

Proteção de Dados Pessoais e Técnicas de Anonimização

Prof. Msc. Francisco de Assis de Souza Rodrigues

UniFAP
CENTRO UNIVERSITÁRIO PARAÍSO

🎓 Objetivos

- Compreender os princípios da proteção de dados pessoais.
- Conhecer a legislação vigente (ex: LGPD).
- Identificar os tipos de dados pessoais e sensíveis.
- Apresentar técnicas de anonimização e pseudonimização.
- Aplicar exemplos práticos de anonimização de dados.

🏷️ O que são dados pessoais?

Dados pessoais são qualquer informação relacionada a uma pessoa natural identificada ou identificável



Exemplos:

- Nome completo
- CPF, RG
- Número de telefone
- Endereço
- E-mail
- Data de nascimento
- Foto, biometria, reconhecimento facial
- Localização via GPS

🏷️ O que são dados pessoais?

Dados sensíveis (subconjunto de dados pessoais que merecem maior proteção):

- Raça ou etnia
- Convicções religiosas
- Opiniões políticas
- Filiação sindical
- Dados genéticos ou biométricos
- Dados de saúde
- Orientação sexual



🏷️ O que são dados pessoais?

Dado pessoal ≠ dado pessoal sensível:

O primeiro é qualquer dado que identifica, o segundo é um dado que, além de identificar, pode gerar **discriminação, estigmatização ou danos mais sérios** se for exposto.

“Vocês conseguem pensar em outros exemplos de dados pessoais?”

“Se vazar seu CPF, qual seria o problema? E se vazarem dados sobre sua religião?”

🔒 Por que proteger dados pessoais?

1. **Privacidade:**
   Garantir que a pessoa tenha controle sobre seus próprios dados.

2. **Prevenção de fraudes:**
   Vazamento de CPF, nome e dados bancários pode gerar:

   - Abertura de contas falsas
   - Contratação de serviços em seu nome
   - Golpes financeiros

🔒 Por que proteger dados pessoais?

3. **Prevenção de discriminação:**
   Exposição de dados sensíveis pode gerar:
   - Discriminação em processos seletivos
   - Exclusão social
   - Riscos à segurança física e psicológica

4. **Cumprimento da Lei:**
   A **LGPD** exige que empresas, órgãos públicos e pessoas que tratam dados pessoais garantam segurança, privacidade e evitem vazamentos.

5. **Reputação:**
   Empresas que vazam dados perdem credibilidade, sofrem processos e podem ter multas altíssimas.

⚠️ Exemplos reais de vazamento de dados e suas consequências

Facebook / Cambridge Analytica (2018)

- Dados de **87 milhões de usuários** foram coletados indevidamente.
- A empresa de marketing político usou informações pessoais para criar **perfis psicológicos** e direcionar anúncios políticos manipulativos, incluindo nas eleições dos EUA e referendo do Brexit.
- Impacto: Multas bilionárias, queda no valor das ações, CEO no Congresso dos EUA.

⚠️ Exemplos reais de vazamento de dados e suas consequências

🔓 Vazamento de dados da Serasa (2021)

- Dados de **220 milhões de brasileiros** vazaram (nome, CPF, renda, score de crédito, endereço, telefone).
- Dados estavam sendo vendidos na dark web.
- Consequências: Milhares de fraudes, golpes via WhatsApp e abertura de contas falsas.

⚠️ Exemplos reais de vazamento de dados e suas consequências

🏥 Vazamento do Ministério da Saúde (2021)

