

Nº 179688

Pequenas empresas LGPD: princípios fundamentais

Silvana Bentlei Murbak Cerqueira

Adriana Camargo Brito

*Palestra apresentada no LGPD e
Anonimização de dados, TEAMS,
30/06/2025. 59p.*

A série “Comunicação Técnica” compreende trabalhos elaborados por técnicos do IPT, apresentados em eventos, publicados em revistas especializadas ou quando seu conteúdo apresentar relevância pública.

PROIBIDO A REPRODUÇÃO, APENAS PARA CONSULTA.

Instituto de Pesquisas Tecnológicas do Estado de São Paulo
S/A - IPT
Av. Prof. Almeida Prado, 532 | Cidade Universitária ou
Caixa Postal 0141 | CEP 01064-970
São Paulo | SP | Brasil | CEP 05508-901
Tel 11 3767 4374/4000 | Fax 11 3767-4099

www.ipt.br

Pequenas Empresas

LGPD: Princípios Fundamentais



Silvana Bentley Murbak Cerqueira
Adriana Camargo de Brito



LGPD — significa **Lei Geral de Proteção de Dados** é a Lei 13.709/2018

Essa é a Lei que cuida dos **DADOS PESSOAIS** de
TITULARES DE DADOS

Mas, quem é titular de dados?

Titular de dado é toda **PESSOA**
FÍSICA, todos que possuem
DADOS PESSOAIS

O QUE SÃO **DADOS PESSOAIS**?

Dados que permitem a
IDENTIFICAÇÃO DE UMA PESSOA

Exemplos:

- Nome
- RG
- CPF
- Endereço
- E-mail com o nome do titular
- Endereço de IP

O QUE SÃO **DADOS PESSOAIS SENSÍVEIS**?

Dados que podem levar a
discriminação de titulares de dados
Exemplos:

- Dados de saúde
- Dados biométricos
- Dados sobre a vida sexual
- Opinião política, filiação sindical e religião

Se Aplica a LGPD

Em qualquer operação de tratamento realizada por pessoas físicas ou jurídicas de direito público ou privado, sempre que:

- A operação de tratamento seja realizada no território nacional;
- A atividade de tratamento tenha por objetivo a oferta ou o fornecimento de bens ou serviços ou o tratamento de dados de indivíduos localizados no território nacional; ou
- Os dados pessoais objeto do tratamento tenham sido coletados no território nacional.

Não se Aplica a LGPD

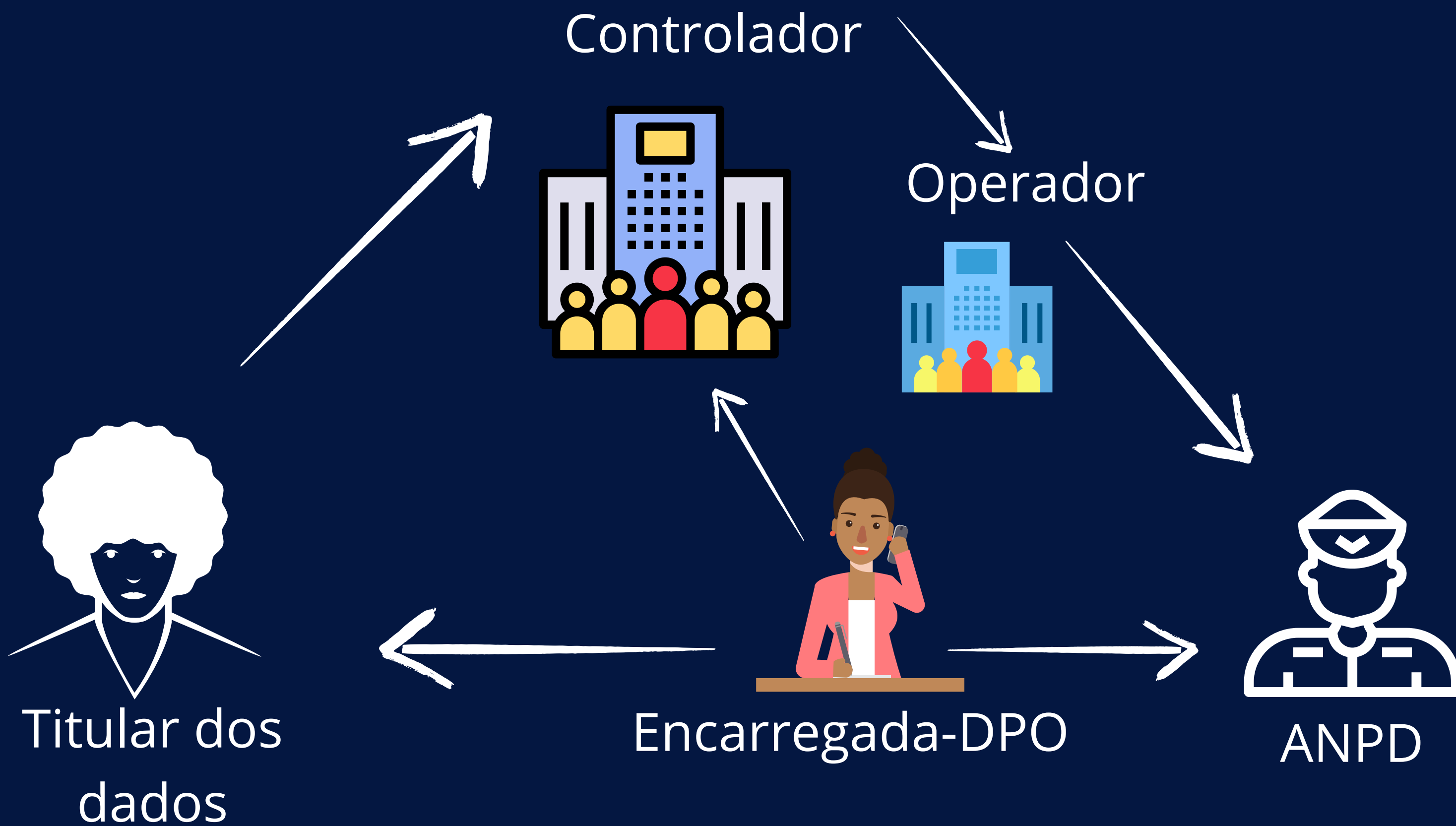
No tratamento de dados pessoais:

