



Conheça o ChatTCU



Introdução

O ChatTCU é uma ferramenta de inteligência artificial desenvolvida pelo Tribunal de Contas da União (TCU) para auxiliar nas atividades diárias dos servidores e colaboradores. Ele funciona como um chatbot, ou seja, uma aplicação que simula uma conversa humana, permitindo que usuários obtenham respostas e orientações rápidas sobre serviços administrativos, jurisprudência selecionada, normativos, entre outros temas relevantes ao TCU e sua atuação. Além disso, o ChatTCU é também uma plataforma de agentes de IA, que permite a automação de tarefas e a execução de fluxos de trabalho mais complexos, por meio da orquestração de agentes especializados que podem interagir entre si e com sistemas internos, ampliando significativamente suas capacidades além da simples conversação.



Versões e Evoluções

1. Versão 1:

- Lançamento: março de 2023.
- Funcionalidade: Chatbot que usava o modelo GPT-3.5 da OpenAI para responder aos prompts do usuário.
- Plataforma de uso: Apenas Microsoft Teams.
- Limitações: Não havia garantia de sigilo de dados. Usava direto o modelo da OpenAI.

2. Versão 2:

- Lançamento: julho de 2023.
- Evolução: Passou a usar o modelo disponível na Azure (nuvem Microsoft), garantindo o sigilo das informações. Passou a ser uma aplicação web.
- Plataforma de uso: Microsoft Teams, Navegadores Web e Dispositivos Móveis.
- Novas funcionalidades: Feedback dos usuários.

3. Versão 3:

- Lançamento: outubro de 2023.
- Evolução: Passou a usar o modelo GPT-4 32K. Integração com bases de dados de sistemas do TCU.
- Plataforma de uso: Microsoft Teams, New Teams, Navegadores Web e Dispositivos Móveis.
- Novas Funcionalidades: Respostas sobre jurisprudência selecionada e serviços administrativos. Sumarização de documentos e peças processuais.

4. Versão 4:

- Lançamento: abril de 2024
- Evolução: Passou a usar o modelo GPT-4 Turbo 128K com maior capacidade de processamento. Integração com novas bases.
- Plataforma de uso: Microsoft Teams, New Teams, Navegadores Web e Dispositivos Móveis.
- Novas Funcionalidades: Respostas sobre normativos. Upload de documentos.

5. Versão 5:

- Lançamento: outubro de 2024
- Evolução: Melhorias de performance da aplicação para centenas de chats. Mais rápido para indexar arquivos grandes para perguntas. Passou a usar o novo modelo da OpenAI (ada-003-large) para melhora de respostas a perguntas. Botão 'Parar' adicionado.
- Plataforma de uso: Todas as versões do Teams, Navegadores Web e Dispositivos Móveis.
- Novas funcionalidades: Arquivar e desarquivar chats. Compartilhar chats e continuar conversas com ou sem arquivos. Integração com o e-TCU e TCU Mobile via API. Possibilidade de escolher especialistas para interação.

6. Versão 6:

- Lançamento: maio de 2025
- Evolução: Mais personalização e informação dos modelos de linguagem: exibição do "raciocínio" dos LLMs que possuem esta habilidade; exibição do "raciocínio" dos Agentes; personalização do prompt de sistema, temperatura (criatividade da IA) e da quantidade de trechos utilizados para contextualizar o LLM. Passou a disponibilizar LLMs provindos de três nuvens (Azure, AWS e Google Cloud), possibilitando o acesso a todos os principais modelos de linguagem do mundo em um só lugar.
- Plataforma de uso: Microsoft Teams, Navegadores Web e Dispositivos Móveis.
- Novas funcionalidades: Agentes especialistas de IA (GABI, AEQ e Minuta de Voto). Input do usuário aceitando imagens e texto simultaneamente. 17 integrações com aplicações do TCU via API (entre prontas e em desenvolvimento).

roadmap ChatTCU

Lançamento da primeira versão do ChatTCU, com uso apenas na plataforma Teams, sem sigilo garantido

09/03/2023

Criação do Núcleo de IA, uma estrutura mais robusta com o objetivo de se tornar um centro de referência para IA dentro do Tribunal

18/04/2023

Utilizando a tecnologia de GPT-4 32K, é lançada a versão 3 do ChatTCU. A ferramenta agora se integra aos demais sistemas do TCU, além de ter opções de formatação mais agradáveis

24/10/2023

Lançamento do ChatTCU v2, utilizando a nuvem Azure da Microsoft, além de se tornar acessível como aplicação web. A ferramenta agora tem novas funcionalidades

