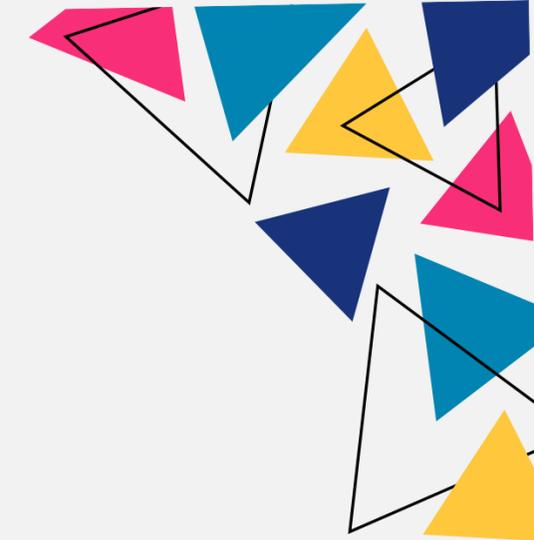


Uso de IA no Desenvolvimento de Aplicações

PALESTRANTE:
WAGNER BORGES



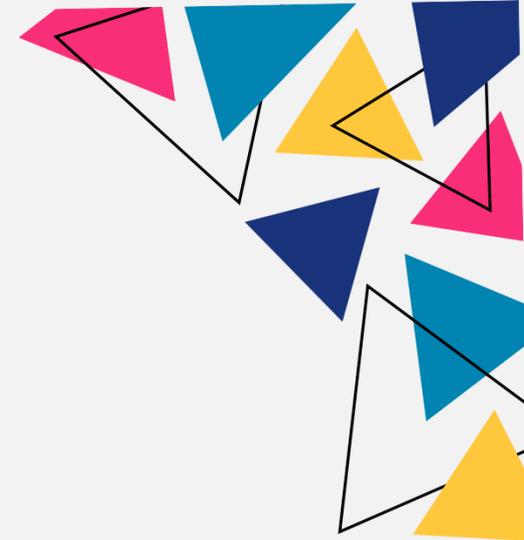
Caso Prático

Processo de Concessão de Isenção do IPVA: Deferimento Automático



Contexto

- **Projeto IPVA Digital**
 - Isenção de IPVA.
- Para solicitar a **isenção do IPVA**, o contribuinte deve **anexar uma série de documentos comprobatórios** que atestam a legitimidade da concessão do benefício.
- Estes documentos **precisam ser validados pelos analistas da SEF**, o que gera um trabalho manual e bem demorado na maioria dos casos.
- Somente após a completa validação dos documentos, os benefícios poderão ser concedidos ao contribuinte.



Contexto

Este é o formulário onde o contribuinte anexa os documentos comprobatórios para solicitar a **isenção do IPVA**

Note que vários documentos precisam ser anexados, cada um com suas peculiaridades

The screenshot shows the 'Portal de Serviços' interface for an IPVA exemption request. The progress bar at the top indicates the current step is 'Documentação' (4), with previous steps 'Tipo de Isenção', 'Dados Pessoais', and 'Dados do Veículo' completed, and 'Verificação' (5) remaining. The main section is titled 'Envio de documentos necessários do beneficiário'. It includes a dropdown menu for 'Documento Representando Legal/Procurador' and a list of documents to be uploaded, each with an 'Anexar' button and specific requirements (e.g., file size and format).

Documento	Requisitos	Ação
Comprovante de Representação	(Máx.: 3 MB, formatos: PDF, JPG e PNG)	Anexar
Comprovante de Inscrição no INSS	(Máx.: 8 MB, formatos: PDF)	Anexar
Extrato Previdenciário INSS	(Máx.: 8 MB, formatos: PDF)	Anexar
Carteira Nacional de Habilitação	(Máx.: 8 MB, formatos: PDF)	Anexar
Autorização de Isenção de IPI - Receita Federal	(Máx.: 8 MB, formatos: PDF)	Anexar
Comprovante de Exercício da Profissão (Município)	(Máx.: 8 MB, formatos: PDF)	Anexar
Comprovante de Residência	(Máx.: 8 MB, formatos: PDF)	Anexar

At the bottom of the form, there are three buttons: 'SALVAR E CONTINUAR DEPOIS', 'ETAPA ANTERIOR', and 'FINALIZAR SOLICITAÇÃO'.