- Realizado por pessoa natural para fins exclusivamente particulares e não econômicos;
- Realizado para fins exclusivamente jornalísticos, artísticos ou acadêmicos; e
- Realizado para fins exclusivos de segurança pública, defesa nacional, segurança do Estado ou atividades de investigação e repressão de infrações penais.

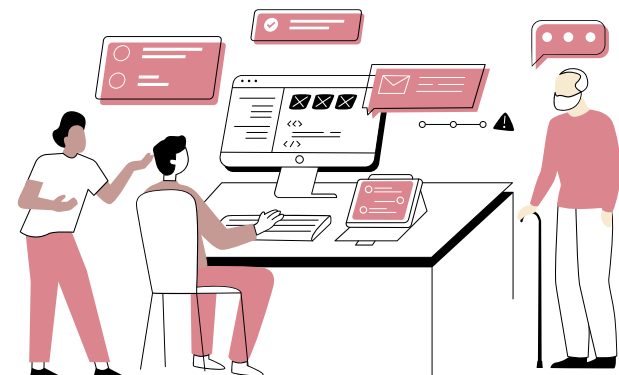
Objetivo da Lei

- Proteger os dados pessoais estipular as regras para as atividades de tratamento de dados pessoais,
- Garantir segurança jurídica nas operações

A LGPD não proíbe o tratamento de dados pessoais



Tratamento



Tratamento



Princípios da LGPD

Finalidade

Deixar Claro para qual finalidade está coletando o dado.

Necessidade

Limitação do tratamento ao mínimo necessário para a realização de suas finalidades

Adequação

Compatibilidade do tratamento com as finalidades informadas ao titular, de acordo com o contexto do tratamento

Princípios da LGPD

BOA FÊ

Prevenção
Adotar
medidas para
evitar danos.

Não
Discriminação
Evitar uso
discriminatório
dos dados.

Responsabilização
Demonstrar
conformidade
com as normas.

Princípios da LGPD

**Livre
Acesso:**
Garantir
acesso fácil
e gratuito
aos dados

**Qualidade
dos Dados:**
Manter
dados
precisos e
atualizados

Segurança:
Proteger os
dados
contra
acessos não
autorizados.

Hipóteses legais para Tratamento de Dados Pessoais

Obrigação Legal ou Regulamentar

Para cumprir uma obrigação legal ou regulatória.

Administração Pública

Para a execução de políticas públicas previstas em leis ou regulamentos.

Estudos por Órgão de Pesquisa

Para a realização de estudos, garantida a anonimização dos dados sempre que possível.

Execução de Contrato:

Necessário para a execução de um contrato ou procedimentos preliminares relacionados a um contrato do qual o titular é parte.

Exercício Regular de Direitos

Para o exercício regular de direitos em processos judiciais, administrativos ou arbitrais.

Hipóteses legais para tratamento de dados pessoais

Proteção da
Vida ou
Incolumidade
Física

Para
proteger a
vida ou a
integridade
física do
titular ou de
terceiros.

Tutela da
Saúde

Exclusivamente
em
procedimentos
realizados por
profissionais de
saúde, serviços
de saúde ou
autoridade
sanitária.

Proteção do
Crédito

Para a proteção
do crédito,
conforme a
legislação
pertinente.

Consentimento

Quando o titular
dos dados dá
permissão
explícita para o
tratamento.

Interesses
Legítimos:

Quando
necessário para
atender aos
interesses
legítimos do
controlador ou
de terceiros,
exceto se
prevalecerem
direitos e
liberdades
fundamentais do
titular.



Direitos dos Tituláres de Dados

- Direito de confirmar se seus dados estão sendo tratados.
- Direito de acessar os dados pessoais que estão sendo tratados.
- Direito de solicitar a correção de dados incorretos.
- Direito de solicitar a anonimização, bloqueio ou eliminação de dados que não estejam sendo tratados conforme a lei.
- Portabilidade - Direito de transferir os dados a outro fornecedor de serviço ou produto.
- Direito de solicitar a eliminação dos dados tratados com base no consentimento, exceto nas hipóteses previstas na lei.
- Direito de ser informado sobre as entidades públicas e privadas com as quais o controlador realizou uso compartilhado de dados.
- . Direito de ser informado sobre a possibilidade de não fornecer consentimento e sobre as consequências dessa negativa.
- . Direito de revogar o consentimento a qualquer momento.



