



LLMS APLICADOS À CLIMATOLOGIA E AGROMETEOROLOGIA

Foco na Amazônia e no Cerrado

Estudante: Ana Clara Guerra Torres

Orientador: Daniel Louzada

Coorientador: Júlio Reis

Outubro/2025

Sumário

1

Contextualização
do Tema

2

Motivações

3

Objetivos

4

Revisão
Bibliográfica

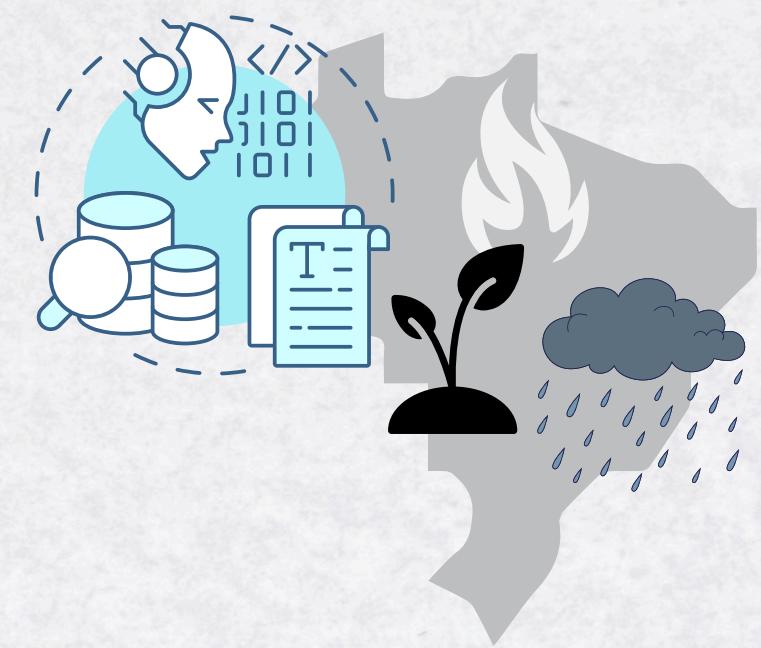
5

Metodologia
Proposta

6

Cronograma

Contextualizando o tema



- Problema Central
 - Desmatamento altera o ciclo da estação chuvosa
 - Impacto direto na produção agrícola
- Uso de LLMs (Large Language Models):
 - Modelos de IA treinados em grandes volumes de texto
 - Capazes de entender e gerar linguagem natural com contexto
 - Podem ser adaptados a **domínios específicos** (ex.: agricultura, clima)
- Conexão
 - Aplicar LLMs para analisar dados climáticos e ambientais
 - Apoiar decisões em climatologia e agrometeorologia

Motivações



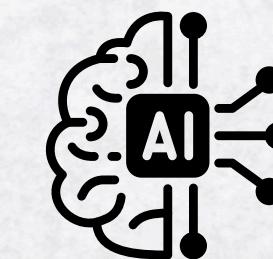
Relevância Climática

Amazônia e Cerrado influenciam o regime de chuvas em toda a América do Sul



Relevância Econômica

Agricultura é a base da economia brasileira. Então previsões climáticas confiáveis são essenciais para reduzir perdas



Relevância Científica

Há poucos modelos de IA adaptados ao domínio climático brasileiro

Objetivos

- **Objetivo Geral:**
 - Investigar a aplicação de LLMs na climatologia e agrometeorologia, com foco na análise das mudanças climáticas e seus impactos na estação chuvosa na Amazônia e no Cerrado
- **Objetivos Específicos:**
 - Mapear o estado da arte do uso de LLMs nas geociências
 - Analisar metodologias de adaptação de LLMs a domínios especializados
 - Identificar lacunas em trabalhos sobre LLMs e mudanças climáticas
 - Propor um framework conceitual para análise da estação chuvosa