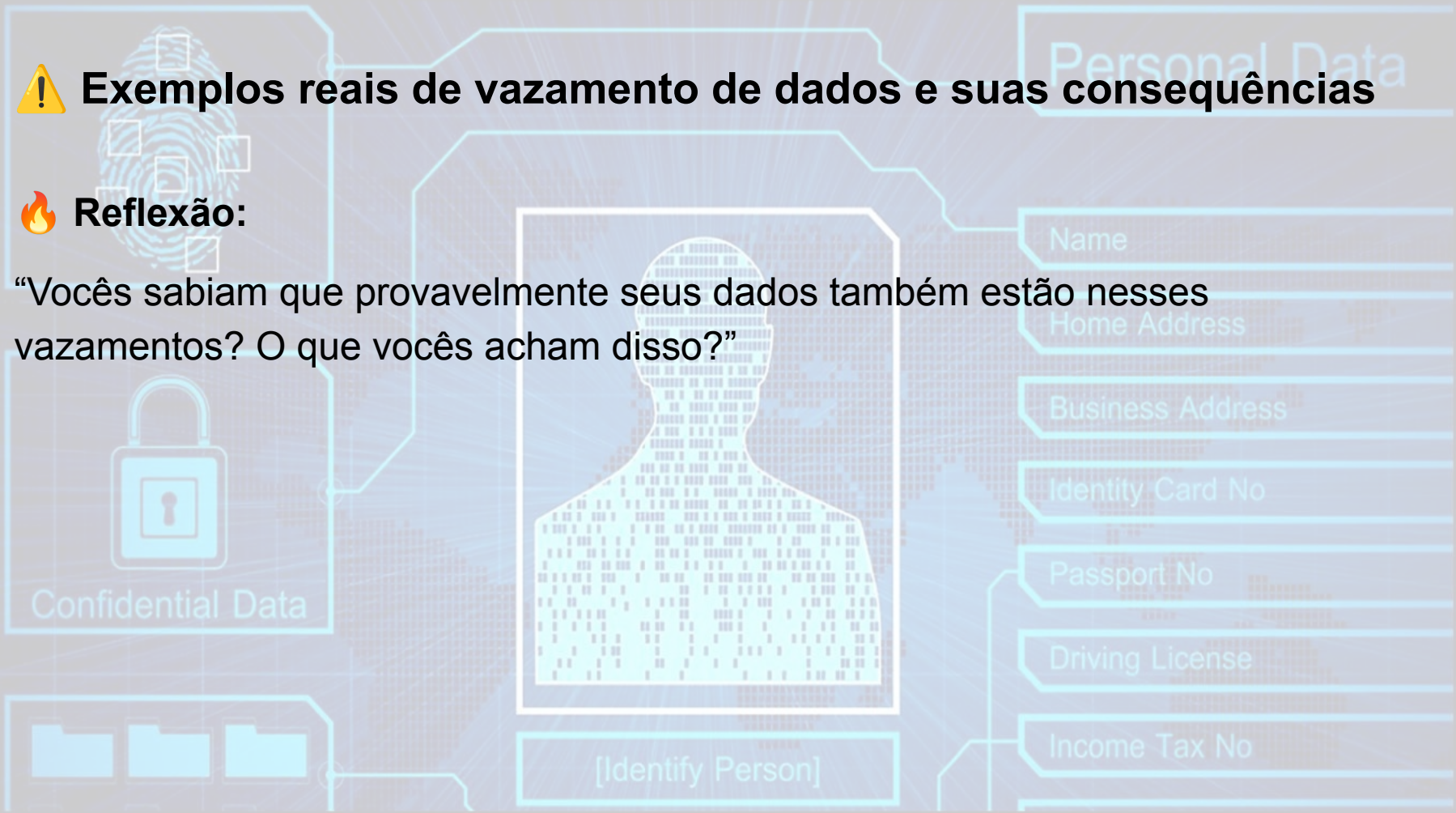
- Dados de milhões de brasileiros que acessaram o SUS foram expostos (nome, CPF, registros de vacinação e exames).

- Consequências: Risco à privacidade médica, possibilidade de chantagens, discriminação e golpes.

⚠️ Exemplos reais de vazamento de dados e suas consequências

🔥 Reflexão:

“Vocês sabiam que provavelmente seus dados também estão nesses vazamentos? O que vocês acham disso?”



📜 2. Legislação e Princípios da Proteção de Dados

🚩 LGPD (Lei Geral de Proteção de Dados) – Lei nº 13.709/2018 – Brasil

✔️ Finalidade da LGPD

- Proteger os **direitos fundamentais de liberdade, privacidade e o livre desenvolvimento da personalidade** da pessoa natural.

- Estabelece regras claras sobre **coleta, armazenamento, tratamento e compartilhamento de dados pessoais**.

- Visa aumentar a **segurança jurídica**, promover a **transparência** e criar um ambiente de **confiança digital** entre empresas e cidadãos.

📜 2. Legislação e Princípios da Proteção de Dados

⚖️ Bases Legais para Tratamento de Dados

A LGPD permite o tratamento de dados apenas se houver uma **base legal adequada**. As principais são:

- **Consentimento** do titular (deve ser claro, informado e inequívoco).
- **Cumprimento de obrigação legal ou regulatória.**
- **Execução de políticas públicas** (órgãos governamentais).
- **Estudos por órgão de pesquisa** (dados anonimizados sempre que possível).
- **Execução de contrato ou procedimentos preliminares.**

📜 2. Legislação e Princípios da Proteção de Dados

- **Estudos por órgão de pesquisa** (dados anonimizados sempre que possível).
- **Execução de contrato ou procedimentos preliminares.**
- **Exercício regular de direitos** em processos judiciais, administrativos ou arbitrais.
- **Proteção da vida ou da incolumidade física** do titular ou de terceiros.
- **Tutela da saúde**, exclusivamente, por profissionais ou entidades de saúde.
- **Interesse legítimo do controlador**, exceto quando prevalecerem direitos e liberdades do titular.
- **Proteção do crédito**, como para análise e concessão de crédito.

