negócios, abordando conteúdos que são relevantes para iniciantes no mercado e até mesmo, outras de informações que são de extrema importância para investidores com uma maior experiência.

Por fim, o portal Inteligência Financeira possui um *design* moderno e limpo. Neste site é agrupado as informações de diversas fontes. Possui 4 principais categorias: Economia, Empresas, Mercado e Política. As notícias são apresentadas como listas nas categorias, para deixar a leitura do usuário mais fluida. Ao acessar a notícia é permitido compartilhar por meio das redes sociais.

A seção seguinte aborda a estratégia tecnológica por meio da qual este projeto foi realizado, que envolve o uso de Python e o *framework* Django. Essas ferramentas são fundamentais para coletar e disponibilizar informações valiosas sobre o mercado financeiro, oferecendo aos usuários notícias relacionadas ao nicho financeiro para decisões informadas e estratégicas.

### 3. METODOLOGIA DE DESENVOLVIMENTO

No contexto do presente artigo, voltado para o desenvolvimento de uma API de notícias financeiras, foi fundamental estabelecer uma sólida metodologia de desenvolvimento que englobou tanto as linguagens de programação quanto os processos utilizados. A escolha adequada de linguagens de programação e a implementação de processos eficazes são cruciais para o sucesso do projeto. Neste tópico, foi abordada a Metodologia de Desenvolvimento adotada, incluindo as linguagens de programação escolhidas e as práticas ágeis de gestão de projeto, como a metodologia Scrum.

#### 3.1. LINGUAGENS

Neste projeto, foram utilizadas duas linguagens principais: Python com o *framework* Django. Python é uma linguagem de programação amplamente utilizada no desenvolvimento web, conhecida por sua simplicidade, clareza de código e vasta gama de bibliotecas disponíveis. Será utilizada a biblioteca BeautifulSoup, que segundo a sua documentação oficial (BEAUTIFULSOUP, 2023) fornece uma interface amigável para extrair dados de documentos HTML e XML (*Extensible Markup Language*).

O Django é um *framework web* de alto nível, escrito em Python, que segundo sua documentação oficial (DJANGO PROJECT, 2023), foi criado por desenvolvedores mais experientes, com foco em escrever o código sem precisar reinventar a roda, permitindo o desenvolvimento rápido e seguro de aplicações web. Será utilizado para a construção do *web service* que disponibilizará os dados coletados pelo *web scraper* por meio de uma API.

A opção pelo Django se deu por estar atrelado ao Python e facilitar a comunicação com o *web scraper* feito com a biblioteca BeautifulSoup, além de ter artefatos facilitadores como a criação de *models*, simples integração com banco de dados por conta da geração de *migrations* e permitir fácil adaptação para outros bancos caso necessário. Além de tudo a integração com o Django Rest Framework (DRF) é possível gerar filtros e manipular as rotas da aplicação com maior facilidade e praticidade.

A combinação da raspagem de dados, que coleta as notícias financeiras relevantes, junto a API disponibilizada, proporciona ao usuário a possibilidade de

integração das notícias a ambientes internos da empresa, podendo ainda escolher quais as fontes desejadas dentre as opções disponibilizadas.

### 3.2. PROCESSOS

O desenvolvimento deste projeto seguirá o processo de desenvolvimento ágil, baseado no *framework* Scrum. Conforme Soares (2004), o Scrum é uma abordagem iterativa e incremental que enfatiza a colaboração, a comunicação e a entrega contínua de valor. São realizadas reuniões regulares, como reuniões diárias de acompanhamento (*daily scrum*), reuniões de planejamento de *sprint*, revisões e retrospectivas.