10/07/2023

Agora com a tecnologia de GPT-4 128K, a versão 4 do ChatTCU conta com um sistema robusto de envio de documentos para perguntas, pode responder sobre normativos e portarias do TCU, além de gerar consultas SQL à partir de linguagem natural. Pode ser integrado a diversas outras ferramentas

30/04/2024

31/10/2024

Contando com novas funcionalidades, o ChatTCU 5.0 agora garante maior sigilo das informações fornecidas. Além disso, é possível escolher qual base deseja utilizar para realizar sua consulta, migrar entre modelos de aprendizado, e muito mais

Com novos modelos e agentes especializados, o ChatTCU 6.0 permite agora o uso de imagens nas interações e amplia ainda mais suas funcionalidades. Além disso, traz melhorias na oferta de modelos de linguagem, recursos exclusivos por perfil de usuário e novas integrações com soluções internas do TCU via API.

12/05/2025

Visão Futura

Não paramos por aqui. O ChatTCU continua avançando e sendo desenvolvido para a nossa comunidade como uma ferramenta importante na produtividade do dia a dia

Principais Funcionalidades

1. Pesquisa e Recuperação de Informações:

- O modelo identifica a intenção do usuário e busca informações nas bases do TCU ou na sua base de conhecimento geral.
- Utilização de tecnologia RAG (Retrieval-Augmented Generation). Os dados das bases são indexados em um banco textual e vetorial, os trechos mais relevantes são recuperados desse banco e são passados para o modelo gerar as respostas dos prompts dos usuários.

2. Integração com Bases de Dados Internas:

- O ChatTCU está integrado com várias bases de dados do TCU, permitindo acesso e consultas rápidas sobre jurisprudência selecionada, normativos (Instruções Normativas, Portarias, Resoluções e Decisões Normativas) e serviços administrativos.
- Permite também que o usuário solicite resumo de documentos e peças processuais do sistema e-TCU.

3. Sumarização de documentos:

- O ChatTCU auxilia a compreensão de documentos extensos por meio da funcionalidade de sumarização.
- O usuário pode solicitar o resumo de documentos e peças processuais do sistema e-TCU.

4. Upload de Documentos:

- Permite que os usuários façam upload de arquivos PDF, DOCX, XLSX e CSV para serem utilizados em consultas.
- Nessa funcionalidade também é usada a técnica de RAG. A ferramenta divide o documento em trechos, faz a indexação desses e, no momento da consulta, recupera os trechos mais relevantes para serem usados na resposta gerada pelo LLM.
- Todas as respostas apresentadas ao usuário trazem referências explícitas aos trechos originais dos documentos, garantindo transparência e rastreabilidade das informações.
- A partir da versão 6, a funcionalidade passou a oferecer suporte ao upload de arquivos MP3 e MP4, permitindo gerar transcrições,

relatórios e realizar consultas baseadas no conteúdo desses arquivos. Além disso, passou a aceitar links do YouTube, possibilitando as mesmas funcionalidades de transcrição, análise e geração de respostas com base no conteúdo dos vídeos.

5. Feedback do Usuário:

- A ferramenta permite que os usuários deem feedback sobre as respostas geradas, ajudando a melhorar continuamente a precisão e relevância das respostas.

6. Interface do Usuário:

- Possui uma interface de diálogo textual (chat) bastante simples para interações rápidas, com texto gerado pelo modelo, mas enriquecido de referências a documentos confiáveis para fácil verificação da veracidade das informações pelo usuário.

7. Compartilhamento de Chats:

- Permite aos usuários compartilhar conversas, respostas do modelo e prompts com colegas.

8. Arquivamento de Chats:

- A fim de melhorar a usabilidade, a ferramenta possibilita que chats sejam arquivados e retirados da lista de chats ativos do usuário.

9. Geração de Atas de Reunião:

- Por meio do agente GABI, permite que o usuário tenha acesso a atas de reunião a partir de transcrições ou de vídeos fornecidos à ferramenta.

10. Geração de Minuta de Voto:

- A ferramenta, por meio de um de seus agentes, disponibiliza uma minuta de voto para recursos de reconsideração a partir de peças do e-TCU indicadas pelo usuário.

11. Integração com Outras Aplicações:

- A API do ChatTCU permite que outras aplicações do TCU tenham acesso às funcionalidades da ferramenta.

