

INOVAÇÃO

GPTs Pós*BR: a Inteligência Artificial na Pesquisa em Teses e Dissertações

Por João Alberto de Oliveira Lima
joaoli13@gmail.com

Introdução

Desde o dia 6 de novembro de 2023, a OpenAI abriu novas possibilidades com a criação de GPTs (Transformadores Generativos Pré-treinados) especializados. Esta inovação ofereceu uma oportunidade para explorar e utilizar de maneira eficiente a vasta base de dados abertos da CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) de teses e dissertações.

Com um acervo de mais de 1,4 milhão de produções acadêmicas, a base de dados da Capes foi a fonte primária para a criação de 100 GPTs, conforme a seguir: 91 GPTs especializados em cada área do conhecimento; 8 GPTs que abrangem as grandes áreas do conhecimento e 1 GPT que considera todas as áreas (ver: <https://joaoli13.github.io/GPTs.html>). Paralelamente, um trabalho similar foi realizado para a base de teses do Reino Unido (EThOS), que possui mais de 627 mil teses distribuídas no período entre 1768 e 2023, em 20 áreas do

conhecimento. Este projeto resultou na criação de GPTs adicionais disponíveis no endereço: <https://joaoli13.github.io/GPTsUK.htm>. Este texto de divulgação, contudo, focará nos GPTs criados a partir da base de teses e dissertações da CAPES.



João Lima é Analista de Informática Legislativa no Senado Federal e criador do GPTs Pós*BR.

INOVAÇÃO

GPTs Pós*BR - Produções Acadêmicas da Base de Teses e Dissertações da Capes

Aqui você encontra GPTs que dão acesso a 1.401.067 produções (dissertações e teses) acadêmicas do Brasil no período de 1987 a 2021

Estes GPTs utilizam dados provenientes do 'Catálogo de Teses e Dissertações - Brasil', um recurso fornecido pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), disponível sob a licença Creative Commons Attribution (CC BY).

Toda as Áreas do Conhecimento		BASE DE TESES E DISSERTAÇÕES									
		PósBTDOTBR (Pesquisa Convencional: mais lenta e dados completos)					PósResumoTextoBR (Pesquisa Textual: mais rápida e dados mais relevantes)				
Grande Área do Conhecimento	Ciências Exatas e da Terra	Ciências Biológicas	Engenharias	Ciências da Saúde	Ciências Agrárias	Ciências Sociais Aplicadas	Ciências Humanas	Linguística, Letras e Artes	Multidisciplinar		
Área do Conhecimento			Engenharia Civil								
			Engenharia de Minas				Direito				
							Administração				
			Biologia Geral	Engenharia de Materiais e Metalúrgica			Economia	Filosofia			
			Genética			Agronomia	Economias Agrárias e dos Recursos Naturais	Sociologia			
		Matemática	Botânica	Engenharia Elétrica	Medicina	Recursos Florestais e Engenharia Florestal	Arquitetura e Urbanismo	Antropologia			
		Probabilidade e Estatística	Zoologia	Engenharia Mecânica	Farmacologia			Arqueologia			
		Ciência da Computação	Ecologia	Engenharia Química	Enfermagem	Engenharia Agrícola	Planejamento Urbano e Regional	História	Linguística	Interdisciplinar	
		Astronomia	Morfologia	Engenharia Sanitária	Nutrição	Zootecnia		Geografia	Letras	Ensino	
		Física	Fisiologia	Engenharia de Produção	Medicina Veterinária	Medicina Veterinária	Demografia	Psicologia	Literatura Comparada	Materiais	
		Química	Bioquímica	Engenharia de Produção	Saúde Coletiva	Fonoaudiologia	Ciência da Informação	Tratamento e Prevenção Psicológica	Artes	Biotecnologia	
		Geociências	Biofísica	Engenharia Nuclear	Fisioterapia e Terapia Ocupacional	Recursos Pesqueiros e Engenharia de Pesca	Museologia	Educação		Ciências Ambientais	
		Oceanografia	Farmacologia	Engenharia de Transportes	Educação Física	Ciência e Tecnologia de Alimentos	Comunicação	Ciência Política			
			Imunologia	Engenharia de Transportes			Serviço Social	Teologia			
			Microbiologia	Engenharia Naval e Oceânica			Economia Doméstica				
			Parasitologia	Engenharia Naval e Oceânica			Desenho Industrial				
				Engenharia Aeroespacial			Turismo				
				Engenharia Biomédica							

Conheça também:

🚩 GPTs Teses UK

Descubra GPTs que oferecem acesso a 627.035 teses acadêmicas do Reino Unido, cobrindo o período de 1768 a 2023. Esses GPTs utilizam dados do 'ETHOS - Serviço Online de Teses do Reino Unido', um recurso fornecido pela Biblioteca Britânica, disponível sob a licença Creative Commons Atribuição (CC BY 4.0).

