

PDD - Extração de Cadastro Nacional da Pessoa Jurídica (CNPJ)

1. Introdução

Proposta do documento

O documento descreve a sequência de ações executadas como parte do processo de negócio, as condições e regras do processo anterior à automação (AS IS), bem como a nova sequência de ações que o processo seguirá como resultado da preparação para a automação (TO BE).

O PDD é um documento de comunicação entre:

- O RPA Business Analyst e o SME/Proprietário do Processo. O objetivo é garantir que o RPA Business Analyst tenha o entendimento correto do processo e o represente com precisão.
- O RPA Business Analyst e a equipe de Development team (representada pelo Solution Architect e RPA Development Lead). O objetivo é garantir que o processo seja documentado adequadamente e com um nível de detalhe suficiente para que o Solution Architect possa então criar a solução com base no conteúdo do PDD

2. Objetivos

Os objetivos de negócios e benefícios esperados pelo Business Process Owner após a automação do processo de negócios selecionado são:

- Retirar a atividade manual do colaborador
- Extrair o conjunto de dados mensalmente
- Direcionar a análise humana para eventuais pendências

3. Contatos chave

O documento de especificações inclui requisitos concisos e completos do processo de negócios e é construído com base nas entradas fornecidas pelo **key user deste processo**.

Espera-se que o **Process Owner** o revise e forneça aprovação quanto à precisão e conclusão das ações, contexto, impacto e um conjunto de exceções do processo. Os detalhes devem ser incluídos na tabela abaixo.

Função	Nome	Email
Stakeholder (Diretor(a))	X	xxxxx.xxxxx@xxx.com.br
BA (Analista de Negócio)	Matheus Martins	xxxxx.xxxxx@xxx.com.br
Process Owner (Dono do Processo)	XX	xxxxx.xxxxx@xxx.com.br
Key user (ponto focal)	XXX	xxxxx.xxxxx@xxx.com.br

4. ROI

Relacionar nos campos abaixo as informações que auxiliarão no cálculo e avaliação do ROI.

Nome processo

Processo Modelo

Volumetria	37 arquivos
Periodicidade	Mensal
Qtd. Executores	1
Perfil Executor	Analista Jr
Tempo p/ Execução	30 segundos a 6 minutos por arquivo

5. Fora do escopo

- Nome de arquivos que não foram mapeados
- Envio do conteúdo dos arquivos para o a área de negócio e ao departamento de TI
- Reprocessamento automático dos arquivos em caso de falha

6. Riscos

Descrever neste campo se existe e quais os possíveis riscos para o processo.

- Não existe prazo legal para a entrega
- Não há problema na quantidade de vezes que os arquivos são baixados. Pode-se utilizar estes dados para desenvolvimento e teste
- O conteúdo dos arquivos não contém dados sensíveis, pois são informações públicas
- Independente do resultado da execução, a área de negócio precisa ser informada para ter tempo hábil para analisar

7. AS IS - Overview

Abaixo seguem informações e descritivo do processo em seus formato AS IS.

Item	Descrição
Nome completo	Extração de informações de CNPJ
Área	xxx
Departamento	xxx
Objetivo deste processo	Efetuar a baixa dos arquivos mapeados, descompactar e informar o resultado

7.1. Regras de Negócio

Descrever as regras de negócio aplicadas ao processo.

- A área de negócio deve ser informada sobre o status de cada arquivo
- Verificar a data da atualização dos dados e caso seja posterior a data do último download, prosseguir com a execução
- Quando o conteúdo do arquivo .zip é grande, o site divide em outros arquivos .zip e coloca o index (0, 1, 2, 3...) no nome do arquivo
- Descompactar arquivos com a extensão .zip e remover o arquivo compactado
- Após a conclusão da execução enviar e-mail ao departamento de TI xxxx.xxxx@xxx.com.br
- Caso ocorra mais de uma atualização para o mesmo período, sobrescrever o arquivo

7.2. Detalhamento [↗](#)

Descrever as etapas de execução do processo em seu formato AS IS (com prints das telas que são acessadas).