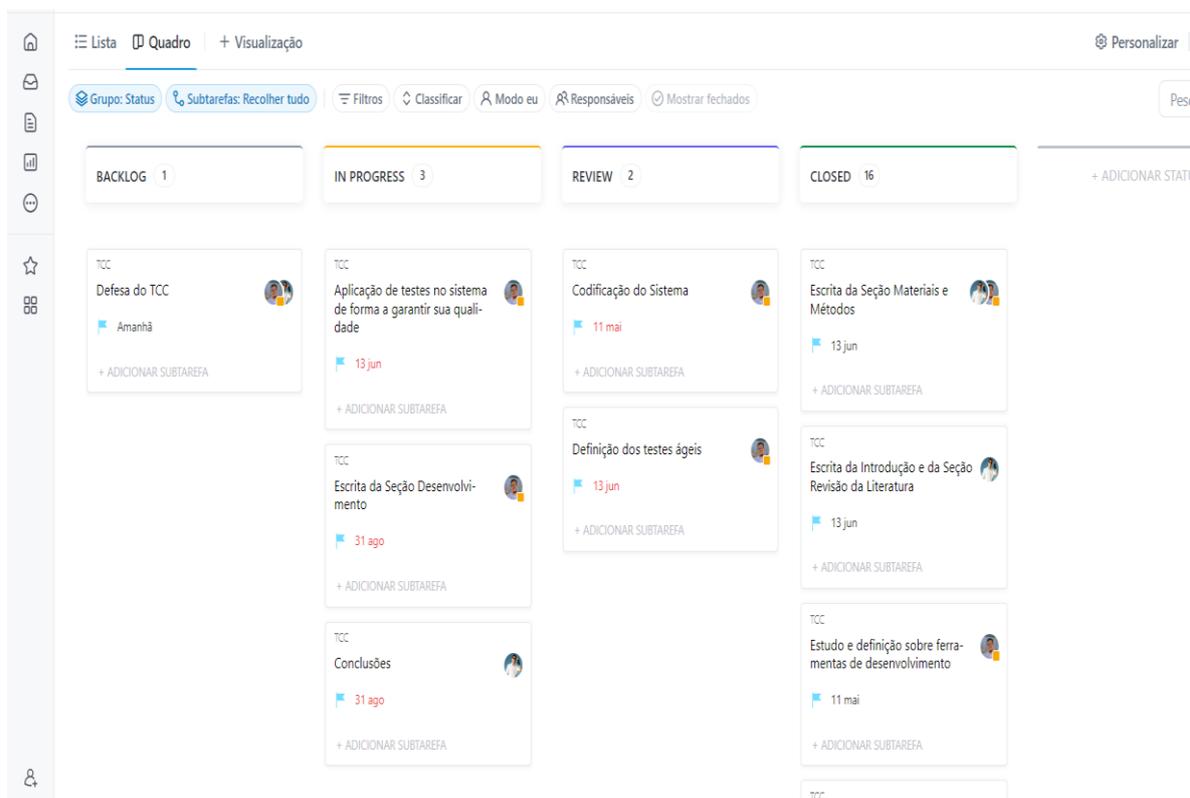
A metodologia de desenvolvimento adotada também incluiu a prática de testes automatizados. Foram desenvolvidos testes unitários e testes de integração para garantir a qualidade do *software* desenvolvido. Foram utilizadas as bibliotecas de teste do Python, como o unittest, para implementar os testes.

Além disso, foram utilizadas práticas de versionamento de código utilizando o Git e o GitHub, seguindo padrão de *commit* e de *branch*. O Git, alinhado a sua documentação oficial (GIT, 2023), é um sistema de controle de versão distribuído que permite o gerenciamento eficiente do código-fonte e a colaboração em equipe. O GitHub é uma plataforma baseada em nuvem que facilita o compartilhamento e a colaboração em projetos Git (GITHUB GUIDES, 2023).

O desenvolvimento seguirá uma abordagem iterativa e incremental, com entregas parciais ao longo do tempo. Foram definidas iterações ou *sprints*, com duração estabelecida, e ao final de cada *sprint*, foram feitas revisões do progresso e retrospectivas para identificar oportunidades de melhoria.

A metodologia adotada permitiu uma abordagem flexível e adaptável, garantindo a entrega contínua de valor ao longo do desenvolvimento do projeto.

Figura 1 - Tarefas Kanban



Fonte - Os autores

## 4. DESENVOLVIMENTO

As etapas de desenvolvimentos foram separadas em módulos de forma a seguir o seguinte fluxo: Primeiramente foi feita a extração de informações dos webjornais, a partir desses dados extraídos criaremos o banco de informação, em seguida ocorrerá a categorização por tipos de notícias, na sequência foi desenvolvida uma API de consumo desses dados já categorizados e por fim integrar com sistemas de comunicações internos (Slack).

Todas as etapas estão disponíveis, com os códigos e arquivos, no repositório GitHub do presente projeto, para acessar pelo o link: <https://github.com/LucasBolela/tcc>

### 4.1. EXTRAÇÃO DE INFORMAÇÃO

Uma etapa fundamental no desenvolvimento do sistema proposto é a extração das informações dos sites de notícias financeiras por meio do *web scraper*.

Nesta seção, será explicado como o *web scraper*, implementado em Python utilizando a biblioteca BeautifulSoup, foi utilizado para coletar os dados das principais notícias do dia e padronizá-los no formato JSON.

