

# Mineração de dados em redes sociais - TWITTER

---

1ª Jornada de Ciência de Dados e Inteligência Artificial de Brasília  
02/12/2021

# Quem sou?

## @AlexandreVazRoriz

Mestre em Ciência da Computação pela Universidade de Brasília. Especialista em Análise de Dados para o Controle pelo ISC/TCU. Especialista em Gestão de Segurança da Informação pela UPIS. Graduado em Engenharia Mecatrônica pela Universidade de Brasília.

Desde 2021 é servidor do Tribunal de Contas da União onde atua em projetos relacionados a aplicações de técnicas de Ciência de Dados e Inteligência Artificial no âmbito do Controle Externo.

Desde 2017 é professor em disciplinas dos cursos "Jogos Digitais" e "Ciência de Dados e Inteligência Artificial" no Centro Universitário IESB.

<https://www.linkedin.com/in/alexandre-vaz-roriz-07880724/>

alexandre.vaz@gmail.com

# Agenda

Regras do jogo

Configuração da conta de  
desenvolvedor do Twitter

Por que minerar redes sociais?

Explorar a API do Twitter

Estudo de caso

---

# Regras do jogo

# Regras do jogo

Material disponibilizado em:

[https://github.com/DSALemDoBasico/JornaCIA\\_2021\\_DataMiningSocialNetworks](https://github.com/DSALemDoBasico/JornaCIA_2021_DataMiningSocialNetworks)

*Importante ter uma noção mínima de python, mas não é pré-requisito.*

Muito importante ter uma conta no Twitter.

Faremos um intervalo de 10 minutos às 18hrs.

Quanto mais participação melhor. :)

# Configuração da conta de desenvolvedor do Twitter

# Para que a conta de desenvolvedor??

Para ter acesso à API do Twitter.

O que seria API???

É o acrônimo de Application Programming Interface, que seria uma camada intermediária que permite que a comunicação de dois aplicativos. Quando uma aplicação deseja oferecer informações ou serviços, pode fazer isso por meio da disponibilização de uma API. Para garantir a segurança dos dados, as APIs requerem uma chave para autenticar a solicitação. A documentação da API geralmente contém as informações necessárias para instruções e requisitos de acesso.

No caso da API do Twitter, é necessária a criação de uma conta de desenvolvedor, na qual serão disponibilizadas as chaves de acesso para identificar as eventuais aplicações desenvolvidas. Essas chaves devem ser mantidas em sigilo, dado que se tratam de uma senha que poderá enviar tweets a partir do usuário.

Por que minerar  
dados de redes  
sociais?



# O que são redes sociais?

**Rede social** é uma estrutura social composta por pessoas ou organizações, conectadas por um ou vários tipos de relações, que compartilham valores e objetivos comuns. Uma das fundamentais características na definição das redes é a sua abertura, possibilitando relacionamentos horizontais e não hierárquicos entre os participantes.

*(Wikipedia)*

# Primeira rede social???



The screenshot shows the SixDegrees website interface. At the top left is the SixDegrees logo, a 3x3 grid of colored squares (red, green, blue, yellow, purple) followed by the text "sixdegrees®" and the tagline "you'd be surprised who you know". Below the logo is a button that says "check out our newest service: recommendations" with a small icon. To the right of the logo is a "member login" section with input fields for "e-mail address:" and "password:", and a "submit" button. Below the login section are two lines of text: "If you're not a member yet (it's free), [click here](#)." and "If you can't remember your password, [click here](#)." Below the recommendations button is a "sixdegrees news" section with a list of news items: "4/13 Check out this month's column on 'MegaNetworking.'", "3/18 Do you have multiple sixdegrees accounts? Merge 'em.", "3/1 Recommendations gets off to a flying start...", and "2/25 The brand new FAQ page - what you need to know about sixdegrees...". To the right of the news section is a "publicareas" section with links: "joinnow" (with a "welcome" button), "about sixdegrees" (with a magnifying glass icon), and "whitepages" (with a book icon). At the bottom of the page is a decorative border consisting of a row of small, repeating icons.

**sixdegrees®**  
you'd be surprised who you know

check out our newest service:  
**recommendations**

**sixdegrees news**

- 4/13 Check out this month's column on "MegaNetworking."
- 3/18 Do you have multiple sixdegrees accounts? Merge 'em.
- 3/1 Recommendations gets off to a flying start...
- 2/25 The brand new FAQ page - what you need to know about sixdegrees...

**member login**

e-mail address:

password:

submit

If you're not a member yet (it's free), [click here](#).

If you can't remember your password, [click here](#).

**publicareas**

- ▶ **joinnow** **welcome**
- ▶ **about sixdegrees** 🔍
- ▶ **whitepages** 📖

## Valor explorado nas redes???

- Marketing, recomendações
- Descoberta de sentimentos da sociedade
- Descoberta de relações
- *Quais mais??*

Daí a importância...

**Muito valor a ser obtido.**

Pode ser feita de forma por scrap (coletando dados de forma automatizada com a ajuda de “robôs” que simulam um usuário):

- Por meio de API disponibilizadas pelas redes;
- Incentivadas pelas redes, pois permitem que a comunidade crie ferramentas que incrementam o uso delas;
- Entretanto, não entregam o ouro... Ou seja, estabelecem limitações seja de quantidade de informação que pode ser obtida, seja pela frequência que a consulta pode ser feita.

# Por que minerar dados do Twitter??

Rede social extremamente popular. Utilização massiva. Dados frequentemente públicos.

## ***Tweet***

- Texto (micro blog) publicado na rede, originalmente, de até 140 caracteres. A partir de 2017, o limite passa a 280.

Alguns números do Twitter:

- 500 milhões de tweets mandados por dia
- Cerca de 463 milhões de usuários ativos mensalmente

# Por que minerar dados do Twitter??

## Jornalistas

Identificar a repercussão de algum assunto/pauta, talvez aplicando uma análise de sentimento.

## Profissionais de marketing

Verificar reação do público frente a novos produtos. Quais termos estão sendo conectados a ele?? Identificar fundamentos para uma estratégia...

## Analistas financeiros

Capturar tweets que possam influenciar cenários econômicos, compondo modelos de precificação.