Jornada

- **Uso de Solução baseada em OCR**
- Em FEV/24, começamos a analisar o que havia de tecnologia para buscar informações em documentos.
- Inicialmente avaliamos tecnologias opensource para OCR (**PyTesseract**).
- **Testamos Soluções proprietárias**
 - AWS Textract
 - Azure Document Intelligence e
 - Oracle Vision OCR
- **Por quê?**
 - Os documentos anexados pelos contribuintes em sua grande maioria são em formato de imagem.



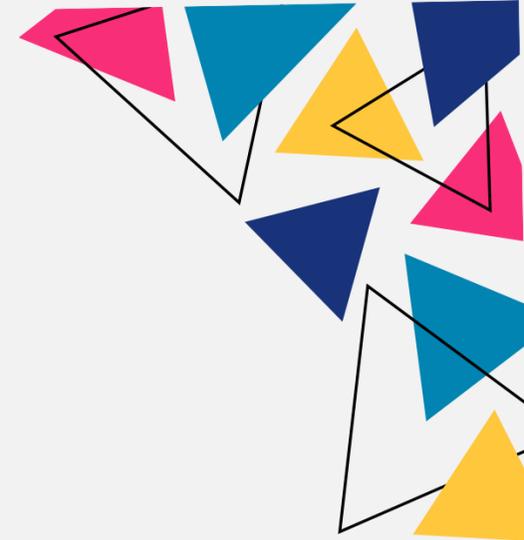
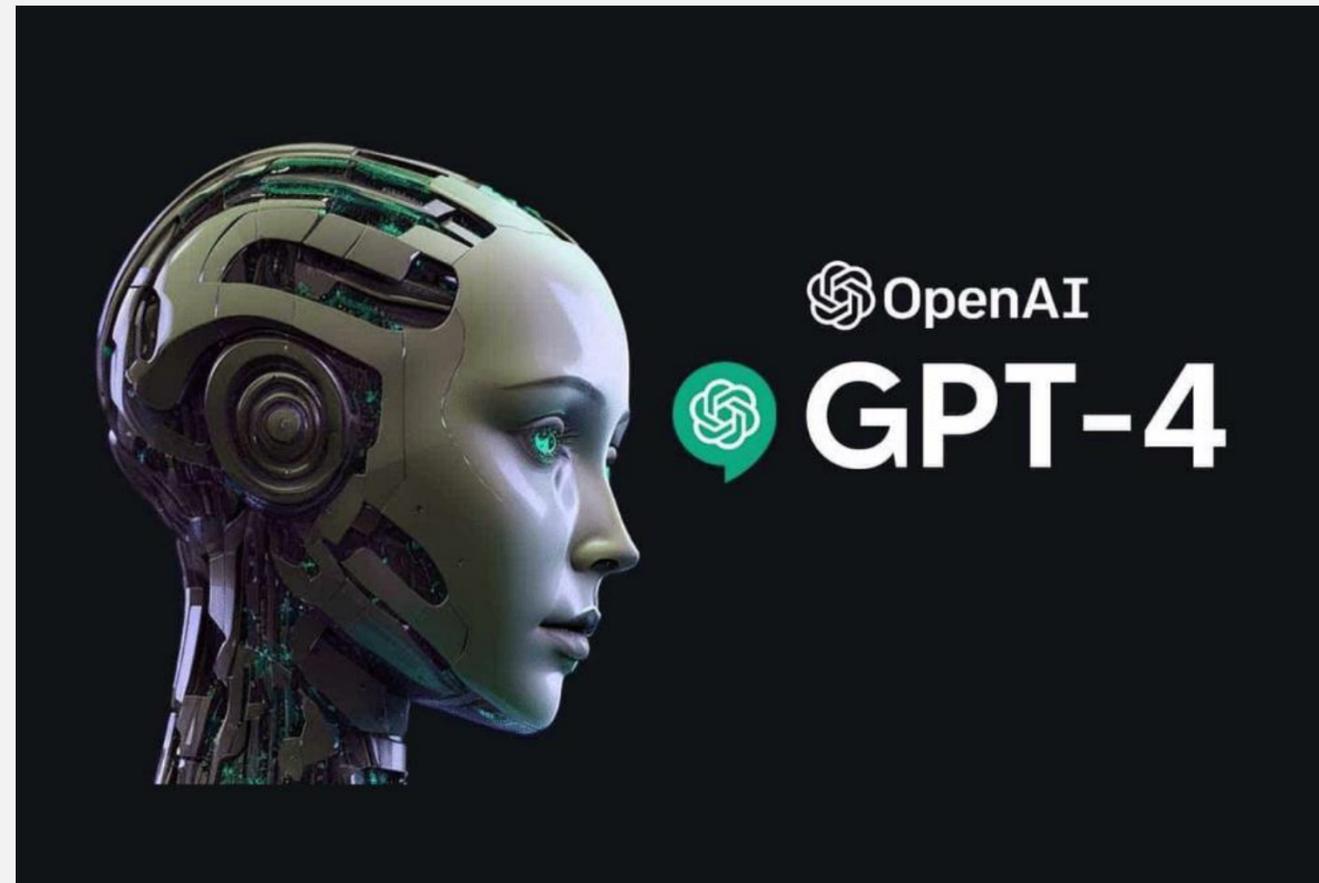
Jornada

