# Projeto: Otimização de prompt de IA visando diminuir o Consumo de Água e Energia, conscientizando o usuário

## 1. Introdução

Com o crescente uso das “inteligências artificiais”, ou os Large Language Models – LLM, verificadas no dia a dia da população, seja para consultas simples feitas via celular ou para estudos acadêmicos e geração de imagens e gráficos científicos, foi possível verificar a popularização desse tipo de uso.

Tendo em vista que o sucesso na resposta dos LLMs, seja em veracidade ou em exatidão da resposta, está diretamente relacionada com a formulação das sentenças no prompt de entrada. Faz-se necessária dominar a formulação dos prompts de forma a serem eficazes.

A engenharia de prompt que remonta dos anos de 2010 com os estudos do processamento da linguagem natural (Natural Language Processing – NLP), teve entre os anos de 2020 e 2021 destaque renovado pela popularização das LLMs através do GPT3.

Estima-se que são feitas atualmente mais de 1 bilhão de interações com o chat GPT por dia e que isso representa um consumo de 10 a 25 milhões de litros d´água diariamente, segundo o Exploding Topics 2025 visto na página <https://www.ufsm.br/2025/09/04/como-o-uso-de-inteligencias-artificiais-consome-agua>, em 30/10/2025.

É nesse contexto que o projeto está inserido apresentando como objetivo conscientizar os usuários sobre o consumo de recursos naturais, como água e energia, durante interações com modelos de linguagem (LLMs).

A aplicação propõe uma análise prévia do prompt escrito pelo usuário visando três pilares fundamentais da engenharia de prompts: clareza, contexto e direção. Como resultado da análise do prompt serão identificadas partes desnecessárias no texto e estimado o impacto ambiental causados por elas.

## 2. Objetivo Central

Criar uma aplicação que permita que o usuário insira um prompt e, com base em uma base de dados e regras pré-definidas, a aplicação identifique saudações, frases redundantes e trechos supérfluos para uma pesquisa em LLMs.

Essas partes são associadas a um consumo estimado de água e energia, e o sistema mostra o impacto total gerado.

## 3. Funcionamento do Sistema

O sistema é dividido em três etapas principais:

1. Entrada: o usuário insere o prompt.

2. Processamento: o sistema identifica partes desnecessárias e calcula o consumo.

3. Saída: exibe as partes supérfluas, o consumo de recursos e uma versão otimizada do prompt.

## 4. Arquitetura do Sistema

A aplicação é composta pelos seguintes componentes:

|  |  |  |
| --- | --- | --- |
| Componente | Tecnologia sugerida | Função |
| Front-end | Angular | Interface para digitar e visualizar resultados |
| Back-end | Python com FastAPI | API REST para análise e cálculo |
| Banco de dados | SQLite ou PostgreSQL | Armazena frases e valores médios de consumo |
| Visualização | Chart.js / Recharts | Exibe consumo em gráficos (Talvez) |

## 5. Estrutura do Banco de Dados

Exemplo da tabela de frases desnecessárias:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| id | tipo | texto | consumo\_agua\_ml | consumo\_energia\_wh |
| 1 | saudacao | Olá | 0.5 | 0.02 |
| 2 | saudacao | Tudo bem? | 0.8 | 0.03 |
| 3 | redundancia | Por favor, me explique detalhadamente | 3.0 | 0.12 |
| 4 | encerramento | Obrigado pela ajuda | 1.2 | 0.05 |

## 6. Cálculo de Consumo

Os valores de consumo são baseados em estudos sobre o impacto ambiental de modelos de linguagem. Por exemplo, uma única consulta pode consumir aproximadamente 500 ml de água e 0.5 Wh de energia. O sistema distribui esse consumo proporcionalmente ao tamanho e redundância do prompt.

## 7. Exemplo de Funcionamento

Prompt: "Olá, tudo bem? Me explique detalhadamente o que é um átomo, por favor."  
Partes desnecessárias: 'Olá, tudo bem?', 'detalhadamente', 'por favor'.  
Consumo desnecessário estimado: 1,2 litros de água e 0,7 Wh de energia.  
Versão otimizada: 'Explique o que é um átomo.'

## 8. Interface Proposta

A interface deve conter um campo para o usuário digitar o prompt, um botão de análise e uma área de resultado com o consumo e as partes desnecessárias destacadas. Também pode incluir gráficos mostrando economia total.