O processo de extração de informações consiste em percorrer os sites de notícias financeiros selecionados, acessar as páginas relevantes, identificar os elementos HTML que contêm as informações desejadas e extrair essas informações de forma estruturada. A biblioteca BeautifulSoup foi utilizada para analisar o código HTML das páginas web e facilitar a extração dos dados desejados.

Uma vez coletados os dados das notícias financeiras, foi fundamental tornar as informações padronizadas de maneira coerente e estruturada para facilitar o processamento posterior pela API. O formato JSON se apresenta como uma escolha adequada para esta finalidade, dadas suas características de simplicidade, legibilidade e flexibilidade.

A estrutura do JSON utilizada para padronizar as informações extraídas está representada na Figura 2, contendo as principais informações da notícia.

**Figura 2** - Formatação estrutural em JSON dos dados

#### Formatação pós raspagem de dados

```
JSON v Copy Caption ...
{
  "title": String,
  "subtitle": String,
  "content": String,
  "link": String,
  "tag": Object {
    "name": String,
    "link": String
  },
  "fonte": String
}
```

Tipagem da formatação pós coleta dos dados nos sites de notícias

#### Fonte - Os autores

Nesse formato, cada notícia é representada como um objeto JSON, contendo os seguintes campos:

- **title** (Título da notícia): Este campo armazena o título da notícia, fornecendo uma identificação clara do conteúdo da matéria.
- **subtitle** (Subtítulo ou resumo da notícia): Aqui, são registrados eventuais subtítulos ou resumos que auxiliam na contextualização da notícia.

- **content** (Conteúdo completo da notícia): Este campo é reservado para o conteúdo integral da notícia, permitindo o acesso às informações detalhadas.
- **link** (URL da notícia): Este campo contém a URL que leva diretamente à página web da notícia, possibilitando a referência externa.
- **tag** (Informações sobre a categoria ou tag da notícia): Este campo é representado como um objeto JSON aninhado, composto por dois campos:
  - **name** (nome da tag) e **link** (URL relacionada à tag). Esse formato permite a associação das notícias a categorias ou tags específicas, facilitando a categorização e a recuperação de conteúdo relacionado.
- **fonte** (Nome do site): Este campo armazena o nome do site que foi feita a extração dos conteúdos.

A padronização dos dados nesse formato JSON viabiliza a interoperabilidade e a consistência no armazenamento e na distribuição das notícias financeiras coletadas. Além disso, facilita a integração com sistemas de gerenciamento de banco de dados, bem como o desenvolvimento de aplicações que utilizarão essas informações de forma eficiente e eficaz.

Durante o processo de extração, o web scraper percorre as páginas dos sites de notícias financeiras, localiza os elementos HTML correspondentes aos campos desejados e utiliza os métodos da biblioteca BeautifulSoup para obter os valores relevantes. Esses valores são então armazenados em objetos JSON, seguindo a estrutura mencionada acima. Essa abordagem permite a extração eficiente das informações necessárias, garantindo a padronização dos dados coletados.

Viabilizando essa estratégia de raspagem de dados, foram identificados padrões dentro da estrutura das fontes de notícias, onde estes listam as notícias utilizando blocos semelhantes e a partir desses blocos é possível pegar o *link* da notícia por completo para extrair o conteúdo e padronizá-los. Essa estrutura pode ser notada conforme a Figura 3 mostra as notícias na fonte de informação InfoMoney, já alocadas na categoria economia, essas são apresentadas com um título, uma imagem e junto possui um *link* que direciona a notícia completa, segue a Figura 4 mostrando detalhes da notícia "Usiminas (USIM5) tem R\$ 346 milhões bloqueados pela Justiça por poluição em Ipatinga (MG)" de onde o restante das informações são coletadas.

Ao estudar os *websites* notamos que esse comportamento ocorre nos outros portais de notícias, permitindo uma estrutura padronizada com alteração apenas das classes que farão a coleta das informações em cada um dos portais.

Figura 3 - Listagem de notícias no site InfoMoney

## Economia

---



Indústria

**Usiminas (USIM5) tem R\$ 346 milhões bloqueados pela Justiça por poluição em Ipatinga (MG)**

---

## Mais sobre Economia



Câmbio

**Dólar hoje opera de lado antes de decisões do Federal Reserve e do Banco Central brasileiro**

39 minutos atrás.

---



Análise do BC-Br

**IBC-Br mostra economia brasileira resiliente, mesmo com cenário de juros altos, dizem analistas**

2 horas atrás.