Revisão Bibliográfica

A Survey of Large Language Models

- O artigo faz uma revisão ampla sobre LLMs, abordando evolução, técnicas de treinamento, adaptação (instruction tuning, RLHF, RAG) e principais desafios
- Contribuições
 - Estrutura conceitual dos LLMs (pré-treino, ajuste e uso)
 - Técnicas de especialização e alinhamento ético
- Relevância para meu trabalho
 - Base teórica para compreender o funcionamento e a adaptação de LLMs
 - Fundamenta o uso de RAG e fine-tuning no modelo especializado em climatologia
 - Justifica criar um LLM específico para domínio climático
- Limitações
 - Abordagem genérica – não trata de aplicações ambientais

Revisão Bibliográfica

- **Modelos de Uso Geral**

- Vantagem: Grande poder de generalização
- Limitações: Não especializados em temas técnicos, com risco de imprecisões ("alucinações")

- **LLMs e Geoprocessamento**

- **GeoGPT:** Integra LLMs a sistemas GIS para automatizar fluxos de trabalho espaciais
- **GeoLLM:** Utiliza LLMs para extrair conhecimento de dados geoespaciais e socioeconômicos
- Limitação Comum: Focam em geoprocessamento genérico ou aplicações urbanas, não em análises climáticas e de estação chuvosa.

Revisão Bibliográfica

Modelos de Domínio Específico

- **AgriGPT**

- Ponto-chave: Ecossistema especializado em agricultura
- Estratégia: Construído a partir de um corpus de artigos científicos

- **ClimateChat**

- Ponto-chave: O modelo de maior relevância para este projeto
- Estratégia: Adaptação de LLMs para mudanças climáticas usando um Corpus Especializado e RAG
- Vantagens: Aumenta a factualidade, superando modelos generalistas