Pequenas Empresas- ANPD Resolução 2 2022

Microempreendedor Individual (MEI) Empresário individual com atividade formalizada Até R\$ 81 mil/ano Máx. 1 funcionário Vendedor de bolos, cabeleireira autônoma

● Microempresa (ME) Empresa de pequeno porte já constituída Até R\$ 360 mil/ano Pequeno comércio, gráfica, consultoria local

Empresa de Pequeno Porte (EPP) Empresa de maior porte, mas ainda considerada de pequeno porte R\$ 360 mil a R\$ 4,8 milhões/ano Loja de materiais de construção, restaurante



Pequenas Empresas- ANPD Resolução 2 2022

Startup Empresa inovadora com atuação recente Até 10 anos de CNPJ
Até R\$ 16 milhões/ano
Foco em inovação Aplicativo de entregas, fintech, edtech

○ Pessoa Natural / Entidade sem fins lucrativos Realiza tratamento de dados no exercício de atividade econômica principal Deve atuar de forma estruturada, mesmo que sem CNPJ
Associação de moradores, professor online

Pequenas Empresas- ANPD Resolução 2 2022

CUIDADOS

Tratar dados com finalidade legítima e informada:

- Informar claramente o motivo da coleta dos dados.

Ex: Coletar e-mails para envio de promoções? Precisa informar e obter consentimento, se não houver outra base legal.

Verificar se está fazendo tratamento de alto risco - A empresa perde o tratamento diferenciado se:

- Tratar grande volume de dados (em escala massiva). Usar dados para decisões automatizadas que afetem o titular (ex: crédito negado por algoritmo). Tratar dados sensíveis (como saúde, religião, biometria), sem medidas de segurança adequadas.

✎ Dica: Se faz isso, precisa agir como uma empresa de grande porte em termos de proteção.

Pequenas Empresas- ANPD Resolução 2 2022

CUIDADOS

Garantir a segurança da informação

- Mesmo com estrutura simplificada, a empresa deve adotar medidas técnicas e administrativas para proteger os dados

Divulgação clara da política de privacidade:

- Pode ser simples mas tem que existir - Deve informar: Quais dados são tratados; Para qual finalidade; Como o titular pode exercer seus direitos.

Pequenas Empresas- ANPD Resolução 2 2022

CUIDADOS

Nomear um Encarregado (DPO)?

- A Resolução permite não nomear formalmente, mas:
 - A empresa deve oferecer um canal de comunicação com o titular de dados (e.g., e-mail ou WhatsApp para dúvidas sobre dados pessoais)

Registro de Tratamento - Pode ser simplificado, mas precisa existir.

- A Resolução permite não nomear formalmente, mas:
 - A empresa deve oferecer um canal de comunicação com o titular de dados (e.g., e-mail ou WhatsApp para dúvidas sobre dados pessoais)

Pequenas Empresas- ANPD Resolução 2 2022

CUIDADOS

Atendimento aos direitos dos titulares:

- A Resolução permite prazo diferenciado, mas não o descumprimento.

Avaliação de riscos

- Para tratamento de Alto Risco

ATENÇÃO!

Se a empresa realiza tratamento que possa causar danos significativos aos titulares, ela não poderá usar as flexibilizações. Exemplos:

- Clínica estética ou médica que trata dados de saúde (dados sensíveis)
- Plataforma de monitoramento por geolocalização contínua
- Empresa que avalia crédito ou faz scoring automatizado
- Escola com sistema de vigilância e reconhecimento facial
- Startup que oferece análise de perfil psicológico ou biométrico
- B



Exceções e tratamentos diferenciados em relação à aplicação da Lei Geral de Proteção de Dados (LGPD).



- **Relatório de Impacto à Proteção de Dados:** Micro e pequenas empresas estão isentas de elaborar este relatório complexo.
- **Encarregado de Dados:** Não é obrigatório nomear um DPO, mas devem fornecer um canal de comunicação para titulares de dados.
- **Registro de Operações de Tratamento:** Podem estar dispensadas de manter registros detalhados, desde que os dados não apresentem alto risco.

Fiscalização ANPD

Primeira Multa

Publicações e Relatórios ANPD

Informações sobre Fiscalização ANPD



Punições para o Uso Indevido de Dados Pessoais



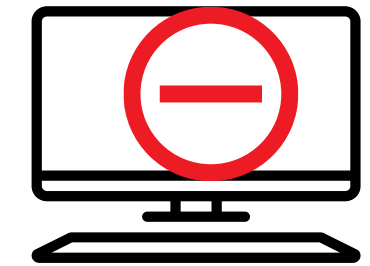
Advertência com indicação de prazo para adoção de medidas corretivas



Multa simples, de até 2% do faturamento limitado a R\$ 50.000.000,00 ou multa diária observando o limite acima



Publicização a infração após devidamente apurada e confirmada a sua ocorrência



Bloqueio dos dados pessoais a que se refere a infração até a sua regularização



Eliminação dos dados pessoais a que se refere a infração



Suspensão do exercício da atividade de tratamento dos dados pessoais a que se refere a infração pelo período máximo de 6 meses prorrogável por igual período



Proibição parcial ou total do exercício de atividades relacionadas a tratamento de dados.