GPT LearningLadder - Uma Ferramenta de Aprendizagem

Embarkue na jornada da LearningLadder de 10 degraus, onde você pode mergulhar em um tópico de interesse e enfrentar questões cada vez mais complexas. Ganhe pontos parciais por cada tentativa e receba explicações abrangentes para enriquecer seu conhecimento.

Interface do Diretório de GPTs da Base de Teses e Dissertações

1. Coleta e Transformação dos Dados

Os GPTs Pós*BR dão acesso a um acervo de 1.401.067 produções acadêmicas brasileiras, que abrangem o período de 1987 a 2021. Estes dados são provenientes do 'Catálogo de Teses e

Dissertações - Brasil', um recurso fornecido pela CAPES e disponível sob a licença Creative Commons Attribution (CC BY).

Os primeiros desafios envolveram a preparação dos dados para serem utilizados na criação dos GPTs. O processo de preparação dos dados consistiu, inicialmente, na concatenação das bases, pois a CAPES oferece 4 arquivos com quantidade de campos variável. Por isso, foi necessária a compatibilização dos campos para o conjunto de campos mais recente. Além disso, foi necessária também a normalização das datas, uma tarefa complexa, pois as datas eram registradas em diversos formatos.

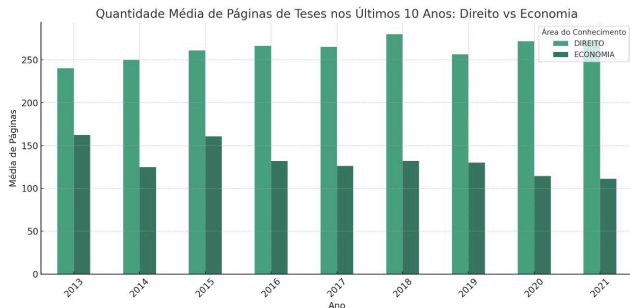
Após essa etapa, os dados foram particionados por área e grandes áreas do conhecimento, gerando um arquivo CSV para cada divisão. Esta etapa segmentou os metadados para a alimentação dos GPTs com dados relevantes de cada área de conhecimento. Para finalizar o processo, o GPT Builder foi utilizado para fazer o upload do arquivo CSV, definir as instruções de operação e criar os GPTs propriamente ditos.

O GPT Builder é uma ferramenta avançada de desenvolvimento de IA que simplifica a criação de aplicações de diálogo especializadas, utilizando o modelo de linguagem GPT-4, como os GPTs Pós*BR. Essa ferramenta permite aos usuários definir parâmetros

INOVAÇÃO

You
Crie um gráfico para comparar a quantidade média de páginas de teses nos últimos 10 anos, comparando o Direito com a Economia. Desconsidere produções com mais de 10.000 páginas.
< 2 / 2 >

PósCiênciasSociaisAplicadaBR



Aqui está o gráfico comparando a quantidade média de páginas de teses nos últimos 10 anos nas áreas de Direito e Economia. Cada barra representa a média anual de páginas para cada uma dessas áreas, permitindo uma comparação visual entre as tendências desses dois campos ao longo do tempo. [-]

📄 📄 📄 📄

Prompt para criar um gráfico comparativo da quantidade de páginas média em Direito e Economia com o GPT
PósCiênciasSociaisAplicadaBR.

Uma das funcionalidades mais interessantes é a possibilidade de exportar os metadados para uma planilha no formato Excel. Após filtrar e apresentar alguns dados de um tópico específico, o usuário pode, no comando seguinte, solicitar a geração da planilha com todos os dados.

You
Liste as teses que tratam do tema "mutação constitucional". Apresente título, instituição e ano e autor.

