

Proposta de Projeto 2025-2026

LICENCIATURA EM ENGENHARIA INFORMÁTICA

PROPOSTA N.º

TÍTULO*	Orquestração de Chamadas a APIs de Large Language Models para Processamento de Dados de Saúde Pública no âmbito do projeto HfPT
ORIENTADOR PRINCIPAL*	Dennis Lourenço Paulino(INESC TEC);
COORIENTADORES	Gonçalo Campos Gonçalves (INESC TEC); Raul Manuel Viana (INESC TEC); Hugo Alexandre Paredes Guedes da Silva (INESC TEC / UTAD)
ALUNOS(S)	1
ÁREA DE INVESTIGAÇÃO	Integração de Sistemas
CENTRO DE INVESTIGAÇÃO	INESC TEC - HUMANISE
DEPENDÊNCIAS	Clique aqui para introduzir texto.
APRESENTAÇÃO*	<p>Os Large Language Models (LLMs), como o BERT e o GPT, têm vindo a assumir um papel crescente no processamento de dados de saúde, em particular na extração e estruturação de informação a partir de fontes heterogéneas. No contexto da saúde pública, muitos indicadores continuam a ser disponibilizados através de fontes instáveis e pouco estruturadas, como documentos PDF, páginas HTML ou outros formatos semi-estruturados, o que dificulta a sua integração em processos analíticos consistentes. O projeto HfPT (https://www.healthfromportugal.pt/pt/) enquadra um conjunto alargado de linhas de investigação nesta área, sendo este trabalho um subprojeto focado na aplicação de LLMs a processos de ETL (Extract, Transform and Load) de dados de saúde pública. O trabalho de licenciatura proposto incidirá no desenvolvimento de funcionalidades na API deste subprojeto, nomeadamente mecanismos de <i>throttling</i> de pedidos HTTP e de <i>chunking</i> de ficheiros de grande dimensão, de forma a melhorar a robustez, escalabilidade e fiabilidade do processamento de dados provenientes de múltiplas fontes.</p>
OBJETIVOS*	<ul style="list-style-type: none">- Análise do fluxo de interação com APIs de LLMs: identificar onde ocorrem múltiplas chamadas, limites de taxa (rate limits) e constrangimentos no envio de ficheiros de grande dimensão.- Desenho de estratégias de chunking orientadas a LLMs: definir como documentos grandes são divididos, enviados e recombinaados, tendo em conta limites de contexto e custos.- Implementação de mecanismos de controlo de pedidos a LLMs: gerir throttling, filas e temporização de chamadas para respeitar limites das APIs externas.- Gestão de erros e repetição segura: lidar com falhas parciais, timeouts e respostas inválidas das APIs de LLMs durante o processamento.- Avaliação e demonstração: testar o comportamento do sistema com documentos grandes e múltiplos pedidos concorrentes, demonstrando ganhos de fiabilidade e controlo.

* Campos de preenchimento obrigatório

NOTA: a totalidade deste documento (exceto esta linha) não deve exceder uma página.