Exemplo:



- **A anonimização é uma ferramenta poderosa:**
 - Mas deve ser usada corretamente.
 - Os dados devem ser tratados de forma legítima desde o início
 - Deve ser informada e compatível com a finalidade



Anonimização



Anonimização: fundamentos jurídicos

Princípios Básicos

- **Legitimidade:**

- Antes de anonimizar, o tratamento dos dados deve estar de acordo com a lei.

- **Finalidade:**

- O tratamento dos dados deve ter um propósito claro, legítimo e informado ao titular dos dados.

- **Adequação:**

- O tratamento dos dados deve ser compatível com a finalidade informada.

- **Necessidade:**

- Só coletar e tratar dados necessários para atingir a finalidade informada.



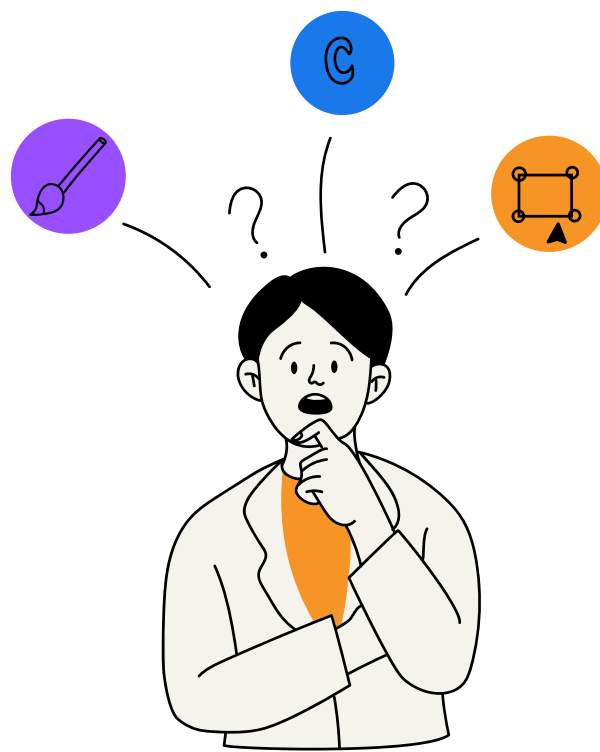
Anonimização e Pseudonimização

- **Anonimização:**

- Dados pessoais perdem a possibilidade de associação a um indivíduo.
- Os dados **não podem ser revertidos** para identificar um indivíduo.

- **Pseudonimização:**

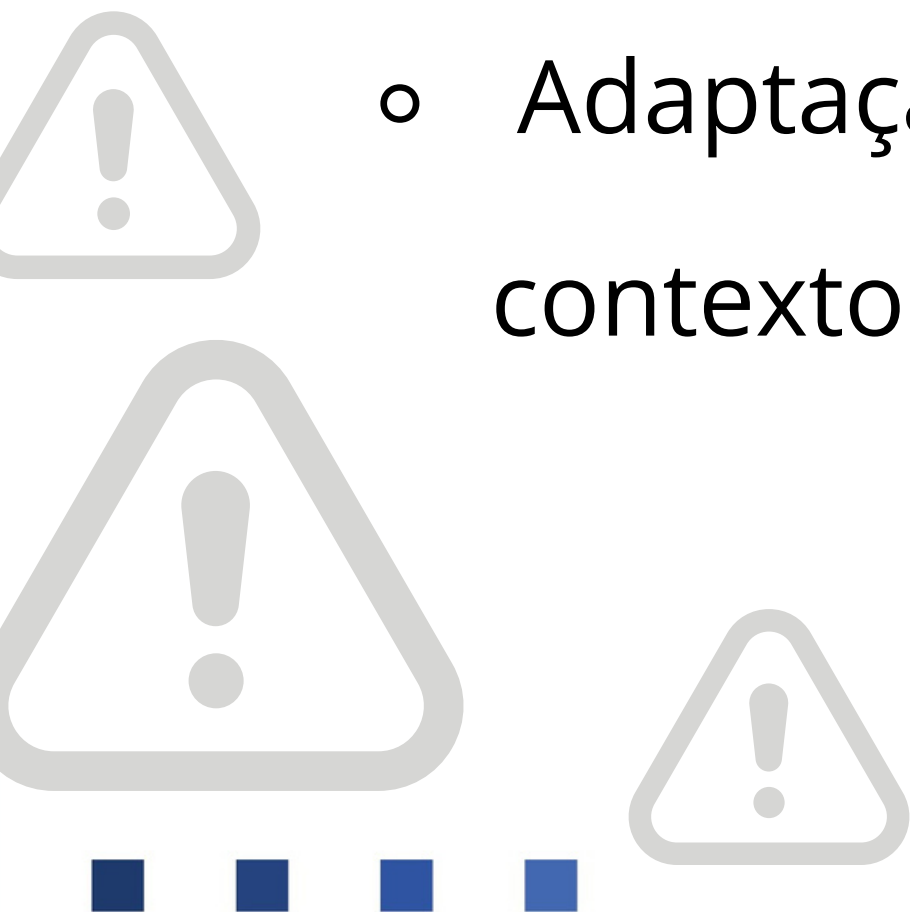
- Dados pessoais perdem a possibilidade de associação a um indivíduo.
- Uso de informação adicional mantida separadamente pelo controlador em ambiente controlado e seguro permite **acesso aos dados originais**.