7.3. Entrada [↗](#)

É a relação fixa de arquivos abaixo:

Nome do recurso	Nome do arquivo	Extensão
Dicionário de Dados do Cadastro Nacional da Pessoa Jurídica	cnpj-metadados	PDF
DADOS ABERTOS CNPJ	Cnaes	ZIP
DADOS ABERTOS CNPJ	Empresas	ZIP
DADOS ABERTOS CNPJ	Estabelecimentos	ZIP
DADOS ABERTOS CNPJ	Motivos	ZIP
DADOS ABERTOS CNPJ	Municipios	ZIP
DADOS ABERTOS CNPJ	Países	ZIP
DADOS ABERTOS CNPJ	Qualificacoes	ZIP
DADOS ABERTOS CNPJ	Simples	ZIP
DADOS ABERTOS CNPJ	Socios	ZIP

7.4. Processamento [↗](#)

Abertura e verificação da data [↗](#)

O acesso é pelo link fixo <https://dados.gov.br/dados/conjuntos-dados/cadastro-nacional-da-pessoa-juridica---cnpj>



Caso a barra lateral estiver fechada, precisa abrir

Portal de Dados Abertos

dados.gov.br/dados/conjuntos-dados/cadastro-nacional-da-pessoa-juridica---cnpj



Ministério da Fazenda

O Ministério da Fazenda é o órgão da estrutura administrativa da Repúbl

137 conjuntos de dados 1 reusos

+ Seguir Contato

> Descrição

Recursos 5

Verificar se a data da atualização é posterior a data do último download, caso não seja, **encerrar o processo**

PDF

Dicionário de Dados do Cadastro Nacional da Pessoa Jurídica

O arquivo relaciona as informações que possibilitam organizar, classificar, relacionar e inferir novos dados sobre o conjunto de dados do Cadastro Nacional da Pessoa Jurídica.

Data de criação: 24/11/2022 Downloads: 55795

Acessar o recurso

Efetuar o download do arquivo PDF



Conjunto 2: Dados abertos do CNPJ

Acessar recurso

ZIP

DADOS ABERTOS CNPJ

O Cadastro Nacional da Pessoa Jurídica (CNPJ) é um banco de dados gerenciado pela Secretaria Especial da Receita Federal do Brasil (RFB), que armazena informações cadastrais das pessoas jurídicas e outras entidades de interesse das administrações tributárias da União, dos Estados, do Distrito Federal e dos Municípios.

Data de criação: 15/08/2024 Downloads: 10326

Acessar o recurso

Clicar no período mais recente (mês anterior a data da execução)

Portal de Dados Abertos

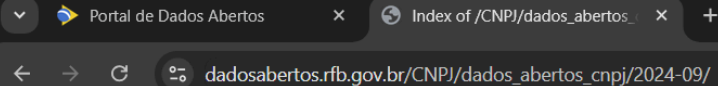
Index of /CNPJ/dados_abertos_cnpj

dadosabertos.rfb.gov.br/CNPJ/dados_abertos_cnpj/?C=N;O=D



















Index of /CNPJ/dados_abertos_cnpj

Name	Last modified	Size	Description
Parent Directory		-	
2024-09/	2024-09-18 16:20	-	
2024-08/	2024-08-14 11:42	-	
2024-07/	2024-08-15 15:32	-	
2024-06/	2024-08-16 09:58	-	
2024-05/	2024-08-20 14:23	-	