📜 2. Legislação e Princípios da Proteção de Dados

👤 Direitos dos Titulares de Dados

O titular de dados possui diversos direitos garantidos, como:

- **Confirmação** da existência do tratamento.
- **Acesso** aos dados.
- **Correção** de dados incompletos, inexatos ou desatualizados.
- **Anonimização, bloqueio ou eliminação** de dados desnecessários, excessivos ou tratados em desconformidade.

📜 2. Legislação e Princípios da Proteção de Dados

- **Portabilidade** dos dados a outro fornecedor (respeitados segredos comerciais).
- **Eliminação dos dados tratados com consentimento.**
- **Informação** sobre compartilhamento com terceiros.
- **Revogação do consentimento.**
- **Oposição** ao tratamento em caso de descumprimento da lei.

📜 2. Legislação e Princípios da Proteção de Dados

👥 Papéis na LGPD

- **Controlador:** Pessoa ou empresa que **toma as decisões** sobre o tratamento dos dados.

- **Operador:** Quem **realiza o tratamento dos dados** em nome do controlador.

- **Encarregado (DPO – Data Protection Officer):**
  - Canal de comunicação entre controlador, titulares e ANPD (Agência Nacional de Proteção de Dados).
  - Recomenda boas práticas, orienta e verifica o cumprimento da lei.

Exemplo prático:
Uma universidade:

📜 2. Legislação e Princípios da Proteção de Dados

- Controlador → A própria universidade.
- Operador → Empresa terceirizada que processa os pagamentos.
- DPO → Responsável designado para comunicação sobre dados.

📜 2. Legislação e Princípios da Proteção de Dados

🌍 GDPR – Contextualização Internacional

- Lei europeia de proteção de dados, em vigor desde 2018.
- **Inspirou diretamente a LGPD.**
- Aplica-se globalmente se uma empresa tratar dados de cidadãos da União Europeia, mesmo sem ter sede lá.
- GDPR é mais rígido em alguns pontos, como exigência de **Privacy by Design e Privacy by Default.**

📜 2. Legislação e Princípios da Proteção de Dados

🔍 Dados Pessoais vs Dados Sensíveis

- **Dados Pessoais:** Qualquer dado que **identifique ou possa identificar** uma pessoa.
  Exemplos: nome, CPF, e-mail, endereço, número de telefone, foto, IP.
- **Dados Sensíveis:** Dado pessoal que revela informações íntimas e que podem gerar discriminação.
  Exemplos: origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato, dados de saúde, dados biométricos, dados genéticos, vida sexual.
- **Dados anonimizados:** Deixam de ser considerados dados pessoais, desde que não possam ser revertidos.

🔥 3. Riscos no Tratamento de Dados

🚩 1. Identificação Indireta (Reidentificação)

- Mesmo após remover dados diretamente identificáveis (nome, CPF, e-mail), **um indivíduo pode ser reidentificado** combinando outros dados aparentemente inofensivos.

- 🔍 **Exemplo clássico:**
  Um estudo mostrou que **87% da população dos EUA pode ser identificada apenas com:**
  - Data de nascimento
  - CEP
  - Gênero

🔥 3. Riscos no Tratamento de Dados

- 🔗 Quando datasets são cruzados com bases externas (dados públicos, redes sociais, cadastros), há alto risco de descobrir quem é quem.
- ⚠️ **Consequência:**
  Acreditar que dados são "anônimos" sem testes rigorosos leva a exposições acidentais e violações legais.

🔥 3. Riscos no Tratamento de Dados

🚩 2. Compartilhamento Indevido

- **Risco:** Dados são compartilhados com terceiros sem o consentimento do titular ou sem garantias adequadas de segurança.

- 🚫 Pode acontecer de forma:
  - Intencional (venda ilegal de dados, espionagem corporativa)
  - Acidental (e-mails enviados para destinatários errados, uploads públicos inadvertidos)

🔥 3. Riscos no Tratamento de Dados

- 💣 **Exemplos reais:**
  - Vazamento de dados do Serasa
  - Empresas que vendem dados de localização de usuários de aplicativos

- 🔐 **Implicação:** Além de sanções legais (como multas da LGPD), há perda de reputação e confiança por parte dos clientes.

🔥 3. Riscos no Tratamento de Dados

🚩 3. Armazenamento Inseguro

- Dados sensíveis armazenados sem criptografia, em servidores desprotegidos ou acessíveis publicamente.