PósDireitoBR
Encontrei algumas teses que tratam do tema "mutação constitucional". Aqui estão os detalhes de algumas delas:

- Título:** "MUDANÇA CONSTITUCIONAL: O BRASIL PÓS-88"
Autor: Marcus Vinicius Martins Antunes
Instituição: Pontifícia Universidade Católica do Rio Grande do Sul
Ano: 2001
- Título:** O sigilo bancário enquanto norma constitucional no novo cenário econômico-social
Autor: Melissa Folmann
Instituição: Pontifícia Universidade Católica do Paraná
Ano: 2003
- Título:** MUTAÇÃO CONSTITUCIONAL, DEMOCRACIA E A AMPLIAÇÃO DOS PODERES DO SUPREMO TRIBUNAL FEDERAL
Autor: Carlos Henrique Ramos
Instituição: Universidade Gama Filho
Ano: 2012
- Título:** MUTAÇÃO CONSTITUCIONAL E A DEFESA DA FORÇA NORMATIVA DA CONSTITUIÇÃO: UM ESTUDO NO CONTEXTO DO CONSTITUCIONALISMO MODERNO E DO CONSTITUCIONALISMO CONTEMPORÂNEO
Autor: Adriana Inomata
Instituição: Centro Universitário Autônomo do Brasil
Ano: 2012
- Título:** TEORIA DA MUTAÇÃO CONSTITUCIONAL E A JURISPRUDÊNCIA DO SUPREMO TRIBUNAL FEDERAL
Autor: Gustavo Martins Bainei
Instituição: Universidade Federal do Rio Grande do Sul
Ano: 2012

Estes são apenas alguns exemplos e há mais teses disponíveis no dataset sobre este tema. [-]

📄 📄 📄 📄

Message PósDireitoBR...

ChatGPT can make mistakes. Consider checking important information.

Solicitação de teses sobre "mutação constitucional"

INOVAÇÃO

You
Gere uma planilha com todos os resultados da pesquisa acima.

PósDireitoBR
Para gerar uma planilha com todos os resultados da pesquisa sobre teses que tratam do tema "mutação constitucional", seguirei os seguintes passos:

1. Preparar os dados filtrados.
2. Salvar os dados em um arquivo CSV.
3. Disponibilizar o link para download do arquivo CSV.

Vou começar agora com a preparação e salvamento dos dados.

A planilha contendo todos os resultados da pesquisa sobre teses que tratam do tema "mutação constitucional" foi criada. Você pode baixá-la através do seguinte link:

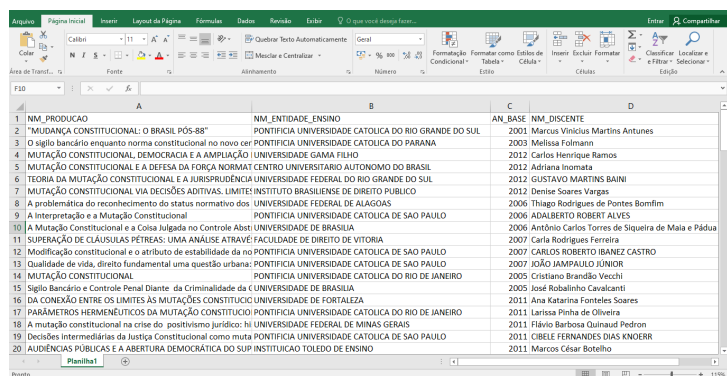
[Download da Planilha - Teses sobre Mutação Constitucional \[-\]](#)



Message PósDireitoBR...

ChatGPT can make mistakes. Consider checking important information.

Solicitação de geração do resultado completo numa planilha



Resultado da solicitação dos resultados em planilha com metadados

Em meados de dezembro de 2023, o projeto deu um passo significativo com o desenvolvimento de uma maneira mais rápida de pesquisar a base de dados completa, considerando o campo Resumo. Antes dessa inovação, a técnica anterior levava mais de um minuto para ler os dados do CSV e a base mais geral não possuía o campo

Resumo. Esta limitação foi superada com a criação de um servidor externo com um serviço de busca textual que indexa os campos das produções científicas e que se comunica com um GPT específico.

A tabela a seguir resume um comparativo das características do GPT PósBDTBR (mais lento e com respostas completas) e GPT PósResumoTextoBR (mais rápido e com as respostas mais relevantes).

Característica	GPT PósBDTBR	GPT PósResumoTítuloBR
Pesquisa em resumos e palavras-chave	Não	Sim
Necessidade de busca em servidor externo	Não	Sim
Velocidade de resposta	Lenta, de acordo com o tamanho da base	Rápida
Relevância vs Completude	Retorna resultados completos	Retorna os resultados mais relevantes
Exportação de metadados para planilhas	Exporta os dados completos	Exporta apenas dados mais relevantes

Comparação PósBDTBR vs PósResumoTítuloBR

Em resumo: é possível pesquisar a base completa de duas maneiras: com a pesquisa no CSV, por meio do GPT PósBDTBR, ou com a pesquisa no banco