Revisão Bibliográfica

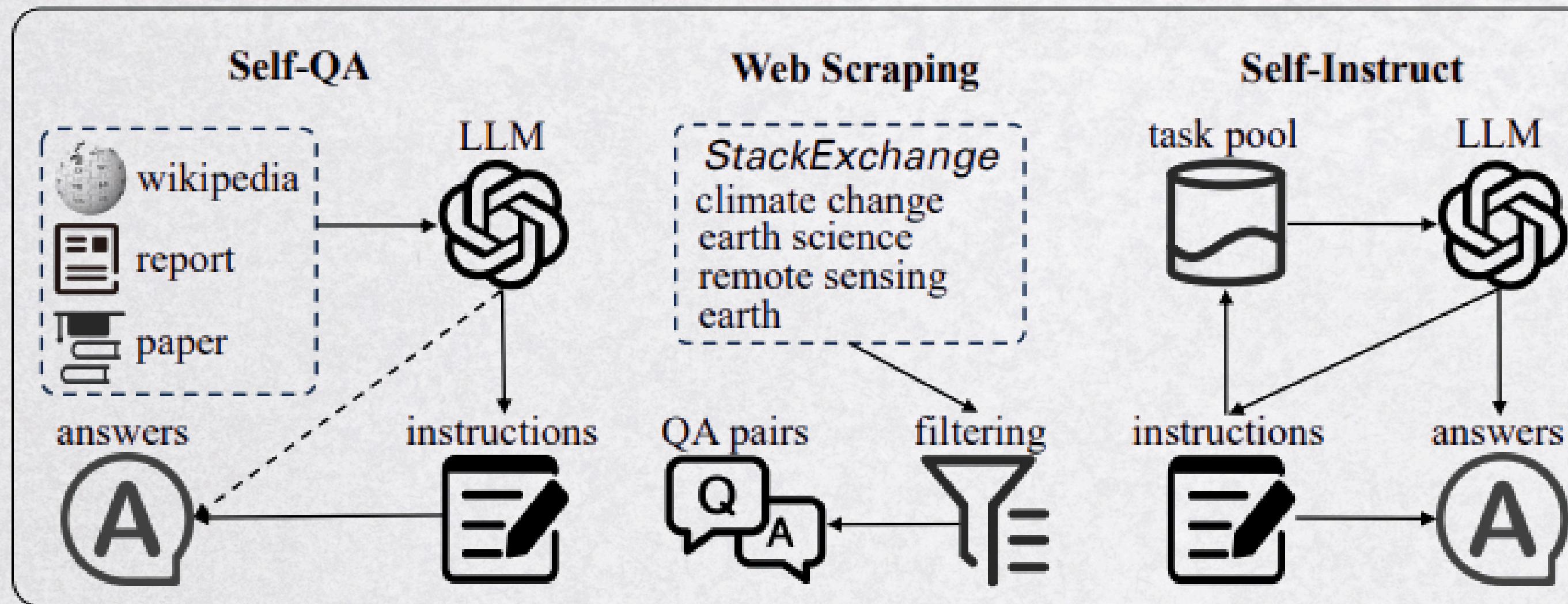


Figure 1: Construction of the ClimateChat-Corpus.