Funcionalidades em desenvolvimento e Próximos passos

- 1. Expansão do uso da API do ChatTCU;**
- 2. Suporte aos protocolos MCP (Model Context Protocol) e A2A (Agent-to-Agent);**
- 3. Criação da Loja de Agentes:**
 - A ferramenta disponibilizará um ambiente onde os usuários poderão acessar, escolher e utilizar diferentes agentes especializados de IA, cada um criado para resolver tarefas ou atender demandas específicas.
 - Na prática, a Loja de Agentes funcionará como um catálogo ou marketplace interno, onde os servidores poderão buscar e utilizar agentes prontos, otimizando seu trabalho. Além disso, ela facilitará o compartilhamento de agentes criados por diferentes áreas do TCU, promovendo a colaboração, padronização e reutilização de soluções de IA.
- 4. Descentralização da criação de agentes:**
 - Disponibilização de um ambiente que permitirá aos usuários desenvolverem seus próprios agentes personalizados de forma autônoma.

Tecnologias Utilizadas Atualmente

- 1. Arquitetura de comunicação entre frontend e backend: Microserviços (REST/JSON);**
- 2. Frontend: TypeScript com a biblioteca principal React;**
- 3. Backend: Python, com o framework rest FastAPI e os framework IA Langchain e LangGraph;**
- 4. Auditoria dos dados: Elasticsearch on-premise;**
- 5. Nuvens: Microsoft Azure, Google Cloud Platform (GCP), AWS;**
- 6. Principais serviços da nuvem Azure: Azure AI services (Azure OpenAI, AI Search), Azure Container Apps, Azure Functions, Azure API Management, Azure CosmosDB, Azure Storage Account**
- 7. Modelo de embeddings para transformar texto em vetores: text-embedding-ada-003-large.**
- 8. LLMs suportados: GPT-4o, GPT-4.1, GPT-o1, GPT-o3, GPT-o4-mini, Gemini 2.5 Pro, Gemini 2.5 Flash, Claude 3.7 Sonnet, DeepSeek R1.**

Perguntas Frequentes

Quem pode utilizar o ChatTCU?

A ferramenta está disponível para todos os servidores e colaboradores do TCU.

Servidores de outros órgãos podem acessar o ChatTCU?

O ChatTCU está restrito ao público interno do TCU. Entretanto, o TCU está disponibilizando o código-fonte da ferramenta para que outros órgãos possam ter acesso as funcionalidades desenvolvidas.

Como solicitar a cessão do código-fonte do ChatTCU?

Para solicitar o código-fonte, é necessário enviar um Ofício à presidência do TCU, assinado pela autoridade máxima do órgão solicitante. O Ofício deve conter uma solicitação expressa da cessão do código-fonte, além de indicar os e-mails, nomes e telefones de contatos para resposta. Ressaltamos que documentos enviados ao TCU devem ser protocolados via Protocolo Eletrônico, pelo link: [Protocolo eletrônico | Portal TCU](#). Após o protocolo, solicitamos que uma cópia do Ofício protocolado, contendo o número do registro de entrada, seja enviada por e-mail para setid@tcu.gov.br.

Posso utilizar o ChatTCU em dispositivos móveis?

Sim, a ferramenta está disponível para uso em navegadores, dispositivos móveis, além do Microsoft Teams e New Teams.

Como a ferramenta garante a confidencialidade dos dados?

A partir da versão 2, o ChatTCU passou a utilizar a nuvem Microsoft Azure, que oferece garantias contratuais de sigilo das informações trafegadas.

Como ter certeza de que a informação trazida pela ferramenta sobre jurisprudência selecionada e/ou normativos é verdadeira?

As respostas fornecidas pelo ChatTCU sobre jurisprudência e sobre os normativos são baseadas em trechos reais desses documentos e as respostas sempre trazem referência a esses trechos, permitindo aos usuários verificarem a origem da informação diretamente.

O que é a temperatura do modelo e como ela afeta as respostas?

A temperatura é um parâmetro que controla o nível de aleatoriedade nas respostas do modelo. No ChatTCU, a temperatura é ajustada para zero para minimizar aleatoriedade e melhorar a precisão.

Como os documentos são processados e indexados?

Os documentos são quebrados em trechos menores, indexados em um banco de dados vetorial e textual (atualmente usamos o AI Search), permitindo uma recuperação eficiente durante as consultas.

Como é feita a atualização das bases de dados?

As informações das bases de dados (jurisprudência selecionada, normativos e serviços administrativos) as quais o ChatTCU está integrado são atualizadas no AI Search periodicamente e de forma automatizada para garantir que as informações estejam sempre corretas e atualizadas.



Contato

Para mais informações ou suporte, entre em contato com a equipe do NIA (Núcleo de Inteligência Artificial) do TCU pelo e-mail nia@tcu.gov.br.



TRIBUNAL DE CONTAS DA UNIÃO 