---

**Fonte** - InfoMoney. Disponível em: <https://www.infomoney.com.br/economia/>.  
Acessado em: 17 set. 2023.

Figura 4 - Detalhes da notícia "Usiminas (USIM5) tem R\$ 346 milhões bloqueados pela Justiça por poluição em Ipatinga (MG)"

InfoMoney

IBC-Br Copom Fomc Investir no Exterior Aula Magna com Paulo Guedes Seguros Ime

-0,78% PETR4 R\$ 34,47 +1,00% VALE3 R\$ 68,74 +0,03% ITUB4 R\$ 27,18 -0,37% SIMULE AGORA

Indústria

## Usiminas (USIM5) tem R\$ 346 milhões bloqueados pela Justiça por poluição em Ipatinga (MG)

Siderúrgica afirmou que irá recorrer da liminar e que segue compromissos assumidos de redução das emissões de partículas sedimentáveis

Por [Estadão Conteúdo](#)  
19 set 2023 14h00-Atualizado 2 minutos atrás

[f](#) [t](#) [p](#) [w](#) [l](#) [e](#)

A Justiça de Minas Gerais bloqueou R\$ 346 milhões da Usiminas ([USIM5](#)) por conta da poluição atmosférica na unidade de Ipatinga (MG), na região do Vale do Aço, devido à emissão de pó preto. A decisão é liminar e atende a uma ação civil pública (ACV) protocolada pelo Ministério Público de Minas Gerais (MPMG) em julho deste ano.

Na ação, o MPMG requereu reparação por dano moral coletivo em razão da emissão de poluentes em desacordo com os padrões especificados pela legislação ambiental em vigor, causando poluição atmosférica.

De acordo com a decisão, que foi realizada no dia 15, mas divulgada apenas nesta terça-feira, 19, "as provas colhidas pelo MPMG demonstram que, desde a fundação da empresa, em setembro de 1966, a Usiminas pratica condutas agressivas ao meio ambiente, seja através das emissões atmosféricas geradas nas diversas áreas da empresa, da emissão de partículas sedimentáveis em desacordo com o padrão estabelecido pela legislação ambiental, conhecida

Fonte - InfoMoney. Disponível em:

<https://www.infomoney.com.br/mercados/usiminas-usim5-tem-r-346-milhoes-bloqueados-pela-justica-por-poluicao-em-ipatinga-mg/>. Acessado em: 17 set. 2023.

Ao final do processo de extração de informações utilizando o *web scraper*, os dados coletados estarão estruturados, conforme a Figura 5 exemplifica um objeto coletado com dados reais da fonte de informação InfoMoney, e prontos para serem utilizados na próxima etapa do desenvolvimento do sistema, que envolve a organização e apresentação dessas informações por meio da API de notícias financeiras para empresas brasileiras.

**Figura 5** - Exemplo de dados reais padronizados

```
JSON v Copy Caption ...
[
  {
    "title": "Usiminas (USIM5) tem R$ 346 milhões bloqueados pela Justiça por",
    "subtitle": "Siderúrgica afirmou que irá recorrer da liminar e que segue c",
    "content": "A Justiça de Minas Gerais bloqueou R$ 346 milhões da Usiminas",
    "link": "https://www.infomoney.com.br/mercados/usiminas-usim5-tem-r-346-mi",
    "tag": {
      "name": "Economia",
      "link": "https://www.infomoney.com.br/economia/"
    },
    "fonte": "infomoney"
  },
  {
    "title": "Dólar hoje opera de lado antes de decisões do Federal Reserve e",
    "subtitle": "Percepção de analistas é que dólar tende a se valorizar frente",
    "content": "Após o real ganhar força na última semana frente à moeda norte-",
    "link": "https://www.infomoney.com.br/mercados/dolar-hoje-opera-de-lado-an",
    "tag": {
      "name": "Economia",
      "link": "https://www.infomoney.com.br/economia/"
    },
    "fonte": "infomoney"
  },
  {
    "title": "IBC-Br mostra economia brasileira resiliente, mesmo com cenário",
    "subtitle": "Economistas continuam a prever perda de ritmo na atividade ec",
    "content": "A alta de 0,44% na atividade econômica em julho, captada pelo",
    "link": "https://www.infomoney.com.br/economia/ibc-br-mostra-economia-bras",
    "tag": {
      "name": "Economia",
      "link": "https://www.infomoney.com.br/economia/"
    },
    "fonte": "infomoney"
  }
]
```