- Comparativo entre as Soluções baseadas em OCR

	PyTesseract	AWS Textract	Azure Document Intelligence	Oracle Vision ORC
Fornecedor	Open source	Amazon	Microsoft	Oracle
Extração de textos	Sim	Sim	Sim	Sim
Extração de manuscritos	Muito ruim	Sim	Sim	Sim
Deteção de assinaturas	Não (*)	Sim	Não (*)	Não (*)
Preço básico	R\$ 0,00	USD 1,50 por cada 1.000 páginas	USD 1,50 por cada 1.000 páginas	R\$ 5,01 por cada 1.000 páginas
Preço deteção de assinaturas	Não se aplica	USD 3,00 por 1.000 páginas	Não se aplica	Não se aplica

Jornada

- Até o GPT 3.5, somente era possível enviar texto.
- A partir do GP4 foi possível enviar texto e imagem
 - As tecnologias OCR ficaram obsoletas



Detalhamento do Processo com IA

- Para **reduzir** ou **eliminar totalmente** o trabalho manual dos analistas, propõe-se a utilização da **Inteligência Artificial (IA)** para validar os documentos enviados pelos contribuintes.
- Neste contexto, existem dois tipos de tratamentos:

Tratamento	Detalhes	Validação humana
Deferimento automático	A IA garantirá que o documento é realmente válido	Opcional
Pré-validação	A IA validará informações relevantes do documento, retornando se determinado campo está ou não conforme o esperado	Necessária

Detalhamento do Processo com IA

- Qual o ganho?

– Os benefícios observados com a implementação da ideia estão categorizados por cada uma **das *personas envolvidas***, conforme detalhado a seguir:

Persona	Deferimento automático	Pré-validação
Contribuinte	Será possível conhecer o status de validação do documento de forma imediata	Será possível saber se algum detalhe precisa ser corrigido, antes mesmo do documento chegar até o analista
Analista	Será possível eliminar o trabalho manual necessário para a validação do documento	Será possível reduzir o trabalho manual necessário para a validação do documento

Detalhamento do Processo com IA

- Algumas estatísticas relevantes

- São gerados **aproximadamente 45.000 protocolos de isenção de IPVA** por ano
- **Pelo menos 7 documentos** comprobatórios são necessários em cada protocolo
- Se o **tempo para analisar cada documento for de 5 minutos**
 - Conta de Padeiro:
 - 45.000 protocolos x 7 docs x 5 minutos = 1.575.000 minutos
 - 1.575.000 minutos = 26.250 horas
- Trabalho manual estimado para a análise de documentos no processo de isenção de IPVA, **26.250 horas/ano**

Detalhamento do Processo com IA

- **Deferimento Automático**

– Documentos em que o processo de deferimento automático obteve 100% de precisão:

Documento	Detalhes	Estágio atual
Autorização da Receita Federal para aquisição de veículo com isenção do IPI	A IA recupera o CPF e o Código de Validação localizado no rodapé do documento. Com base nestas informações, um robô acessa o site SISEN da receita federal e realiza o processo de validação .	Implementado/testado
Carteira Nacional de Habilitação	A IA recupera o Nome, N° Registro, CPF, Código de Validação e Nome da Mãe . Na sequência, um robô acessa o portal de serviços do SENATRAN utilizando uma chave GOV.BR . Este portal retorna se a CNH é válida ou não.	Implementado/testado
Comprovante de Inscrição do INSS Extrato previdenciário do INSS	A IA recupera o CPF, Ocupação e Código de validação localizado no rodapé do documento. Na sequência, um robô acessa o site MEU.INSS.GOV.BR para validar as informações	Implementado/testado

Detalhamento do Processo com IA

- Deferimento Automático

MINISTÉRIO DA ECONOMIA
SECRETARIA ESPECIAL DA RECEITA FEDERAL DO
DELEGACIA DA RECEITA FEDERAL DO BRASIL EM RECIFE

DATA DE
29/12/2021

AUTORIZAÇÃO DE ISENÇÃO DE IPI PARA TÁXI - CONDUTOR AUTÔNOMO

PROTOCOLO SISEN Nº: _____ DATA DA TRANSMISSÃO: 29/12/2021 17:37:49

IDENTIFICAÇÃO DO REQUERENTE

NOME: _____ CPF: _____

AUTORIZAÇÃO

Tendo em vista o requerimento apresentado pelo(a) interessado(a) acima identificado(a) e os documentos constantes do processo supracitado, **reconheço o direito à isenção do imposto sobre Produtos Industrializados (IPI)**, prevista na Lei nº 8.989, de 24 de fevereiro de 1995.