- 🔓 **Vulnerabilidades comuns:**
  - Bancos de dados sem autenticação
  - Backups expostos em buckets na nuvem (ex.: AWS S3 mal configurado)
  - Uso de senhas fracas ou padrões (ex.: "admin", "123456")

🔥 3. Riscos no Tratamento de Dados

🚩 3. Armazenamento Inseguro

🏴‍☠️ Hackers frequentemente utilizam ferramentas automáticas para escanear a internet em busca desses sistemas expostos.

🔥 Casos reais:

- Vazamento de senhas de redes sociais
- Dados de milhões de pacientes de sistemas de saúde expostos devido a configuração incorreta

🔥 3. Riscos no Tratamento de Dados

🚩 3. Armazenamento Inseguro

🏗️ Boas práticas recomendadas:

- Criptografia em repouso e em trânsito
- Políticas rígidas de acesso (privilégio mínimo)
  Backups protegidos
- Monitoramento constante de acessos

4. Técnicas de Anonimização

🔍 O que é Anonimização?

É o processo de remover ou alterar informações pessoais de um conjunto de dados de tal forma que os indivíduos não possam mais ser identificados, direta ou indiretamente.

⚠️ **Importante:** Após a anonimização, os dados deixam de ser considerados dados pessoais segundo a LGPD (se a reidentificação for efetivamente impossível).

4. Técnicas de Anonimização

🚀 Principais Técnicas

1. 🗑️ Supressão

- **Definição:** Remoção de informações diretamente identificáveis.
- **Exemplos:**
  - Remover campos como CPF, nome, e-mail, telefone.
- **Uso:** Quando a informação não é necessária para a análise.

Nome	Idade	Cidade	CPF
João Silva	25	São Paulo	*Removido*

4. Técnicas de Anonimização

2. 🔎 Generalização

- **Definição:** Reduzir a precisão dos dados, tornando-os mais genéricos.

- **Exemplos:**
  - Data de nascimento → Faixa etária (1995-06-15 → 25-30 anos)
  - Localização → Nível de cidade, estado ou país.

- **Aplicação:** Mantém utilidade dos dados para análise estatística, mas dificulta identificação.

Nome	Idade	Cidade
Removido	25-30	Sudeste

4. Técnicas de Anonimização

3. 🎲 Perturbação

- **Definição:** Adiciona ruído (aleatoriedade controlada) aos dados para mascarar valores exatos.

- **Exemplos:**
  Alterar salários adicionando ou subtraindo valores pequenos e aleatórios.

- **Ferramentas:** Técnicas de **Privacidade Diferencial** se enquadram aqui.

Salário real	Salário com ruído
5000	4987

4. Técnicas de Anonimização

4. 👥 Pseudonimização *(Não é anonimização completa!)*

- **Definição:** Substituição de identificadores diretos por pseudônimos.
  → Dados podem ser revertidos mediante chave segura.

- **Exemplos:**
  - Substituir CPF por código interno (A12345)

- **Atenção:** Sob LGPD, ainda é considerado dado pessoal.

Nome	ID Pseudo
João Silva	A12345

4. Técnicas de Anonimização

5. 🔑 Tokenização

- **Definição:** Substituir dados sensíveis por tokens não reversíveis, vinculados a registros em um banco separado e seguro.

- **Uso:** Amplamente utilizado em pagamentos e segurança financeira.

Cartão de Crédito	Token
1234-5678-9012-3456	e8f1acb09x...

4. Técnicas de Anonimização

6. 🔐 K-Anonimato

- **Definição:** Um registro é indistinguível de pelo menos outros **k-1** registros no dataset em relação aos atributos quase-identificadores.

- **Exemplo:** Se k=3, para qualquer combinação de (idade, cidade, gênero) devem existir pelo menos 3 registros iguais.

Vantagem: Protege contra reidentificação direta usando cruzamento de dados.

Idade	Cidade	Sexo
25-30	São Paulo	M
25-30	São Paulo	M
25-30	São Paulo	M

4. Técnicas de Anonimização

7. 🎭 L-diversidade e T-closeness

- **L-diversidade:** Assegura que dentro de cada grupo (gerado pelo k-anonimato), exista uma diversidade suficiente de valores sensíveis.

- **T-closeness:** Garante que a distribuição de um atributo sensível dentro de um grupo é próxima da distribuição total da base de dados.

Por que importa? K-anonimato protege contra reidentificação, mas não contra inferência de atributos sensíveis.

Grupo	Doença
A	Gripe, Câncer, Asma

4. Técnicas de Anonimização

📌 Considerações

- **Anonimização não é perfeita.** Existe sempre um risco de reidentificação, especialmente se combinada com dados externos.