Exemplo Json de dados reais InfoMoney

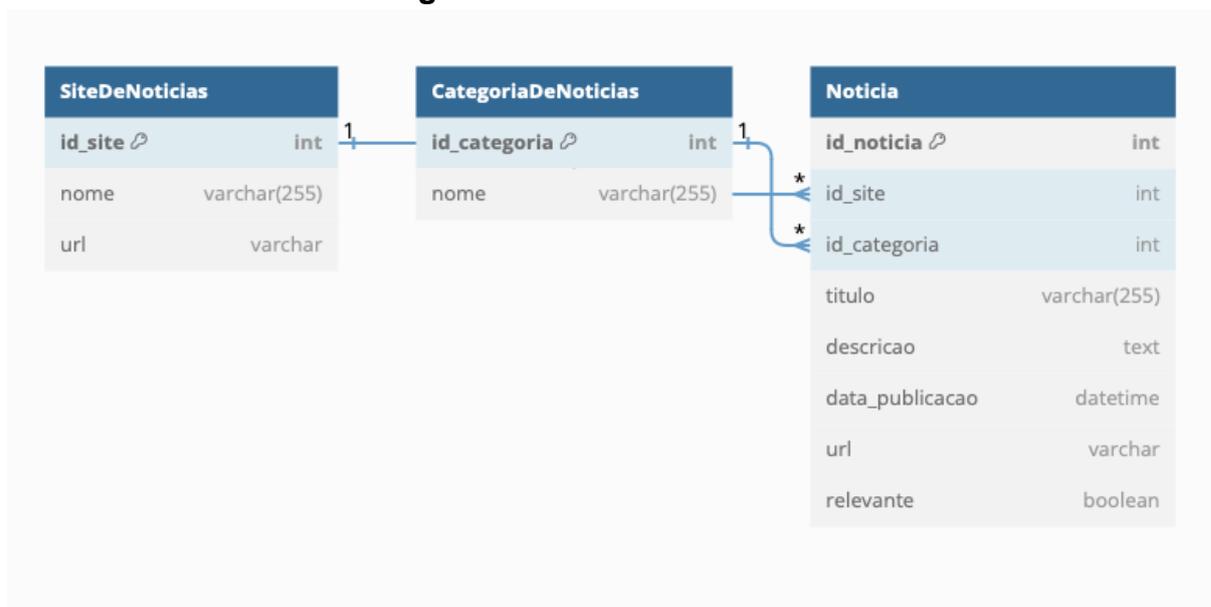
**Fonte** - Os autores

## 4.2. BANCO DE INFORMAÇÃO

No âmbito do desenvolvimento do projeto intitulado "API de Notícias Financeiras para Empresas Brasileiras", a estruturação do banco de dados desempenha um papel crucial na organização e disponibilidade de informações relevantes. Neste contexto, optou-se pela utilização do banco de dados PostgreSQL para armazenar e gerenciar essas informações com eficiência e confiabilidade.

A seção dedicada à modelagem do banco foi dividida em três elementos distintos, cada um contribuindo para a criação de um ambiente informacional coeso e eficaz. O conteúdo da modelagem com as migrações utilizando Python com Django estão disponíveis no repositório do GitHub citado no início deste tópico, e a representação do banco de dados descrita abaixo pode ser visualizada na Figura 4.

**Figura 6 - Modelo do Banco de Dados**



**Fonte - Os autores**

Inicialmente, a categorização das notícias, nomeando sua classe modelo de CategoriaDeNoticias, é responsável por armazenar as designações das categorias, permitindo uma segmentação precisa e detalhada dos conteúdos. Isso é viabilizado por meio da geração automática de um identificador numérico único para cada categoria.

A segunda parte do modelo é representada pelo componente denominado SiteDeNoticias, que oferece a capacidade de registrar o nome e a URL dos diferentes sites de notícias. Essa abordagem de separação por sites promove uma estrutura de dados que facilita a categorização e a recuperação de informações específicas de cada fonte.

Para consolidar as informações de categorias e sites, foi concebido o modelo nomeado de Noticia, que alberga os elementos essenciais das notícias, como título, descrição, data de publicação e URL. Ao empregar chaves estrangeiras, este

modelo estabelece conexões entre as informações provenientes das tabelas anteriores, possibilitando uma visão integrada e abrangente das notícias.

Em suma, a modelagem do banco de dados revela-se um elemento estratégico para a construção de um portal de notícias financeiras, oferecendo uma estrutura sólida que permite a categorização, a diferenciação por fontes e a composição coerente das informações, contribuindo assim para uma experiência informativa completa e enriquecedora.

