|  |
| --- |
| **Anonimização de Dados:**  **Legislação e Conflito de Interesses** |



**UNIVERSIDADE DE SÃO PAULO**

**Instituto de Ciências Matemáticas e de Computação**

Anonimização de Dados:

Legislação e Conflito de Interesses

***Diego Henrique Freitas Pinheiro***

USP - São Carlos

2022

Diego Henrique Freitas Pinheiro

Anonimização de Dados: Legislação e Conflito de Interesses

Trabalho de conclusão de curso apresentado ao Centro de Ciências Matemáticas Aplicadas à Indústria do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, como parte dos requisitos para conclusão do MBA em Gestão de Segurança de Dados.

Área de concentração: Segurança de Dados

Orientador: Prof(a). Dr(a). Cláudio F. M. Toledo

USP - São Carlos

2022

Esta página deve conter a ficha catalográfica e deve ser impressa no verso da folha de rosto.

Para elaborar, acesse o endereço:

<https://www.icmc.usp.br/institucional/estrutura-administrativa/biblioteca/servicos/ficha>

ou procure um bibliotecário na Seção de Atendimento ao Usuário da Biblioteca do ICMC

**ERRATA**

A errata é um elemento opcional, que consiste de uma lista de erros da obra, precedidos pelas folhas e linhas onde eles ocorrem e seguidos pelas correções correspondentes. Deve ser inserida logo após a folha de rosto e conter a referência do trabalho para facilitar sua identiﬁcação, conforme a ABNT NBR 14724 (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2011).

|  |  |  |  |
| --- | --- | --- | --- |
| Errata | | | |
| Folha | Linha | Onde se lê | Leia-se |
| 25 | 3 | Elemento opciotnal | Elemento opcional |
| 59 | 7 | Caracteríscas dos elementos | Características dos elementos |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

FOLHA DE AVALIAÇÃO OU APROVAÇÃO

Elemento obrigatório. Deixe esta folha em branco, pois a folha de aprovação será entregue no dia da defesa.

*A minha esposa pela compreensão, carinho e apoio incansável.*

AGRADECIMENTOS

Elemento opcional, texto em que o autor faz agradecimentos dirigidos àqueles que contribuíram de maneira relevante à elaboração do trabalho.

Ao Dr. José Sampaio, que muito me ensinou contribuindo para o meu crescimento científico.

A Dra. Ludmila dos Santos, pelo incentivo e apoio nos momentos de dificuldade, xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx.

Xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx.

Xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx.

Xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx.

Xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx.

EPÍGRAFE

“Elemento opcional, texto em que o autor apresenta uma citação, seguida de indicação de autoria, relacionada à matéria tratada no corpo do trabalho. A fonte deve constar na lista de referências”

Fulano de Tal (2015)

RESUMO

PINHEIRO, D. H. F. **Anonimização de Dados**: Legislação e Conflito de Interesses. 2022. 52 f. Trabalho de conclusão de curso (MBA em Gestão de Segurança de Dados) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2022.

Instituições bancárias sempre geraram muito lucro com pouca evolução tecnológica ao longo de décadas. Como as pessoas dependiam muito de seus serviços, elas não tinham escolha a não ser conviver com o gargalo que seus sistemas lentos e terminais pouco eficientes traziam. Com a competitividade de inúmeros bancos e a popularização de computadores domésticos e dispositivos móveis, a pressão por melhores serviços trouxe para os bancos a necessidade de uma busca mais acelerada pela transformação digital. Isso trouxe inúmeras possibilidades de negócios e também gerou inúmeros problemas. Convivendo nessa situação e tentando evoluir de maneira favorável, muitas empresas foram bombardeadas com a necessidade de regularizar seus sistemas de forma que atendessem a nova Lei Geral de Proteção de Dados o mais rápido possível, sob pena de uma multa bem alta. Grandes bancos investiram muito nessas práticas até então desconhecidas, e que exigiam um salto tecnológico significativo, com uma mudança de arquitetura e envolvimento de todas as áreas da organização. Este trabalho pretende relatar um estudo de convivência no período de três anos de trajetória em uma instituição bancária focada antes, durante e depois da LGDP, pontuando iniciativas que poderiam minimizar o desperdício de recursos e melhores práticas ignoradas por uma gestão ainda inexperiente no assunto. Contrastar metodologias e a influência que o novo paradigma trouxe ao banco.