- A escolha da técnica depende do **contexto**, **risco**, **regulamentação** e **utilidade desejada dos dados**.

5. Aplicações e Desafios

✅ Vantagens da Anonimização

- **Cumprimento da legislação:** Permite que organizações tratem dados sem violar leis como a LGPD, desde que os dados sejam efetivamente anonimizados.

- **Minimização de riscos:** Reduz os impactos em caso de vazamento, já que dados anonimizados não podem, teoricamente, ser ligados a indivíduos.

- **Facilita o compartilhamento de dados:** Dados anonimizados podem ser usados em pesquisas, desenvolvimento de IA, estatísticas e análises, sem infringir a privacidade dos titulares.

5. Aplicações e Desafios

✅ Vantagens da Anonimização

- **Fomenta inovação:** Organizações podem extrair valor dos dados sem expor informações sensíveis, permitindo desenvolvimento de novos produtos, serviços e modelos de negócio.

- **Proteção da reputação:** A adoção de práticas de anonimização demonstra responsabilidade com a privacidade, fortalecendo a imagem da organização perante clientes e parceiros.

5. Aplicações e Desafios

⚠️ **Limitações: Risco de Reidentificação com Dados Externos**

- **Cross-linking (Junção de bases):** Mesmo dados anonimizados podem ser cruzados com outras bases públicas ou privadas, permitindo reidentificação.
  → Exemplo: Um dataset anonimizando o nome, mas contendo CEP, data de nascimento e gênero — pode ser suficiente para identificar uma pessoa específica, dependendo do contexto.

- **Reidentificação histórica:** Dados considerados seguros em um momento podem se tornar identificáveis no futuro, à medida que mais dados externos ficam disponíveis.

5. Aplicações e Desafios

⚠️ **Limitações: Risco de Reidentificação com Dados Externos**

- **Ataques de inferência:** Mesmo sem identificar diretamente uma pessoa, é possível inferir informações sensíveis (como doenças, crenças, renda) com alta probabilidade.

- **Evolução tecnológica:** Avanços em machine learning, IA e big data podem quebrar técnicas que hoje são consideradas seguras.

5. Aplicações e Desafios

🚫 Casos Onde a Anonimização Não é Suficiente

- **Sistemas que exigem personalização:** Como redes sociais, serviços de streaming e aplicativos bancários, onde é necessário saber quem é o usuário para fornecer o serviço.

- **Cenários de segurança pública e saúde:** Onde informações pessoais precisam ser mantidas para controle epidemiológico, atendimento médico, rastreamento de contatos, etc.

5. Aplicações e Desafios

🚫 Casos Onde a Anonimização Não é Suficiente

- **Auditorias e investigações:** Algumas situações legais exigem que a identidade dos envolvidos seja conhecida, o que inviabiliza anonimização completa.

- **Compliance e obrigações legais:** Existem obrigações legais que demandam armazenamento de dados pessoais por determinado período (ex.: registros fiscais, históricos acadêmicos).

- **Limitações na análise de dados:** A anonimização, quando feita de forma excessiva, pode destruir características importantes dos dados, inviabilizando análises estatísticas, preditivas ou modelagens complexas.

5. Aplicações e Desafios

⚖️ **Dilema: Utilidade dos Dados vs. Privacidade**

- **Trade-off inevitável:** Quanto mais os dados são anonimizados, menos úteis eles se tornam para análises detalhadas. Por outro lado, quanto mais ricos e precisos os dados, maior o risco para a privacidade dos indivíduos.

- **Exemplos práticos do dilema:**
  - Empresas querem entender o comportamento dos consumidores, mas isso exige saber quem são e como se comportam.

  - Pesquisas de saúde pública precisam de informações detalhadas (como localização e histórico médico), mas isso pode expor dados extremamente sensíveis.

5. Aplicações e Desafios

⚖️ **Dilema: Utilidade dos Dados vs. Privacidade**

- **Soluções de equilíbrio (parciais):**
  - **Pseudonimização:** Permite que o dado permaneça útil, mas ainda pode ser revertido se necessário (sob controle rigoroso).
  - **Privacidade diferencial:** Técnica avançada que permite extração de insights estatísticos sem expor informações individuais.
  - **Consentimento:** Em alguns casos, é mais viável trabalhar com dados pessoais mediante o consentimento explícito dos titulares.

5. Aplicações e Desafios

⚖️ **Dilema: Utilidade dos Dados vs. Privacidade**

- **Pergunta ética central:**
  "Estamos dispostos a abrir mão de parte da privacidade para obter melhores serviços, mais personalizados e eficientes? Ou preferimos sacrificar um pouco da personalização em favor da proteção dos nossos dados?"