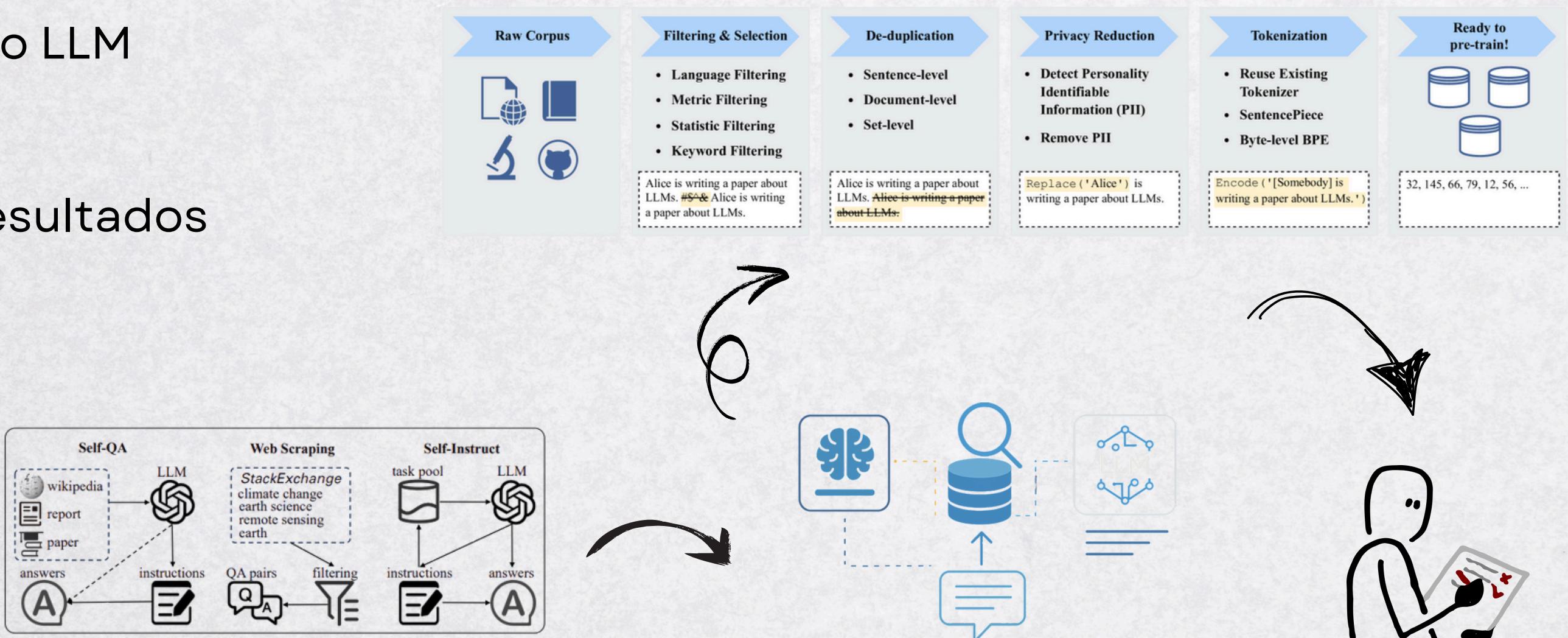
Revisão Bibliográfica

Trabalho	Foco Principal	Método / Dataset	Vantagens	Limitações
Survey of LLMs (ZHAO et al., 2025)	Revisão ampla sobre LLMs	Revisão sistemática; análise de arquiteturas, tuning e avaliação	Síntese abrangente das técnicas e desafios; base conceitual sólida	Não aplica em climatologia;
GPT-3 (BROWN et al., 2020)	Modelo generalista e few-shot	GPT-3 (175B) pré-treinado em grande corpus da web	Capacidade de few/zero-shot; alta generalização	Não especializado; possíveis alucinações e vieses
AgriGPT (YANG et al., 2025)	Agricultura e agrometeorologia	Corpus Agri-342K; Tri-RAG (recuperação densa + esparsa + grafo)	Corpus curado; benchmarks específicos	Foco agrícola amplo; não analisa estação chuvosa regional
ClimateChat (CHEN et al., 2025)	Mudanças climáticas	ClimateChat-Corpus (Self-QA, Web Scraping, Self-Instruct) + RAG	Especialização temática; melhor factualidade via RAG	Não tem recorte Amazônia/Cerrado
GeoGPT (ZHANG et al., 2024)	Integração LLM + GIS	LLM orquestrador + ferramentas GIS (buffer, intersect, clip...)	Automatiza workflows GIS via NL; bom para operações espaciais	Depende de ferramentas externas; não foca análise climática temporal
GeoLLM (MANVI et al., 2024)	Extração de conhecimento geoespacial	LLM + dados auxiliares (ex.: OpenStreetMap)	Melhora previsões espaciais; extrai padrões latentes de LLMs	Aplicações voltadas a ambientes urbanos/socioeconômicos

Quadro 1: Comparação dos principais trabalhos relacionados.

Metodologia Proposta

1. Escolha do LLM Base
2. Construção do Corpus Especializado
3. Integração com o LLM
4. Pipeline Técnico
5. Avaliação dos Resultados



Cronograma

Tabela 2 – Cronograma previsto (setembro/2025 – maio/2026).

Mês/Ano	Revisão Bibliográ- fica	Estudo de Técnicas	Proposição do Fra- mework	Texto Parcial	Defesa
Set/2025	X				
Out/2025	X				
Nov/2025		X			
Dez/2025		X			
Jan/2026		X			
Fev/2026			X		
Mar/2026			X		
Abr/2026			X	X	
Mai/2026					X

Referencias

- BROWN, T. B. et al. Language models are few-shot learners. *Advances in Neural Information Processing Systems*, v. 33, p. 1877–1901, 2020. Disponível em: <https://arxiv.org/abs/2005.14165>. Acesso em: 26 set. 2025.
- CHEN, L. et al. Climatechat: Domain-adapted large language models for climate change research. *Environmental Modelling & Software*, 2025. Preprint arXiv:2406.09876. Disponível em: <https://arxiv.org/abs/2506.13796v1>. Acesso em: 21 set. 2025.
- MANVI, R. et al. Geollm: Harnessing large language models for geospatial knowledge extraction. *Transactions in GIS*, 2024. Preprint arXiv:2401.04567. Disponível em: <https://arxiv.org/abs/2310.06213>. Acesso em: 26 set. 2025.
- VASWANI, A. et al. Attention is all you need. *Advances in Neural Information Processing Systems*, v. 30, 2017. Disponível em: <https://arxiv.org/abs/1706.03762>. Acesso em: 01 out. 2025.
- YANG, X. et al. Agrigpt: A domain-specific large language model for agriculture. *Computers and Electronics in Agriculture*, 2025. Preprint arXiv:2404.01234. Disponível em: <https://arxiv.org/abs/2508.08632v1>. Acesso em: 01 out. 2025.
- ZHANG, H. et al. Geogpt: Integrating large language models with geographic information systems. *International Journal of Geographical Information Science*, 2024. Preprint arXiv:2312.06789. Disponível em: <https://arxiv.org/abs/2307.07930>. Acesso em: 20 set. 2025.
- ZHAO, W. X. et al. A survey of large language models. *Artificial Intelligence Review*, 2025. Preprint arXiv:2303.18223. Disponível em: <https://arxiv.org/abs/2303.18223>. Acesso em: 28 set. 2025.

Obrigada

 ana.torres1@ufv.br