Palavras-chave: Transformação Digital. Big Data. LGPD. Anonimização.

ABSTRACT

PINHEIRO, D. H. F. **Data Anonymization**: Legislation and Conflict of Interest. 2022. 52 f. Trabalho de conclusão de curso (MBA em Gestão de Segurança de Dados) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2022.

Mandatory element, it consists of a sequence of concise and objective sentences, in text form. It should present the objectives, methods, results and conclusions, in a single paragraph with 150 to a maximum of 500 words. It must be followed by the keywords and preceded by the document reference. It must be prepared according to the ABNT NBR 6028 standard. Xxxxxxxxx xxxxx xxxxx xxxxxxxxxx xxxx xxxxxx x xxxx xxxxx xxxx xxxx xxxxx x xxxxxxxxx xxxxxxx xx xxxxxx xxxxx xxxxx xx xxxxxxxx xxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxxxxxxx xxxx xxxxxx x xxxx xxxxx xxxx xxxx xxxxx x xxxxxxxxx xxxxxxx xx xxxxxx xxxxx xxxxx xx xxxxxxxx xxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxxxxxxx xxxx xxxxxx x xxxx xxxxx xxxx xxxx xxxxx x xxxxxxxxx xxxxxxx xx xxxxxx xxxxx xxxxx xx xxxxxxxx xxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxxxxxxx xxxx xxxxxx x xxxx xxxxx xxxx xxxx xxxxx x xxxxxxxxx xxxxxxx xx xxxxxx xxxxx xxxxx xx xxxxxxxx xxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxxxxxxx xxxx xxxxxx x xxxx xxxxx xxxx xxxx xxxxx x xxxxxxxxx xxxxxxx xx xxxxxx xxxxx xxxxx xx xxxxxxxx xxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxxxxxxx xxxx xxxxxx x xxxx xxxxx xxxx xxxx xxxxx x xxxxxxxxx xxxxxxx xx xxxxxx xxxxx xxxxx xx xxxxxxxx xxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxxxxxxx xxxx xxxxxx x xxxx xxxxx xxxx xxxx xxxxx x xxxxxxxxx xxxxxxx xx xxxxxx xxxxx xxxxx xx xxxxxxxx xxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxxxxxxx xxxx xxxxxx x xxxx xxxxx xxxx xxxx xxxxx x xxxxxxxxx xxxxxxx xx xxxxxx xxxxx xxxxx xx xxxxxxxx xxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxxxxxxx xxxx xxxxxx x xxxx xxxxx xxxx xxxx xxxxx x xxxxxxxxx xxxxxxx xx xxxxxx xxxxx xxxxx xx xxxxxxxx xxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xx xxxxxxxx xxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxxxxxxx xxxx xxxxxx x xxxx xxxxx xxxx xxxx xxxxx x xxxxxxxxx xxxxxxx xx xxxxxx xxxxx xxxxx xx xxxxxxxx xxxx xxxxx xxxxx xxxxx xxxxx xxxxx xxxxx.

Keywords: Keyword 1. Keyword 2. Keyword 3. Keyword 4.

LISTA DE ILUSTRAÇÕES

Elemento opcional, elaborada seguindo a mesma ordem apresentada no texto com cada item designado por seu nome e respectivo número de página.

Figura 1 – Variação do teor de gordura com tratamento térmico.................................... 20

Figura 2 – Variação da acidez com tratamento térmico................................................... 23

Figura 3 – Xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx ............................................ 46

Figura 4 –

Figura 5 –

LISTA DE TABELAS

Elemento opcional, elaborada seguindo a mesma ordem apresentada no texto com cada item designado por seu nome e respectivo número de página.

Tabela 1 – Variação do teor de gordura com tratamento térmico....................................20

Tabela 2 – Variação da acidez com tratamento térmico...................................................23

Tabela 3 – Xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx..............................................46

Tabela 4 –

Tabela 5 –

LISTA DE ABREVIATURAS E SIGLAS

Elemento opcional. É composto de uma relação alfabética das abreviaturas e siglas utilizadas no texto seguido do seu significado.