#### 4.3. DESENVOLVIMENTO DA API

A construção de uma API robusta e intuitiva é uma peça fundamental no desenvolvimento do projeto "API de Notícias Financeiras para Empresas Brasileiras". A plataforma de desenvolvimento Python se destaca, proporcionando um ambiente propício para a implementação de soluções de alto desempenho. Nesse contexto, a escolha pelo *framework* Django Rest Framework (DRF) se justifica pela sua ampla adoção na comunidade de desenvolvimento, capacidade de construir APIs RESTful eficientes e estruturadas, e integração nativa com as funcionalidades do Django, conforme citado em sua documentação oficial (DJANGO REST FRAMEWORK DOCUMENTATION, 2023).

A abordagem de montagem da API compreende a criação de rotas de consumo para permitir a busca e a obtenção de notícias provenientes dos sites selecionados pelos usuários. A integração entre os modelos do banco de dados e o DRF proporciona um fluxo de informação dinâmico, no qual os usuários podem realizar consultas personalizadas com base em suas preferências.

Um aspecto notável da API reside na integração da raspagem de dados para carregamento de informações dos sites de notícias. Utilizando técnicas de *web scraping*, a API é capaz de acessar os sites selecionados, extrair os conteúdos das notícias e apresentá-los aos usuários de maneira organizada e coerente. Esse processo se baseia em análises cuidadosas dos elementos HTML das páginas dos sites, permitindo a extração de títulos, descrições, datas de publicação e URLs.

O emprego de *web scraping* possibilita um acesso atualizado e em tempo real às notícias financeiras, garantindo que os usuários obtenham informações pertinentes e oportunas. A interseção entre as capacidades do DRF e as técnicas de raspagem de dados cria um ambiente sinérgico que promove uma experiência enriquecedora.

A aplicação foi construída de maneira que fosse possível ter uma rota principal nomeada `web_news`, que permita passar por parâmetro, conforme a Figura 7 exemplifica, quais sites deseja buscar as informações, caso contrário busca em todos por padrão, e este retornará as últimas notícias relacionadas a finanças dos sites listados anteriormente neste projeto no formato JSON, da maneira que foi representado na Figura 5, de forma que possibilita a integração de maneira simples e concisa para que possa ser feito um front-end próprio ou, como proposto anteriormente, utilizado no sistema de comunicação interna de sua empresa.

Figura 7 - Exemplo do código fonte do web scraping

```
def get_infomoney(self, url="https://www.infomoney.com.br/economia/"):
    page = requests.get(url, headers=self.browsers)

    resposta = page.text

    soup = BeautifulSoup(resposta, "html.parser")

    all_news = soup.find_all("div", {"class": "article-card__content"})

    for news in all_news:
        current_news = news.find("a")

        news_content = requests.get(current_news.get("href"), headers=self.browsers)
        select_news_content = BeautifulSoup(news_content.text, "html.parser")

        content = ""
        for child in (
            select_news_content.find("div", {"class": "single__content"})
            .find("div", {"class": "element-border--bottom"})
            .find_all("p")
        ):
            if child.name == "p" and child.name != "div":
                content += str(child.getText())

        news_page = {
            "title": str(current_news.getText()).replace("\n", "").strip(),
            "subtitle": str(
                select_news_content.find("div", {"class": "single__excerpt"})
                .p.getText()
                .replace("\n", " ")
                .strip()
            ),
            "content": content.replace("\n", " ").strip(),
            "link": str(current_news.get("href")),
            "tag": {
                "name": "Economia",
                "link": url,
            },
            "fonte": self.site.lower(),
        }
        self.news.append(news_page)
```

**Figura 8 - Documentação da rota principal da API**

**Documentation** ×

`https://localhost:3000/api/web_news?fonte=8`

**Parâmetros de filtro:**

Código	Descrição
01	CNN / CNN Brasil
02	O globo
04	Valor Econômico
05	Money Times
06	InfoMoney
07	InvestNews
08	Inteligência Financeira

**Query Params**

fonte	8
-------	---

### Fonte - Os autores

Nesse cenário, a API se destaca como uma peça central do projeto, fornecendo um mecanismo de busca personalizada para notícias financeiras. O uso do Django Rest Framework assegura a construção de uma API escalável e bem estruturada, enquanto o emprego de *web scraping* acrescenta um elemento de atualização constante e relevância das informações. O resultado foi uma interface de usuário que permite o acesso conveniente, confiável e customizado às notícias financeiras de fontes variadas, consolidando o projeto como uma plataforma altamente valiosa para empresas brasileiras e seus interesses financeiros.