"Anonimizar dados não é um fim em si, mas uma das muitas ferramentas no equilíbrio entre inovação, segurança, privacidade e responsabilidade social. Nenhuma técnica, isoladamente, é infalível. Por isso, mais do que apenas usar ferramentas, é essencial ter uma cultura de proteção de dados."

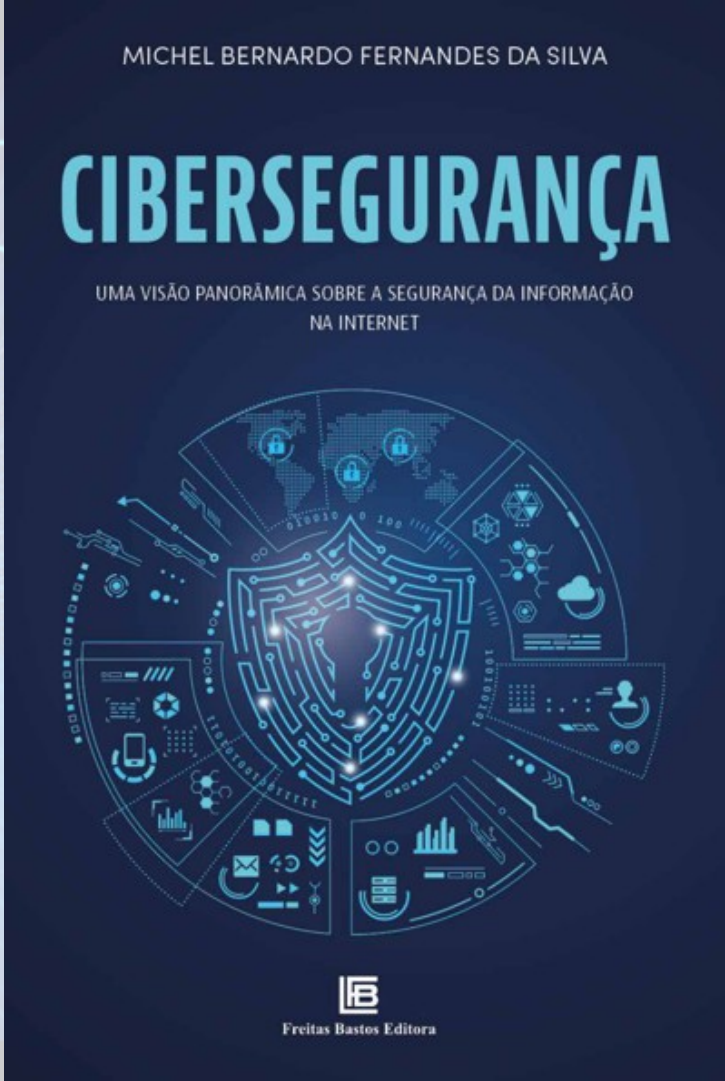
Experimento

https://drive.google.com/file/d/1F2GHnZugSnfNCQUEzGpE2gknEbN-tEPP/view?usp=sharing



Bibliografia

SILVA, Michel Bernardo Fernandes da. Cibersegurança: uma visão panorâmica sobre a segurança da informação na internet. Rio de Janeiro: Freitas Bastos, 2023. E-book. Disponível em: https://plataforma.bvirtual.com.br/Acervo/Publicacao/208076. Acesso em: 25 maio 2025.