Anonimização baseada em risco

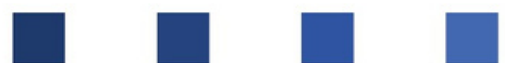
- **Gerenciamento de riscos:**

- Inclui a mensuração do risco de re-identificação após a aplicação de técnicas de anonimização
- Adaptação das soluções com base nas especificidades do contexto



Documentar processos

- Documentar etapas de anonimização
 - Garantir que os dados mantenham suas propriedades estatísticas originais.
 - Ex. ao adicionar ruído, as características dos dados podem mudar, afetando sua qualidade.



Estratégias de Anonimização

- **Generalização**

- Reduz a granularidade dos dados.
- Ex. O logradouro e número viram apenas “bairro”.

ID_Cliente	Nome	Endereço	Número de carros
1	Rodalinda	Rua das Camélias, 133, Aclimação	50
2	100Rodas	Rua das Couves 144, Pinheiros	5
3	MuitaRoda	Rua Mercado, 23, Pinheiros	2
4	Rodaroda	Rua das Amélicas, 45, Aclimação	20



Técnica	ID	Nome	Endereço	Número de carros
Generalização	*	*	Aclimação	50
	*	*	Pinheiros	5
	*	*	Pinheiros	2
	*	*	Aclimação	20

Generalização

```
[7] 1 df_gen = df.copy()
    2 df_gen['Bairro'] = df_gen['Endereço'].apply(lambda x: x.split(',')[ -1].strip())
    3 df_gen = df_gen[['Bairro', 'Carros']]
    4 df_gen
```



	Bairro	Carros
0	Aclimação	50
1	Pinheiros	5
2	Pinheiros	2
3	Aclimação	20



Estratégias de Anonimização

- **Supressão**

- Remove itens identificáveis.
- Ex. Somente número de carros é exibido”.

ID_Cliente	Nome	Endereço	Número de carros
1	Rodalinda	Rua das Camélias, 133, Aclimação	50
2	100Rodas	Rua das Couves 144, Pinheiros	5
3	MuitaRoda	Rua Mercado, 23, Pinheiros	2
4	Rodaroda	Rua das Amélicas, 45, Aclimação	20



Técnica	ID	Nome	Endereço	Número de carros
Supressão	*	*	*	50
	*	*	*	5
	*	*	*	2
	*	*	*	20

Supressão

ChatGPT



```
1 df_sup = df.copy()
2 df_sup = df_sup.drop(columns=['Nome', 'Endereço'])
3 df_sup
```



	ID	Carros
0	1	50
1	2	5
2	3	2
3	4	20



Estratégias de Anonimização

- **Perturbação**

- Adiciona ruído aos dados.
- Ex. O número de carros ganha variação em uma margem adequada.

ID_Cliente	Nome	Endereço	Número de carros
1	Rodalinda	Rua das Camélias, 133, Aclimação	50
2	100Rodas	Rua das Couves 144, Pinheiros	5
3	MuitaRoda	Rua Mercado, 23, Pinheiros	2
4	Rodaroda	Rua das Amélicas, 45, Aclimação	20



Técnica	ID	Nome	Endereço	Número de carros
Perturbação	1	Rodalinda	Rua das Camélicas, 133, Aclimação	48
	2	100Rodas	Rua das Couves 144, Pinheiros	6
	3	MuitaRoda	Rua Mercado, 23, Pinheiros	1
	4	Rodaroda	Rua das Amélicas, 45, Aclimação	22

Perturbação

ChatGPT

```
1 np.random.seed(42)
2 df_pert = df.copy()
3 noise = np.random.randint(-3, 4, size=len(df))
4 df_pert['Carros'] = df_pert['Carros'] + noise
5 df_pert
```



	ID	Nome	Endereço	Carros
0	1	Rodalinda	Rua das Camélias, 133, Aclimação	53
1	2	100Rodas	Rua das Couves 144, Pinheiros	5
2	3	MuitaRoda	Rua Mercado, 23, Pinheiros	3
3	4	Rodaroda	Rua das Amélicas, 45, Aclimação	23



Estratégias de Anonimização

- **Agregação**
 - Agrupa registros em categorias.
 - Ex. Dados condensados por bairro.

ID_Cliente	Nome	Endereço	Número de carros
1	Rodalinda	Rua das Camélias, 133, Aclimação	50
2	100Rodas	Rua das Couves 144, Pinheiros	5
3	MuitaRoda	Rua Mercado, 23, Pinheiros	2
4	Rodaroda	Rua das Amélicas, 45, Aclimação	20

Técnica	ID	Nome	Endereço	Número de carros
Agregação		2 clientes	Aclimação	70
		2 clientes	Pinheiros	7

Agregação

ChatGPT



```
1 df_agg = df.copy()
2 df_agg['Bairro'] = df_agg['Endereço'].apply(lambda x: x.split(',')[-1].strip())
3 agg = df_agg.groupby('Bairro').agg({'ID': 'count', 'Carros': 'sum'}).rename(columns={'ID': 'Clientes'})
4 agg
```



Clientes Carros



Bairro

Aclimação

2

70



Pinheiros

2

7



Estratégias de Anonimização

- **Substituição**

- Substitui dados reais por fictícios.
- Ex. Nomes e endereços inventados.

