

# Web Crawlers e OPENAI



Gileno Alves Santa Cruz Filho

# Gileno, quem?



- Ciência da Computação na UFPE
- Cientista de Dados na Stefanini



**O que são Web  
Crawlers?**

# **Crawler X Scrapping**

# Scraping

- Extrair Dados Estruturados de algo não estruturado
- HTML
- XML
- IMAGENS (OCR)
- PDF

# Crawler

- Visitar Páginas Web
- Seguir regras de visitaço
- Não necessariamente precisa extrair informação estruturada (fazer scraping)

---

# Porque criá-los?



**Informação == dinheiro**



Google



**OpenAI**

- **ChatGPT: vocês já conhecem**
- **Whisper: Transcrição de vídeos, audio...**
- **Entre outros**

**Talk is cheap,  
Show me the  
code**

**Obrigado!**  
**Dúvidas?**



contato@gilenofilho.com.br  
<https://github.com/gileno/gorn2023>