Clicar em cada arquivo mapeado para fazer o download

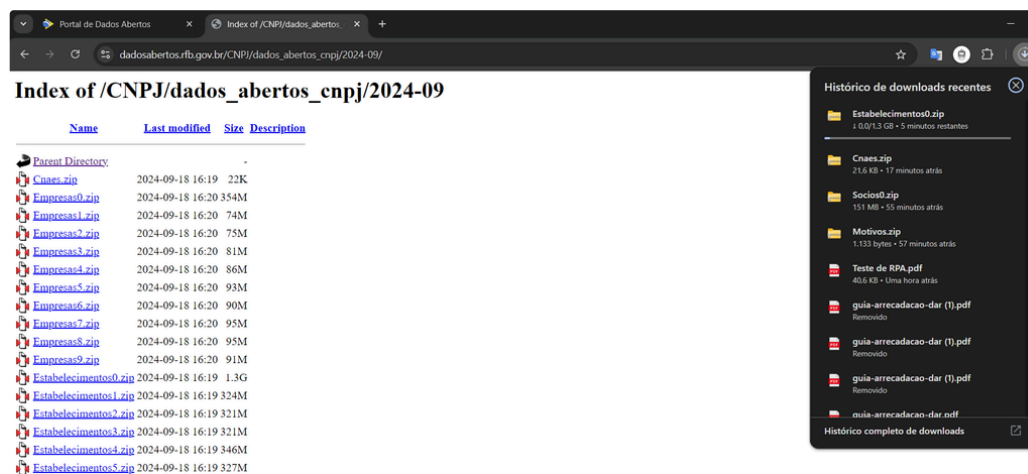


Index of /CNPJ/dados_abertos_cnpj/2024-09

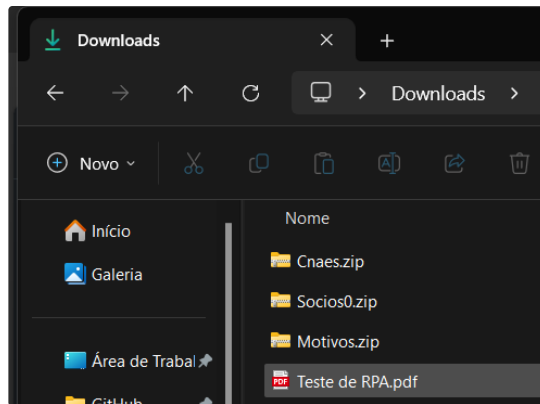
	<u>Name</u>	<u>Last modified</u>	<u>Size</u>	<u>Description</u>
	Parent Directory		-	
	Cnaes.zip	2024-09-18 16:19	22K	
	Empresas0.zip	2024-09-18 16:20	354M	
	Empresas1.zip	2024-09-18 16:20	74M	
	Empresas2.zip	2024-09-18 16:20	75M	
	Empresas3.zip	2024-09-18 16:20	81M	
	Empresas4.zip	2024-09-18 16:20	86M	
	Empresas5.zip	2024-09-18 16:20	93M	
	Empresas6.zip	2024-09-18 16:20	90M	
	Empresas7.zip	2024-09-18 16:20	95M	
	Empresas8.zip	2024-09-18 16:20	95M	
	Empresas9.zip	2024-09-18 16:20	91M	
	Estabelecimentos0.zip	2024-09-18 16:19	1.3G	
	Estabelecimentos1.zip	2024-09-18 16:19	324M	
	Estabelecimentos2.zip	2024-09-18 16:19	321M	
	Estabelecimentos3.zip	2024-09-18 16:19	321M	
	Estabelecimentos4.zip	2024-09-18 16:19	346M	
	Estabelecimentos5.zip	2024-09-18 16:19	327M	
	Estabelecimentos6.zip	2024-09-18 16:19	321M	
	Estabelecimentos7.zip	2024-09-18 16:19	321M	
	Estabelecimentos8.zip	2024-09-18 16:19	343M	
	Estabelecimentos9.zip	2024-09-18 16:19	331M	
	Motivos.zip	2024-09-18 16:19	1.1K	
	Municipios.zip	2024-09-18 16:19	42K	

7.5. Saída

Aguardar o arquivo ser disponibilizado na pasta de download



Para os arquivos com a extensão .zip precisa descompactar



Mover o arquivo descompactado para o destino (não definido)

7.6. Fluxo AS IS [↗](#)

8. Processo - TO BE [↗](#)

Descrever de forma resumida e objetiva a ideia proposta como solução TO BE deste processo.