ABNT – Associação Brasileira de Normas Técnicas

LGPD – Lei Geral de Proteção de Dados

GDPR – General Data Protection Regulation

PAAS – Platform as a Service

MVP – Minimum Viable Product

d\_t0 – Escorregamento relativo à resistência máxima

E\_c – Módulo de elasticidade do concreto

FLA – Flambagem Localizada da Alma

f\_c – Resistência à compressão do concreto

G\_c – Energia de fratura à compressão

h – Largura de banda de fissuras

K\_t – Rigidez tangencial

K\_n – Rigidez normal

μ\_ – Coeficiente de atrito

σ\_ – Tensão normal

τ\_ – Tensão de cisalhamento

L – Conector proposto de superfície lisa

R – Conector proposto de superfície com ranhuras

RP – Conector proposto de superfície com ranhuras e furos

LISTA DE SÍMBOLOS

Elemento opcional. Elaborada seguindo a mesma ordem apresentada no texto com o significado correspondente.

ºC Graus Celsius

K Graus Kelvin

L\* Luminosidade

SUMÁRIO

Elemento obrigatório que consiste na enumeração das partes do trabalho, na mesma ordem e grafia em que aparecem seguidas do número de páginas.

1 **INTRODUÇÃO**............................................................................................................. 17

1.1 **Instruções gerais de apresentação**.............................................................................. 17

1.2 **Apresentação gráfica**................................................................................................... 18

1.3 **Numeração progressiva**............................................................................................... 18

2 **REVISÃO BIBLIOGRÁFICA**...................................................................................... 19

2.1 **Considerações Iniciais**................................................................................................ 19

2.2 **Transformação Digital**................................................................................................ 19

2.3 **Big Data**....................................................................................................................... 21

2.4 **Lei Geral de Proteção de Dados**................................................................................ 21

3 **MODELOS DE CITAÇÃO**.......................................................................................... 23

3.1 **Um autor**................................................................................................................... 23

3.2 **Dois autores**.............................................................................................................. 23

3.3 **Até 3 autores**............................................................................................................. 23

3.4 **Mais de três autores**.................................................................................................. 23

3.5 **Sem autor**................................................................................................................... 24

3.6 **Mesmo autor e mesmo ano**......................................................................................... 24

3.7 **Mesmo sobrenome e ano**........................................................................................... 24

3.8 **Autor entidade**.......................................................................................................... 24

3.9 **Mesmo autor e anos diferentes**.................................................................................. 25

3.10 **Vários trabalhos de diferentes autores**.................................................................... 25

4 **CONCLUSÕES**............................................................................................................ 27

**REFERÊNCIAS**............................................................................................................ 29

**GLOSSÁRIO**................................................................................................................ 33

**APÊNDICE A** – Xxxxxx xxxxx xxxxx......................................................................... 35

**APÊNDICE B** – Xxxxxx xxxxx xxxxx....................................................................... 36

**ANEXO A** – Xxxxxx xxxxx xxxxx............................................................................. 37

**ÍNDICE**......................................................................................................................... 39

**1 INTRODUÇÃO**

Instituições financeiras, em especial, bancos, acumularam décadas de dados pessoais e dados transacionais de seus clientes e parceiros. Durante muito tempo os bancos conseguiam operar e gerar muito lucro apenas com o nome, algum documento de identificação, profissão e estado civil. Inúmeras evoluções tecnológicas trouxeram uma maior competitividade entre as instituições e hoje elas precisam oferecer produtos e serviços cada vez mais atraentes, de forma a fidelizar os clientes e lucrar com o bom atendimento.

Essa situação de muita propaganda e oferta, gerou um incômodo desnecessário às pessoas. Spams, ligações e pop-ups se tornaram as formas mais negativas de se aproximar dos clientes. Técnicas como Big Data, integração de serviços em nuvem, mineração de grandes volumes de dados vieram como uma proposta de oferecer o que o banco sempre ofereceu, porém de forma mais assertiva e menos incômoda.

**1.1 Contextualização**

Agora o banco consegue de fato saber para quem, quando e como oferecer os produtos certos para os clientes que estão dispostos a obtê-los. No entanto, isso também viola um direito intrínseco a qualquer cidadão de não querer ser incomodado com essas ofertas, ou melhor, ter o controle de quem possui seus dados e ter o conhecimento de como são usados.