Autorizo a aquisição, com isenção de IPI, de um automóvel de passageiros de fabricação nacional, equipado com motor de cilindrada não superior a 2.000 cm³ (dois mil centímetros cúbicos), de, no mínimo, 4 (quatro) portas, inclusive a de acesso ao bagageiro, classificado na posição 87.03 da Tabela de Incidência do Imposto sobre Produtos Industrializados (TIPI), movido a combustível de origem renovável, sistema reversível de combustão ou híbrido, ou por propulsão elétrica, nos termos da Lei nº 8.989, de 24 de fevereiro de 1995.

Observações:

1) Fica ressalvado o direito de a Secretaria Especial da Receita Federal do Brasil exigir, a qualquer tempo, o imposto que deixou de ser pago na aquisição do veículo, com os acréscimos legais, se constatar que o beneficiário não atendeu aos requisitos legais, sem prejuízo das sanções penais cabíveis.

2) A presente Autorização não gera direito adquirido (art. 179, § 2º do CTN) e as regras de isenção são aplicáveis no momento da ocorrência do fato gerador (data da saída do veículo constante da nota fiscal). Por esses motivos, deve-se verificar a validade da Autorização, antes da emissão da nota fiscal, o que pode ser feito em consulta ao SISEN pelo interessado, ou pela validação do código de autenticidade presente na Autorização.

ASSINATURA

DELEGACIA DA RECEITA FEDERAL DO BRASIL EM RECIFE

Nome: _____
Cargo: _____
Matri: _____

(* Equipe SISEN - PORTARIA SRRF4 Nº 50.)

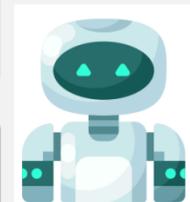
Código de Autenticidade:
DPBÉ.KCKB.....PCBK



IA

CPF

CÓDIGO



ROBÔ

CPF

CÓDIGO

DOCUMENTO VÁLIDO |
DOCUMENTO INVÁLIDO

Receita Federal
Sisen

26/09/2024 - 4.5.1
Tempo restante: 14:54

Validar autorização

Informe CPF do documento:

Insira o código de autenticação:

Voltar Limpar Consultar

Detalhamento do Processo com IA

- **Deferimento Automático**

- Aqui o resultado da validação atestando que o documento analisado está válido

```
Run: check-ipi-auto-ocr x
C:\Users\guilherme.timoteo\PycharmProjects\SEF-OpenAI\.venv\Scripts\python.exe C:\Users\guilherme.timoteo\PycharmProjects\SEF-OpenAI\Auto-IA-Isencao\check-ipi-auto-ocr.py
----- IA>> PROCESSANDO OCR -----
1- Extraindo textos via OCR. Documento: ' ./isencao-ipi-3.pdf '
2- CPF presente no documento
3- Obtendo código de autenticidade... Código recuperado: ██████████
----- VALIDACAO RECEITA FEDERAL-----
4- Acessando site de validação da Receita Federal: https://www.sisen.receita.fazenda.gov.br/sisen/consultaCodigoAutorizacao.jsf
5- Informando CPF: ██████████ e código de autenticidade: ██████████
6- Resultado da consulta na receita federal:
<<DOC_VALIDO>>
-----

Tempo: 29 s

Process finished with exit code 0
```

Detalhamento do Processo com IA

- **Pré-Validação**

- Para alguns tipos de documentos o **deferimento automático não será possível** por razões diversas.

- Porém, vários campos e condições destes documentos podem ser “**pré-validadas**” a fim de reduzir o trabalho dos analistas.

- Para exemplificar este tratamento, iremos abordar o documento “**Comprovante de exercício de profissão fornecido pelo município**” que é um dos requerimentos para a isenção de IPVA para taxistas.

Detalhamento do Processo com IA

- **Pré-Validação**
- O “**Comprovante de exercício de profissão fornecido pelo município**”;
 - Atestar que determinado contribuinte exerce a **profissão de taxista no município** em questão.
- Qual o desafio de tratar este tipo de documento?
 - Minas Gerais tem **853 municípios** e cada um destes municípios tem **seu próprio documento**.
 - Alguns municípios colocam CNH no certificado, outros colocam RG, outros somente colocam o CPF.
- Antes de executar a pré-validação propriamente dita **é preciso saber quais campos devem ser validados pela IA**.
- Identificar os campos mais comuns entre os municípios.