ID_Cliente	Nome	Endereço	Número de carros
1	Rodalinda	Rua das Camélias, 133, Aclimação	50
2	100Rodas	Rua das Couves 144, Pinheiros	5
3	MuitaRoda	Rua Mercado, 23, Pinheiros	2
4	Rodaroda	Rua das Amélicas, 45, Aclimação	20



Técnica	ID	Nome	Endereço	Número de carros
Substituição	1	Joana	Rua A, 10, Bairro Alfa	50
	2	Carla	Rua B, 20, Bairro Beta	5
	3	Marcos	Rua C, 30, Bairro Beta	2
	4	Pedro	Rua D, 40, Bairro Alfa	20

Substituição



```
1 df_sub = df.copy()
2 sub_names = ['Joana', 'Carla', 'Marcos', 'Pedro']
3 sub_end = ['Rua A, 10, Alfa', 'Rua B, 20, Beta', 'Rua C, 30, Beta', 'Rua D, 40, Alfa']
4 df_sub['Nome'] = sub_names
5 df_sub['Endereço'] = sub_end
6 df_sub
```



	ID	Nome	Endereço	Carros
0	1	Joana	Rua A, 10, Alfa	50
1	2	Carla	Rua B, 20, Beta	5
2	3	Marcos	Rua C, 30, Beta	2
3	4	Pedro	Rua D, 40, Alfa	20



```

1 import pandas as pd
2 from faker import Faker
3
4 def anonimizar_nomes_df(df, seed=1234):
5     """
6     Retorna uma cópia do DataFrame com a coluna 'Nome' anonimizadas.
7     Nomes repetidos recebem o mesmo nome fictício.
8     """
9     fake = Faker('pt_BR')
10    Faker.seed(seed)
11
12    df_anon = df.copy()
13    mapping = {} # mapeamento: nome_real -> nome_fictício
14
15    for idx, nome_real in df_anon['Nome'].items():
16        if nome_real not in mapping:
17            mapping[nome_real] = fake.first_name()
18            df_anon.at[idx, 'Nome'] = mapping[nome_real]
19            print(f"Anonimizando '{nome_real}' -> '{mapping[nome_real]}'")
20
21    return df_anon
22
23 df_anon = anonimizar_nomes_df(df)
24 print("\nDataFrame Anonimizado:")
25 df_anon

```

```

Anonimizando 'Rodalinda' -> 'Lunna'
Anonimizando '100Rodas' -> 'Igor'
Anonimizando 'MuitaRoda' -> 'Anthony Gabriel'
Anonimizando 'Rodaroda' -> 'Fernando'

```

DataFrame Anonimizado:

	ID	Nome	Endereço	Carros
0	1	Lunna	Rua das Camélias, 133, Aclimação	50
1	2	Igor	Rua das Couves 144, Pinheiros	5
2	3	Anthony Gabriel	Rua Mercado, 23, Pinheiros	2
3	4	Fernando	Rua das Amélicas, 45, Aclimação	20

Estratégias de Anonimização

- **Mascaramento**

- Mascaramento parcial de dados.
- Ex. Apenas parte dos nomes e endereços visível.

ID_Cliente	Nome	Endereço	Número de carros
1	Rodalinda	Rua das Camélias, 133, Aclimação	50
2	100Rodas	Rua das Couves 144, Pinheiros	5
3	MuitaRoda	Rua Mercado, 23, Pinheiros	2
4	Rodaroda	Rua das Amélicas, 45, Aclimação	20



Técnica	ID	Nome	Endereço	Número de carros
Mascaramento	1	R*****a	Rua das C*****s, **3, Aclimação	50
	2	1**Rodas	Rua das C*****s 1**, Pinheiros	5
	3	M****Roda	Rua M*****do, **, Pinheiros	2
	4	R*****a	Rua das A*****s, **, Aclimação	20

```

1 import pandas as pd
2
3 def mascarar(informacao: str) -> str:
4     """Retorna uma string com asteriscos com o mesmo comprimento da entrada."""
5     return '*' * len(informacao)
6
7 def anonimizar_nomes_df(df: pd.DataFrame) -> pd.DataFrame:
8     """
9     Recebe um DataFrame com colunas 'ID_Cliente' e 'Nome',
10    retorna cópia com nomes mascarados por asteriscos.
11    """
12    df_anon = df.copy()
13
14    # Aplicar máscara em cada valor da coluna 'Nome'
15    df_anon['Nome'] = df_anon['Nome'].apply(mascarar)
16
17    print("Processo de anonimização:")
18    for orig, novo in zip(df['Nome'], df_anon['Nome']):
19        print(f" {orig} -> {novo}")
20
21    return df_anon
22
23 df_anon = anonimizar_nomes_df(df)
24 print("\nDataFrame original:")
25 print(df)
26 print("\nDataFrame anonimizado:")
27 df_anon

```

Processo de anonimização:

Rodalinda -> *****

100Rodas -> *****

MuitaRoda -> *****

Rodaroda -> *****

DataFrame original:

	ID	Nome	Endereço	Carros
0	1	Rodalinda	Rua das Camélias, 133, Aclimação	50
1	2	100Rodas	Rua das Couves 144, Pinheiros	5
2	3	MuitaRoda	Rua Mercado, 23, Pinheiros	2
3	4	Rodaroda	Rua das Amélicas, 45, Aclimação	20

DataFrame anonimizado:

	ID	Nome	Endereço	Carros
0	1	*****	Rua das Camélicas, 133, Aclimação	50
1	2	*****	Rua das Couves 144, Pinheiros	5
2	3	*****	Rua Mercado, 23, Pinheiros	2
3	4	*****	Rua das Amélicas, 45, Aclimação	20

Estratégias de Anonimização

- **Pseudonimização**
 - Substitui identificadores por pseudo-IDs.
 - Ex. Nome, ID e endereço viram pseudônimos, mas rastreáveis internamente.

ID_Cliente	Nome	Endereço	Número de carros
1	Rodalinda	Rua das Camélias, 133, Aclimação	50
2	100Rodas	Rua das Couves 144, Pinheiros	5
3	MuitaRoda	Rua Mercado, 23, Pinheiros	2
4	Rodaroda	Rua das Amélicas, 45, Aclimação	20



Técnica	ID	Nome	Endereço	Número de carros
Pseudonimização	A102	P001	Aclimação	50
	B304	P002	Pinheiros	5
	C507	P003	Pinheiros	2
	D210	P004	Aclimação	20

Pseudonimização

```
[15] 1 df_pseudo = df.copy()
      2 df_pseudo['Pseudo_ID'] = ['A102', 'B304', 'C507', 'D210']
      3 df_pseudo['Nome'] = ['P001', 'P002', 'P003', 'P004']
      4 df_pseudo['Bairro'] = df_pseudo['Endereço'].apply(lambda x: x.split(',')[0].strip())
      5 df_pseudo = df_pseudo[['Pseudo_ID', 'Nome', 'Bairro', 'Carros']]
      6 df_pseudo
```



	Pseudo_ID	Nome	Bairro	Carros
0	A102	P001	Aclimação	50
1	B304	P002	Pinheiros	5
2	C507	P003	Pinheiros	2
3	D210	P004	Aclimação	20



ChatGPT

```
1 from faker import Faker
2
3 #Gerar letras e numeros aleatorios
4 fake = Faker()
5 Faker.seed(0)
6 for _ in range(5):
7     fake_code = fake.bothify(text='Numero do produto: ????-#####') #gera uma letra para cada ? e um número para cada ###
8     print(fake_code)
9
10 for _ in range(5):
11     fake_id = fake.lexify(text='Identificador randomico: ??????????')
12     print(fake_id)
13
14 Faker.seed(0)
15 for _ in range(5):
16     fake_num = fake.numerify(text='Produto id%-%%##K')
17     print(fake_num)
```

```
Numero do produto: FzYt-66048
Numero do produto: isiW-75938
Numero do produto: jtGU-19489
Numero do produto: wBuN-15781
Numero do produto: qdZJ-38778
Identificador randomico: afUzTYQ0aN
Identificador randomico: FvpUuTemKo
Identificador randomico: pZjZICffuG
Identificador randomico: FgtJsThJvI
Identificador randomico: nZMJLsCfMZ
Produto id1-5966K
Produto id5-8676K
Produto id9-3593K
Produto id5-9321K
Produto id2-6841K
```

ChatGPT

```
1 #Criar as referencias de nomes para ele escolher, eles chamam de provider
2 from faker import Faker
3 from faker.providers import DynamicProvider
4
5 nome_sugerido = DynamicProvider(
6     provider_name="sugestao_nome",
7     elements=['roda', 'carro', 'caminhao', 'trem', 'cambio', 'onibus'],
8 )
9
10 fake = Faker()
11
12 # then add new provider to faker instance
13 fake.add_provider(nome_sugerido)
14
15 # now you can use:
16 fake_nome = fake.sugestao_nome()
17 print(fake_nome)
```

roda

Estratégias de Anonimização

- **K-anonimato**

- Garante que cada combinação de dados se repita em pelo menos k registros.
- Ex. Para $k=2$, linhas agrupadas em 2 bairros, e no. de carros em faixas, cada combinação (bairro + faixa) aparece pelo menos 2 vezes.

