

Engenharia de Prompts: Integrando EFD ICMS/IPI com XMLs de NF-e

Especialização em Sistemas e Agentes Inteligentes

Prof. Sanderson Oliveira de Macedo

Hugo Ferreira Ginú
Pedro Moacir de Carvalho
Rafael Ferreira Peixoto

O que é **SPED EFD ICMS/IPI**?

O SPED é Sistema Público de Escrituração Digital.

A **Escrituração Fiscal Digital - EFD** é um **arquivo digital**, que se constitui de um conjunto de escriturações de documentos fiscais e de outras informações **de interesse dos Fiscos das unidades federadas e da Secretaria da Receita Federal do Brasil**, bem como de **registros de apuração de impostos referentes às operações e prestações praticadas pelo contribuinte**.

Este arquivo deverá ser assinado digitalmente e transmitido, via Internet, ao **ambiente Sped**.

Fonte: <http://sped.rfb.gov.br/pagina/show/523>

Fluxo de trabalho e partes envolvidas

- **Empresa Fornecedora:** Emite a NF-e e registra a operação fiscal.
- **SEFAZ (Governo Estadual):** Autoriza e armazena as NF-es emitidas.
- **Empresa Compradora:** Recebe a NF-e, registra os dados no SPED e nos sistemas internos da empresa.
- **Empresa de Consultoria Tributária:** Recebe os arquivos SPED e XML da empresa compradora para análise dos tributos informados. Para esta análise ela precisa fazer o cruzamento das informações do SPED com o XML para validação.



O problema

Em alguns trabalhos relacionados a consultoria tributária, é necessário complementar as informações declaradas no SPED EFD ICMS/IPI com os dados originais dos XMLs das notas fiscais de aquisição.

Ao preencher os dados de aquisição de mercadoria nos sistemas internos da empresa, os colaboradores não respeitam a ordem sequencial dos itens informados na nota fiscal, bem como alteram o código e a descrição dos produtos para os cadastros internos da empresa.



A proposta de solução

- Utilização de LLM
 - Utilização da API da OpenAI *gpt-4o-mini-2024-07-18*
- Bibliotecas Python
 - Dotenv: para carregar as variáveis de ambiente
 - Pydantic: para executar a validação de tipagem de dados.
 - Typing: para utilização de tipos de dados.
 - Json: para conversão de objetos em JSON.
 - Guardrails-ai: biblioteca para validação de entradas e saídas de prompts.

Solução: Bibliotecas python

```
app.py > ...  
1  from guardrails import Guard  
2  from dotenv import load_dotenv  
3  from pydantic import BaseModel, Field  
4  from typing import List, Optional  
5  import json  
6
```

Solução: Prompts

- Utilização dos seguintes conceitos no prompt para resolução do problema:
 - Definição do objetivo:

```
Realize a vinculação dos itens da lista 01: <dadosEfd>"""+df_efd+"""/>dadosEfd> \n  
com os da lista 02: <dadosXml>"""+df_xml+"""/>dadosXml>. \n
```

- Uso de Delimitadores:

```
Para vinculação considere as seguintes regras: \n  
<regra 01>  
  Faça a vinculação pela similaridade considerando apenas o campo 'Descrição' da lista 01 com a lista 02.  
  Não pode existir outra ocorrência com a mesma similaridade.  
</regra 01>  
<regra 02>  
  Faça a vinculação pela aproximação do 'Valor Total', considerando o percentual da diferença entre o valor da lista 01 com o valor  
  Para calcular o percentual, siga as instruções do <exemplo>  
<exemplo>Subtraia o valor total do item 2 do valor total do item 1 (100 - 99), desconsidere o sinal de negativo.
```

Solução: Prompts

- Fornecendo exemplos:

```
<regra 02>
```

```
Faça a vinculação pela aproximação do 'Valor Total', considerando o percentual da diferença entre o valor da lista 01 com o valor da lista 02.  
Para calcular o percentual, siga as instruções do <exemplo>
```

```
<exemplo>Subtraia o valor total do item 2 do valor total do item 1 (100 - 99), desconsidere o sinal de negativo,  
e divida o resultado pelo valor total do item 1 ( $1 / 100 = 0,01$ ) e multiplique por 100 ( $0,01 \times 100 = 1$ ).
```

```
</exemplo>
```

```
Se resultado for menor que 2 (2%) faça a vinculação, desde que não tenha outro item com uma aproximação menor ou igual.
```

```
</regra 02>
```


Solução: Prompts

- Instruções em etapas

Para realizar a vinculação considere as seguintes etapas:

<etapa 1>

Para cada item da lista 01 aplique a regra 01 para todos os itens da lista 02 que ainda não tenham sido vinculados.

Ao realizar a vinculação, retire da lista 01 e da lista 02 o item vinculado.

Caso a lista 01 ou lista 02 tenha apenas um item resolução, realize a vinculação desses itens e considere como resolução a Regra 01.