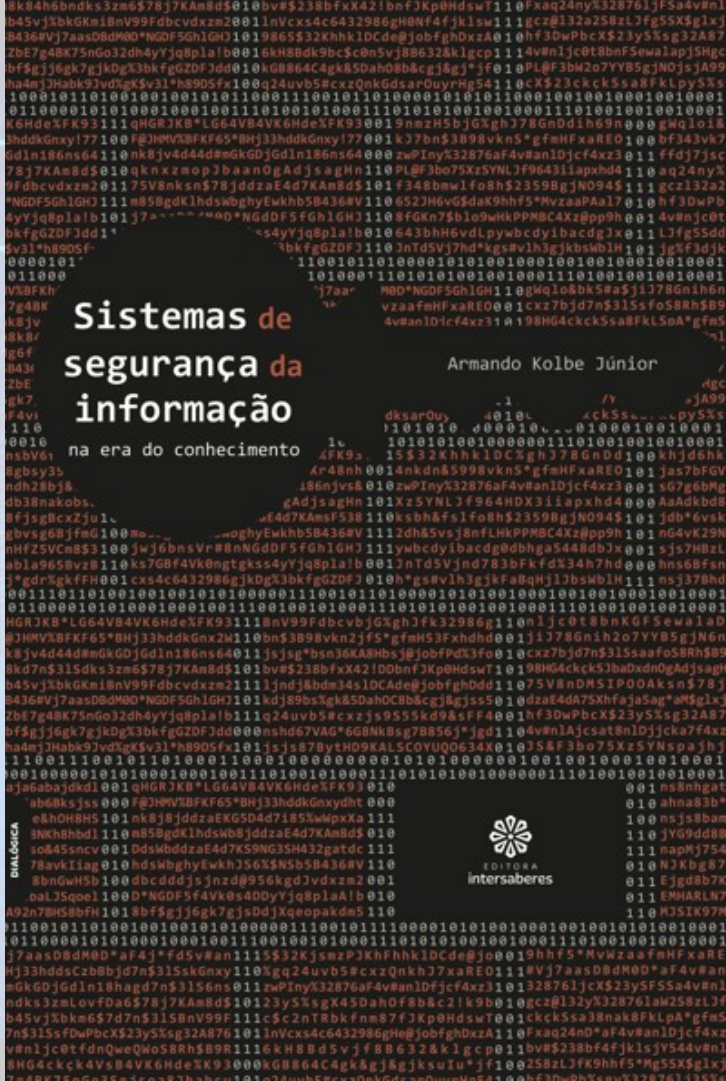
Bibliografia

HINTZBERGEN, Jule et al. Fundamentos de segurança da informação: com base na ISO 27001 e na ISO 27002. Rio de Janeiro: Brasport, 2018. E-book. Disponível em: https://plataforma.bvirtual.com.br/Acervo/Publicacao/160044. Acesso em: 25 maio 2025.



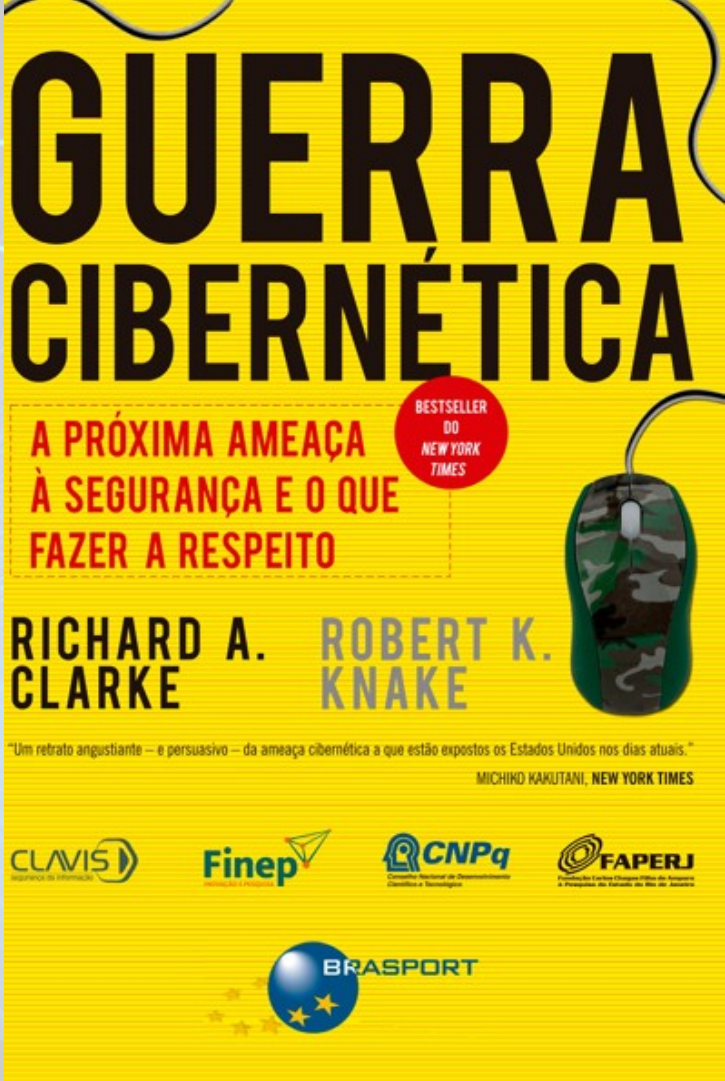
Bibliografia

KOLBE JÚNIOR, Armando. Sistemas de segurança da informação na era do conhecimento. 1. ed. Curitiba: Intersaberes, 2017. E-book. Disponível em: https://plataforma.bvirtual.com.br/Acervo/Publicacao/52012 . Acesso em: 25 maio 2025.