#### 4.4. INTEGRAÇÃO COM SISTEMAS DE COMUNICAÇÃO INTERNA

No contexto do projeto "API de Notícias Financeiras para Empresas Brasileiras", a capacidade de integração com sistemas de comunicação interna surge como um componente estratégico para otimizar a disseminação de informações financeiras relevantes entre os membros de uma organização. Esta seção explora a importância da integração, tanto por meio de documentação de API para sistemas proprietários quanto pela utilização de plataformas prontas para gerar *bots*, como o Slack, com o objetivo de estabelecer uma comunicação interna eficaz e contextualizada.

A integração com sistemas de comunicação interna viabiliza um fluxo contínuo de informações financeiras diretamente para os canais de comunicação preferidos das empresas. Isso é possível por meio da disponibilização de uma documentação detalhada da API, permitindo que sistemas internos personalizados consumam as notícias de maneira automática e direcionada. A flexibilidade oferecida pela API possibilita uma integração personalizada, adaptada às necessidades específicas de cada empresa.

Além disso, a integração com sistemas de *bots*, como o Slack, apresenta uma solução prática para a disseminação de informações financeiras em tempo real. Os *bots* podem ser configurados para receber atualizações automatizadas das notícias mais recentes diretamente da API. Isso promove uma comunicação proativa, onde os colaboradores podem receber notificações instantâneas sobre eventos financeiros relevantes sem interromper suas atividades cotidianas.

Em síntese, a integração com sistemas de comunicação interna se apresenta como um passo estratégico para garantir que as informações financeiras relevantes alcancem os membros da organização de maneira oportuna e eficaz. Através de documentação de API bem elaboradas e sistemas de *bots* inovadores, o projeto estabelece uma conexão direta entre as notícias financeiras e os canais de comunicação interna, promovendo a disseminação de conhecimento e aprimorando a tomada de decisões informadas no contexto empresarial.

## 5. CONCLUSÃO

No âmbito do presente projeto, foi desenvolvida uma API de Notícias Financeiras para Empresas Brasileiras, abrangendo desde a escolha das linguagens de programação até a integração com sistemas de comunicação interna, com base em uma sólida metodologia de desenvolvimento. O presente projeto buscou atender à demanda por uma solução que permitisse o acesso simplificado a informações financeiras relevantes por parte das empresas brasileiras, com foco na atualização contínua e na personalização das notícias.

A escolha das linguagens de programação Python e do *framework* Django se revelou apropriada, fornecendo um ambiente de desenvolvimento eficaz e

facilitando a integração com a biblioteca BeautifulSoup para realizar a raspagem de dados de fontes de notícias financeiras. A padronização dos dados em formato JSON permitiu a interoperabilidade e a consistência no armazenamento e na distribuição das notícias coletadas.

A estruturação do banco de dados com PostgreSQL, juntamente com a modelagem de categorias e sites de notícias, proporcionou uma organização coerente das informações. A API construída com o Django Rest Framework (DRF) ofereceu uma plataforma escalável para a busca personalizada de notícias financeiras, tornando o acesso a informações relevantes mais ágil e eficiente.

A integração com sistemas de comunicação interna, como o Slack, abriu possibilidades para que as empresas compartilhem informações financeiras diretamente em seus canais de comunicação preferidos, promovendo uma comunicação proativa e informada entre os membros da organização.

A relevância deste projeto reside na crescente importância das informações financeiras para as empresas brasileiras. Em um ambiente econômico dinâmico e globalizado, o acesso a notícias financeiras oportunas e relevantes é essencial para a tomada de decisões informadas. Além disso, a capacidade de integrar essas informações com sistemas de comunicação interna permite uma disseminação eficaz e contextualizada, contribuindo para a eficiência operacional e a competitividade das empresas.

Embora o projeto tenha alcançado seus objetivos, há espaço para melhorias e expansões futuras. Algumas áreas de melhoria incluem:

Uma das melhorias sugeridas é o aprimoramento da precisão da raspagem de dados, que envolve refinar os algoritmos de *web scraping* para extrair informações de maneira mais completa e confiável, garantindo a integridade dos dados coletados.

Outra área de desenvolvimento potencial é a integração com um maior número de fontes de notícias financeiras. Isso permitiria uma maior variedade de fontes de informações, enriquecendo ainda mais o conteúdo disponível para os usuários.

A implementação de autenticação e autorização na API é outra medida recomendada. Isso adicionaria camadas de segurança para garantir que apenas usuários autorizados tenham acesso a informações sensíveis, aumentando a proteção dos dados e a privacidade dos usuários.