Detalhamento do Processo com IA

• Pré-Validação

Estatísticas Levantadas	
Número de documentos analisados (amostra)	976
Número de municípios distintos presentes na amostra	231
% de docs que possuem o nome do taxista	100%
% de docs que possuem alguma indicação de data	100%
% de docs que possuem alguma indicação do nome do município	100%
% de docs que possuem alguma indicação de taxi	100%
% de docs que possuem assinatura ou rubrica	94%
% de docs que possuem CPF do taxista	93%
% de docs que possuem a CNH do taxista	14%
% de docs que possuem o RG do taxista	13%

Foram analisados 976 documentos (em PDF) que são emitidos pelas prefeituras de MG. Por esta amostra de documentos foi possível cobrir os modelos emitidos por 231 municípios mineiros

Estes campos devem ser utilizados na pré-validação da IA, pois estão presentes em uma quantidade significativa de documentos

Estes campos **NÃO** devem ser utilizados na pré-validação da IA, pois não tem uma presença significativa nos documentos da amostra

Detalhamento do Processo com IA

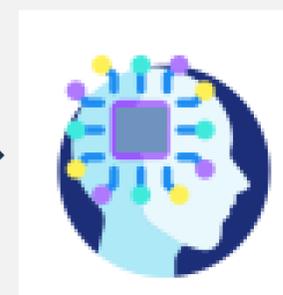
• Pré-Validação

–Fluxo da Pré-Validação



PROMPT

- 1) O município está correto?
- 2) O CPF do taxista está correto?
- 3) O nome do taxista está correto?
- 4) O documento está assinado?
- 5) Qual a data do documento?
- 6) Existe alguma menção a atividade de taxi?



```
Run: check-permissao-prefeitura
C:\Users\guilherme.timoteo\PycharmProjects\SI
1) MUNICIPIO-OK=S
2) CPF-OK=S
3) NOME-OK=S
4) ASSINADO-OK=S
5) 18/11/2021
6) TAXI-OK=S

Tempo: 4 s

Process finished with exit code 0
```

Utilizando a IA generativa conseguimos perguntar em linguagem natural (PROMPT) e recebemos um retorno apropriado.

Exemplo: MUNICIPIO-OK=S, significa que o município está corretamente indicado no documento. MUNICIPIO-OK=N indica o contrário.

Azure OpenAI Service

- **E como a SEF acessa a API da OpenAI (GPT4)?**
 - A SEF possui um **Tenant** dentro do **Azure**.
 - Dentro deste **Tenant** foi habilitado o **Azure AI Studio**.
 - Dentro AI Studio, **publicamos um modelo GPT4** que é uma espécie de clone do GPT4 público (disponível no ChatGPT).
 - No GPT4 público, as informações do prompt enviadas pelos usuários **são guardadas para serem utilizadas nos treinamentos futuros da OpenAI**.
 - No GPT4o privado do Tenant da SEF, as informações do prompt (pergunta + documentos dos contribuintes) **são descartadas e não são utilizadas no treinamento**.
 - Isto é garantido pelo contrato da SEF com a Microsoft.
 - Assim, a **informação sigilosa dos documentos do contribuintes estão totalmente protegidas**, garantindo que tal informação não saia do domínio da SEF.

Como é cobrado a API OpenAI?

- **A cobrança pode ser feita de duas formas:**
 - Por Token
 - Por Reserva de Máquina
- **Por Token**
 - Token de Entrada (Prompt + Imagens)
 - Token de Saída (Retorno da API)
 - Valor Cobrado = **Token de Entrada + Token de Saída**
 - Em média temos em torno **de 500 a 1500 tokens por chamada da API**. Isso gera uma cobrança na conta Azure da SEF de R\$ 0,01 (1 centavo) a R\$ 0,03 (3 centavos)

Como é cobrado a API OpenAI?

- **Por Reserva de Máquinas**

- Neste caso, a cobrança é feita pela reserva de um máquina **no Azure com GPUs.**
- A cobrança neste caso é **feita por tempo de uso da máquina.**
- Vc pode por exemplo, desligar a máquina de madrugada para economizar.
- O valor varia muito de acordo com o **desempenho da máquina**
- **Colocou muita GPU fica caro. Colocou pouca GPU, o modelo fica lento.**



Perguntas?