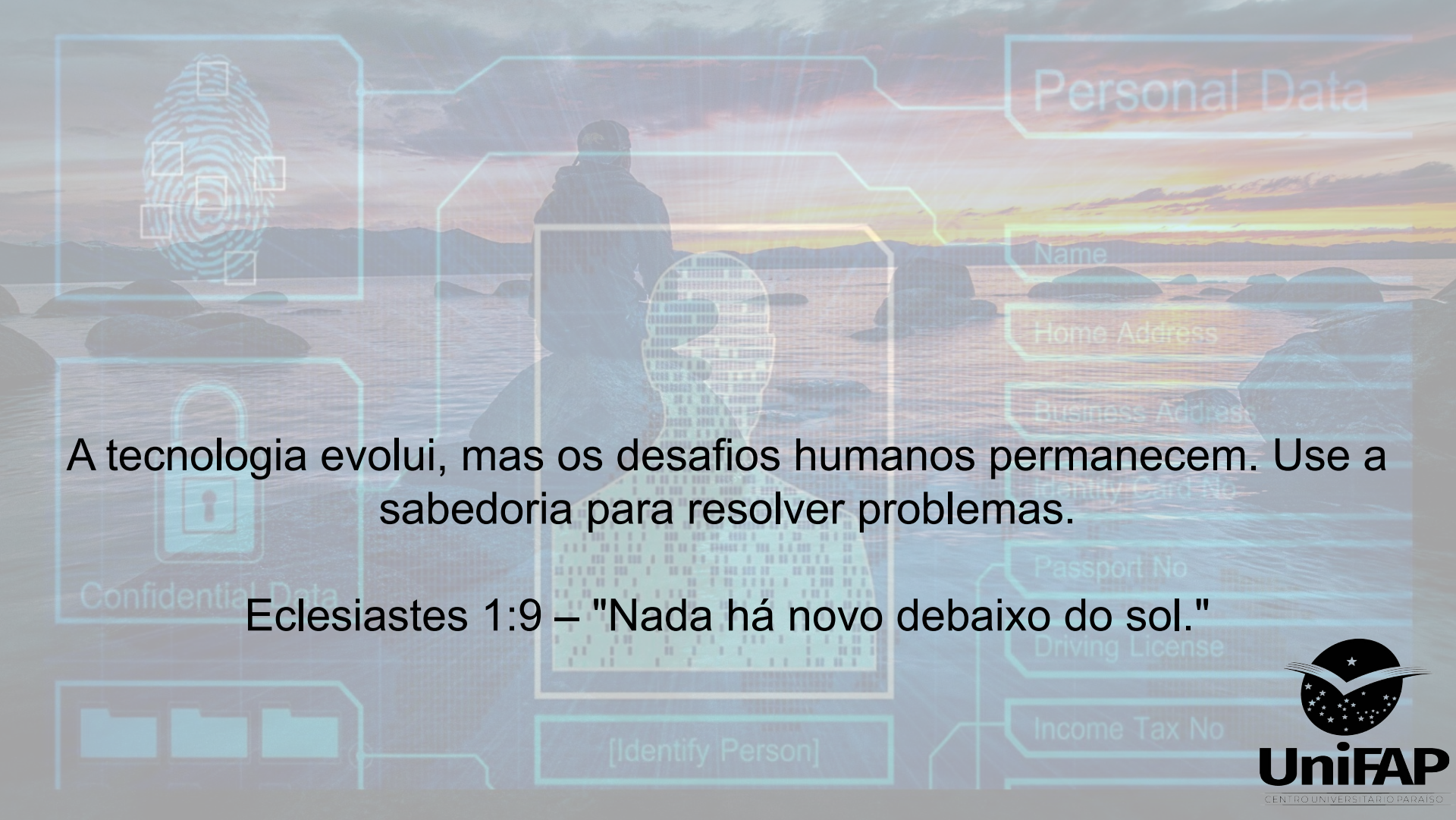
Bibliografia

CLARKE, R. A.; KNAKE, R. K. Guerra cibernética: a próxima ameaça à segurança e o que fazer a respeito. 1. ed. Rio de Janeiro: Brasport, 2015. E-book. Disponível em: https://plataforma.bvirtual.com.br/Acervo/Publicacao/160693. Acesso em: 25 maio 2025.





UniFAP
CENTRO UNIVERSITÁRIO PARAÍSO



A tecnologia evolui, mas os desafios humanos permanecem. Use a sabedoria para resolver problemas.

Eclesiastes 1:9 – "Nada há novo debaixo do sol."

UniFAP
CENTRO UNIVERSITÁRIO PARAÍSO


Personal Data
Name
Home Address
Business Address
Identity Card No
Passport No
Driving License
Income Tax No
Confidential Data
[Identify Person]