O desenvolvimento de uma interface de usuário (UI) amigável é uma melhoria importante para facilitar a experiência do usuário. Uma UI intuitiva permitiria aos usuários configurar suas preferências e acessar notícias de maneira mais simples e eficaz.

Por fim, é crucial realizar testes de desempenho para avaliar a capacidade da API de lidar com um grande volume de solicitações. Se necessário, otimizações podem ser implementadas para garantir um desempenho consistente e eficiente da API em situações de alto tráfego.

Essas áreas de melhoria são essenciais para continuar aprimorando a qualidade e a eficácia da API de Notícias Financeiras para Empresas Brasileiras, tornando-a uma ferramenta cada vez mais valiosa no cenário empresarial.

Em suma, o projeto da API de Notícias Financeiras para Empresas Brasileiras representa um passo significativo no fornecimento de informações financeiras relevantes e atualizadas para as empresas. As melhorias futuras visam aprimorar ainda mais a qualidade e a abrangência das informações disponíveis, tornando-a uma ferramenta ainda mais valiosa no contexto empresarial brasileiro.

## 6. REFERÊNCIAS

- BEAUTIFULSOUP. Beautiful Soup Documentation. Disponível em: <https://www.crummy.com/software/BeautifulSoup/bs4/doc/>. Acesso em: 14 mar. 2023.
- BUSINESS OF APPS. Slack Statistics and Usage in 2021. [S.l.], 2021. Disponível em: <https://www.businessofapps.com/data/slack-statistics/>. Acesso em: 26 mar. 2023.
- DJANGO PROJECT. Django 3.2 Documentation. Disponível em: <https://docs.djangoproject.com/en/4.2/>. Acesso em: 03 jun. 2023.
- DJANGO REST FRAMEWORK DOCUMENTATION. Disponível em: <https://www.django-rest-framework.org/>. Acesso em: 03 jun. 2023.
- GIT. Git documentation. Disponível em: <https://git-scm.com/about>. Acesso em: 02 set. 2023.
- GITHUB GUIDES. Documentação do GitHub. Disponível em: <https://docs.github.com/pt>. Acesso em: 03 jun. 2023.
- KROTOV, Vlad. Legality and Ethics of Web Scraping. In: IEEE International Conference on Data Mining Workshops, 2017, pp. 958-961. Disponível em: [https://www.researchgate.net/publication/324907302\\_Legality\\_and\\_Ethics\\_of\\_Web\\_Scraping](https://www.researchgate.net/publication/324907302_Legality_and_Ethics_of_Web_Scraping). Acesso em: 13 mar. 2023.
- MEN, Linjuan; STACKS, Don. The Effects of Authentic Leadership on Strategic Internal Communication and Employee-Organization Relationships. Journal of Public Relations Research, [S.l.], v. 26, p. 1-18, 2014. DOI: <https://doi.org/10.1080/1062726X.2014.908720>. Disponível em: [https://www.researchgate.net/publication/265173308\\_The\\_Effects\\_of\\_Authentic\\_Leadership\\_on\\_Strategic\\_Internal\\_Communication\\_and\\_Employee-Organization\\_Relationships](https://www.researchgate.net/publication/265173308_The_Effects_of_Authentic_Leadership_on_Strategic_Internal_Communication_and_Employee-Organization_Relationships). Acesso em: 25 mar. 2023.
- NIU, Qingli et al. Web Scraping Tool For Newspapers And Images Data Using Jsonify. Journal of Advanced Software Engineering, [S.l.], v. 26, n. 4, p. 1-10, jun. 2022. Enviado: 28 fev. 2022. Aceito: 8 mar. 2022. Disponível em: [https://doi.org/10.6180/jase.202304\\_26\(4\).0002](https://doi.org/10.6180/jase.202304_26(4).0002). Acesso em: 20 mar. 2023.

SOARES, M.d.S.: Metodologias Ágeis Extreme Programming e Scrum para o Desenvolvimento de Software. Revista Eletrônica de Sistemas de Informação 3(1) (2004). <https://doi.org/10.21529/resi.2004.0301006>

ZHAO, Bo. Web Scraping. [S.l.]: ResearchGate, 2017. Disponível em: [https://www.researchgate.net/profile/Bo-Zhao-3/publication/317177787\\_Web\\_Scraping/links/5c293f85a6fdccfc7073192f/Web-Scraping.pdf](https://www.researchgate.net/profile/Bo-Zhao-3/publication/317177787_Web_Scraping/links/5c293f85a6fdccfc7073192f/Web-Scraping.pdf). Acesso em: 14 mar. 2023.