</etapa 01>

<etapa 2>

Para cada item da lista 01 que não tenha sido resolvido na regra 01, aplique a regra 02 para todos os itens da lista 02 que ainda não tenham sido vinculados.

Ao realizar a vinculação, retire da lista 01 e da lista 02 o item vinculado.

Caso a lista 01 ou lista 02 tenha apenas um item resolução, realize a vinculação desses itens e considere como resolução a Regra 02.

</etapa 2>

<etapa 3>

Para cada item da lista 01 que não tenha sido resolvido na regra 01, aplique a regra 03 para todos os itens da lista 02 que ainda não tenham sido vinculados.

Ao realizar a vinculação, retire da lista 01 e da lista 02 o item vinculado.

Caso a lista 01 ou lista 02 tenha apenas um item resolução, realize a vinculação desses itens e considere como resolução a Regra 03.

</etapa 3>

\n

Solução: Guardrails

Utilização da API guardrails-ai para execução da chamada da LLM e validação do retorno da LLM para garantir que o resultado esteja com a formatação de dados esperado.

```
guard = Guard.for_pydantic(output_class=ModelReturn)

res = guard(
    model="gpt-4o-mini-2024-07-18",
    messages=[{
        "role": "user",
        "content": prompt_analise
    }]
)
```

Execução: Dados de entrada

```
df_efd = ""
Chave NFE|Seq|Código do Item|Descrição|Quantidade|Valor UN|Valor Total
'31240161365557000110550010009047751102632318'|1|000001|MANTEIGA AVIAÇÃO TABL.S/S C/24|24|292,95|7030,8
'31240161365557000110550010009047751102632318'|2|000007|MANTEIGA AVIAÇÃO TABL.C/S C/24|60|284,58|17074,8
'31240161365557000110550010009047751102632318'|3|000002|MANTEIGA AVIAÇÃO LATA C/S C/30|30|268,32|8049,6
'31240161365557000110550010009047751102632318'|4|000004|MANTEIGA AVIAÇÃO POTE S/S C/48|480|189,42|90921,6
'31240161365557000110550010009047751102632318'|5|000008|MANTEIGA AVIAÇÃO POTE C/S C/24|360|164,25|59130
'31240161365557000110550010009047751102632318'|6|000011|DOCE LEITE AVIAÇÃO PT|240|21,42|5140,8
'31240161365557000110550010009047751102632318'|7|001005|REQUEIJÃO COPO AVIAÇÃO ERVAS FINAS|1|147,36|147,36""

df_xml = ""
Chave NFE|Seq|Código do Item|Descrição|Quantidade|Valor UN|Valor Total
'31240161365557000110550010009047751102632318'|1|R-001|REQUEIJAO AVIACAO ERVAS F 180G|1|147,36|147,36
'31240161365557000110550010009047751102632318'|2|M-001|MANT AVIACAO TAB 24X100G C/S|60|284,58|17074,8
'31240161365557000110550010009047751102632318'|3|M-002|MANT AVIACAO POTE 24X500G C/S|360|164,25|59130
'31240161365557000110550010009047751102632318'|4|M-004|MANT AVIACAO LATA 30 X 200G|30|268,32|8049,6
'31240161365557000110550010009047751102632318'|5|M-006|MANT AVIACAO POTE 48X200G S/S|480|189,42|90921,6
'31240161365557000110550010009047751102632318'|6|D-001|DOCE LEITE AVIACAO PT 400G|240|21,42|5140,8
'31240161365557000110550010009047751102632318'|7|M-007|MANT AVIACAO POTE 24X200G S/S|24|292,95|7030,8""

analisar_vinculacao(df_efd, df_xml)
```

Execução: Dados de saída

```
{
  "itens_lista_01": [
    {
      "chv_nfe": "31240161365557000110550010009047751102632318",
      "seq": 1,
      "cod_item": "000011",
      "desc_item": "DOCE LEITE AVIAÇÃO PT",
      "qt_item": 240.0,
      "valor_un": 21.42,
      "valor_total": 5140.8,
      "vinculados": [
        {
          "seq": 6,
          "cod_item": "D-001",
          "desc_item": "DOCE LEITE AVIACAO PT 400G",
          "qt_item": 240.0,
          "valor_un": 21.42,
          "valor_total": 5140.8,
          "forma_vinculacao": "Regra 01"
        }
      ]
    },
    {
      "chv_nfe": "31240161365557000110550010009047751102632318",
      "seq": 2,
      "cod_item": "000007",
      "desc_item": "MANTEIGA AVIAÇÃO TABL.C/S C/24",
      "qt_item": 60.0,
      "valor_un": 284.58,
      "valor_total": 17074.8,
      "vinculados": [
        {
          "seq": 2,
          "cod_item": "M-001",
          "desc_item": "MANT AVIACAO TAB 24X100G C/S",

```

Github

<https://github.com/rafepe/agentes-eprompt>

Obrigado!