ID_Cliente	Nome	Endereço	Número de carros
1	Rodalinda	Rua das Camélias, 133, Aclimação	50
2	100Rodas	Rua das Couves 144, Pinheiros	5
3	MuitaRoda	Rua Mercado, 23, Pinheiros	2
4	Rodaroda	Rua das Amélicas, 45, Aclimação	20



Técnica	ID	Nome	Endereço	Número de carros
K-anonimato	*	*	Aclimação	alto
	*	*	Pinheiros	baixo
	*	*	Pinheiros	baixo
	*	*	Aclimação	medio

K-anonimato (k=2)



Cha

```
1 df_k = df.copy()
2 df_k['Bairro'] = df_k['Endereço'].apply(lambda x: x.split(',')[0].strip())
3 # generaliza número de carros em faixas
4 bins = [0,10, 40, 80]
5 labels = ['baixa','média','alta']
6 df_k['Faixa'] = pd.cut(df_k['Carros'], bins=bins, labels=labels)
7 # suprime Nome, Endereço, ID
8 df_k_final = df_k[['Bairro','Faixa']]
9 # verifica repetição das combinações
10 df_k_final.groupby(['Bairro','Faixa']).size()
```

Bairro	Faixa	
Aclimação	baixa	0
	média	1
	alta	1
Pinheiros	baixa	2
	média	0
	alta	0



Estratégias de Anonimização

- Artigo

- link (<https://www.nature.com/articles/s41597-024-04019-z>)

The screenshot shows a web browser displaying a Nature Scientific Data article. The browser's address bar shows the URL [nature.com/articles/s41597-024-04019-z](https://www.nature.com/articles/s41597-024-04019-z). The page header includes navigation links: "Explore content", "About the journal", and "Publish with us", along with "Sign up for alerts" and "RSS feed". The breadcrumb trail is "nature > scientific data > articles > article". The article is labeled "Article | Open access | Published: 26 November 2024". The title is "An Open Source Python Library for Anonymizing Sensitive Data". The authors are "Judith Sáinz-Pardo Díaz" (with an ORCID icon) and "Álvaro López García". The journal information is "Scientific Data 11, Article number: 1289 (2024)" with a "Cite this article" link. On the right, there is a "Download PDF" button and a sidebar with "Sections", "Figures", and "References". The "Sections" sidebar lists "Abstract" and "Introduction".

Estratégias de Anonimização

- **Frameworks e Bibliotecas**

- **Anjana** (<https://github.com/IFCA-Advanced-Computing/anjana>)
 - Biblioteca de modelos com progressão de proteção
 - Simples anonimato de classes (k)
 - Sofisticados métodos estatísticos (t , β , δ), cobrindo diferentes tipos de ataque, preservando utilidade dos dados.
 - A biblioteca ANJANA permite pipelines de anonimização robustos e customizados

Estratégias de Anonimização

- **Frameworks e Bibliotecas**

- **Py canon** (<https://github.com/IFCA-Advanced-Computing/pycanon>)
 - Framework para avaliar se um dataset atende a diversas técnicas de preservação de privacidade via anonimização
 - Permite identificar os parâmetros mínimos (k , α , ℓ , t , β , δ) necessários para satisfazer cada técnica
 - Fornece insights sobre riscos reais de re-identificação

Estratégias de Anonimização

- **Frameworks e Bibliotecas**

- **Anonympy** (<https://pypi.org/project/anonympy/#description>)
 - A biblioteca para anonimizar diversos tipos de dados
 - Agregação, ruído, arredondamento, mascaramento, substituição, perturbação
 - Anonimiza imagens e PDFs
 - Interface amigável e intuitiva

Estratégias de Anonimização

- **Frameworks e Bibliotecas**

- **Data anonymazer** (<https://pypi.org/project/data-anonymizer/>)
 - Ferramenta para anonimizar arquivos CSV por meio da geração automática de um arquivo-chave para garantir anonimização consistente e manutenção de integridade referencial
 - Utiliza a biblioteca Faker para gerar valores substitutos
- **Faker** (<https://pypi.org/project/Faker/0.7.4/>)
 - permite criar / usar dados sintéticos

Estratégias de Anonimização

- Frameworks e Bibliotecas

- **[anonymization-library](https://github.com/Rayxan/anonymization-library)** (<https://github.com/Rayxan/anonymization-library>)
 - Biblioteca para a linguagem de programação python
 - Permite o desenvolvedor criar códigos capazes de anonimizar dados pessoais brasileiros presentes em arquivos pdf, docx, xlsx, assim como strings.

Estratégias de Anonimização

- **Frameworks e Bibliotecas**

- **Presidio** (<https://github.com/microsoft/presidio>)

- Ferramenta para detecção e anonimização de PII em texto, imagens e dados estruturados, que oferece módulos variados:
 - Presidio Analyzer: identifica entidades sensíveis (nomes, CPF, cartões, endereços, etc.) usando regex, NER
 - Presidio Anonymizer: aplica técnicas como mascaramento, substituição, criptografia
 - Uso em figuras, pdf

Estratégias de Anonimização

- **Softwares**

- **ARX** (<https://arx.deidentifier.org/>)
 - Ferramenta para anonimização de dados estruturados
 - Interface gráfica intuitiva
 - API Java robusta
 - Vários modelos de privacidade e risco (como k-anonimato, ℓ -diversidade, t-closeness, δ -disclosure privacy, β -likeness e até ϵ -differential privacy)

Publicações Relevantes

- [Estudo Técnico sobre anonimização de dados na LGPD](#)
- [Guia orientativo sobre segurança da informação para agentes de tratamento de pequeno porte](#)

Obrigado!

- silvana@ipt.br
- adrianab@ipt.br



[linkedin.com/school/iptsp/](https://www.linkedin.com/school/iptsp/)



[instagram.com/ipt_oficial/](https://www.instagram.com/ipt_oficial/)



[youtube.com/@IPTbr/](https://www.youtube.com/@IPTbr/)

www.ipt.br