8.1. Pré requisitos para a automação [↗](#)

- O navegador precisa abrir o PDF de forma nativa, ou seja, não pode ter outra extensão que abra PDF como o Adobe Acrobat

8.2. Entrada [↗](#)

Nenhuma alteração relevante se comparado ao AS IS

8.3. Processamento [↗](#)

É possível acessar o conjunto de dados pela URL diretamente e mesmo com a atualização de dados, o link é fixo.

- Conjunto 1: Dicionário de Dados do Cadastro Nacional da Pessoa Jurídica > <https://www.gov.br/receitafederal/dados/cnpj-metadados.pdf>
- Conjunto 2: DADOS ABERTOS CNPJ > https://dadosabertos.rfb.gov.br/CNPJ/dados_abertos_cnpj

8.4. Saída [↗](#)

Informar o status de cada arquivo, como no exemplo abaixo:

Nome do recurso	Nome do arquivo	Index	Status	Tipo de erro	Mensagem
Dicionário de Dados do Cadastro Nacional da Pessoa Jurídica	cnpj-metadados		OK		
DADOS ABERTOS CNPJ	Cnaes		OK		
DADOS ABERTOS CNPJ	Empresas	0	OK		

DADOS ABERTOS CNPJ	Empresas	1	OK		
DADOS ABERTOS CNPJ	Empresas	2	OK		
DADOS ABERTOS CNPJ	Empresas	3	ERRO	System exception	Elemento da interface gráfica não encontrado
DADOS ABERTOS CNPJ	Estabelecimentos	0	OK		
DADOS ABERTOS CNPJ	Estabelecimentos	1	OK		
DADOS ABERTOS CNPJ	Estabelecimentos	2	OK		
DADOS ABERTOS CNPJ	Motivos		ERRO	Business exception	Arquivo não encontrado

8.5. Fluxo TO BE [↗](#)

9. Cenários de Erro [↗](#)

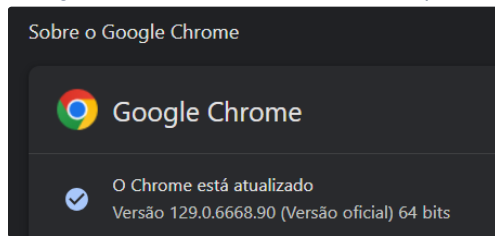
Informar os cenários de erro que forem levantados neste processo.

Ambiente	#	Cenários	Resposta ao erro	Procedimentos Adicionais
Site	1	Instabilidade de acesso e navegação lenta	Lançar uma exceção de sistema	Informar a área de negócio para tentar em outro momento
Site	2	Tempo para realizar o download maior que 10 minutos	Lançar uma exceção de sistema	Verificar manualmente se o problema é ocasionado devido a velocidade de erro ou pelo próprio site. Caso seja o site, informar a área de negócio para tentar em outro momento
Navegador	3	Após clicar no download do arquivo, pode aparecer a janela "Salvar como" ao invés do download direto na pasta download	Mapear no bot para ser capaz de lidar com o cenário	-

10. Aplicativos/ferramentas usados [↗](#)

Elencar sistemas que serão utilizados na proposta.

- Navegador Chrome. Versão utilizada no mapeamento abaixo:



12. Acessos e perfil [↗](#)

Informar os sistemas, perfis, menus/endereços que serão acessados no processo:

Sistema	Perfil Necessário	Menu	Endereço para acesso (pasta de rede)	Ticket aberto
-	-	-	-	-

13. Fontes adicionais de documentação do processo [↗](#)