Buscando defender esse direito, governos de diversos países revisitaram a forma como protegiam seus dados e enxergaram necessidades de uma regulamentação. A União Europeia com o General Data Protection Regulation (GDPR) foi o mais importante e influente, que não só obrigou uma mudança radical nos termos de uso e coleta de dados, como incitou leis semelhantes a todos os países no mundo.

Como profissional de TI, estar ciente do que é a lei, a quem se fundamenta e como deve ser implementada no Brasil, no caso, a Lei Geral de Proteção de Dados (LGPD), garante uma maior chance de sucesso em seu desenvolvimento. Empresas e organizações que trabalham com grandes volumes de dados geralmente estão sempre focados em gerar mais lucro, tentando evoluir seu negócio e extrair maior lucratividade dessas informações. Seus postos de trabalho costumam não ser tão bem definidos, e acabam atuando em inúmeras frentes, freando o desenvolvimento de tecnologias mais eficientes e produtivas.

**1.2 Motivação**

Nesse contexto, analisando uma instituição bancária, é possível enxergar inúmeras obstáculos que precisam ser vencidos para implantação da lei:

**Mudança de mentalidade** - Quando a lei foi anunciada, e o governo definiu um prazo para que as empresas se ajustassem, a fim de evitar uma multa pesada, o banco acreditava que perderia muito dinheiro anonimizando os dados, e queria achar contornos para sua aplicação.

**Evolução tecnológica** - O banco ainda lucra muito com tecnologias menos modernas, como mainframes e bases de dados relacionais simples. A transformação digital trouxe formas diferentes de agregar mais valor com aquilo que o banco já possui, então não havia motivos para investir numa migração digital completa e custosa, mesmo que fosse mais eficiente, algo que a lei provavelmente exigiria para ser corretamente implantada.

**Aprimorar seus principais funcionários** - Há funcionários que assumem cargos de alta importância há décadas, tomando as maiores decisões, sempre muito conservadoras. Pelo desconhecimento, suas preocupações com riscos sensibilizam muitas áreas, e a governança se torna muito engessada, com pouca possibilidade de avanço.

**1.3 Objetivo Geral**

O objetivo deste trabalho é relatar, mediante profunda análise, se no período de 3 anos o banco conseguiu evoluir seus processos para ser capaz de implantar a lei com a qualidade e a atenção exigidas, benefícios ou prejuízos ao fluir do negócio e seus processos.

**1.4 Objetivos Específicos**

Para isso os seguintes objetivos específicos foram estabelecidos:

* Avaliar erros e acertos no emprego dos processos.
* Descrever técnica empregada para anonimizar dados.
* Avaliar o impacto da anonimização e sua aderência à LGPD.
* Comparar com processos realizados em outros países
* Ilustrar o cenário antes e depois da implantação

**2 REVISÃO BIBLIOGRÁFICA**

**2.1 Considerações Iniciais**

Este capítulo inicia o leitor com conceitos que revisam temas comumente usados durante o relatório, como transformação digital, Big Data e LGPD. Essa contextualização se inicia na seção 2.2 com e pretende conduzir o leitor a um melhor entendimento do relatório.

**2.2 Transformação Digital e Big Data**

Shannon et al. (1948) propôs o "Teorema matemático da comunicação" onde defende que os sinais de comunicação devem ser tratados isoladamente do significado das mensagens que transmitem. A quantidade de conhecimento transmitida por um sinal não está diretamente relacionada ao tamanho da mensagem.

Para Kane (2017, p.2):

“Essa implementação de novas tecnologias é apenas uma parte das outras questões, como estratégia, gerenciamento de talentos, estrutura organizacional e liderança, tão importantes, senão mais importantes, que a tecnologia para a transformação digital”.

A transformação digital de um negócio introduz serviços de valor agregado, utilizando novas tecnologias para gerar um impacto direto na entrega de resultados e na experiência do cliente, tornando esta uma das principais diretrizes da empresa. Canais e processos também serão transformados e a necessidade de buscar novos modelos e fluxos de receita com base nas necessidades do cliente se tornará necessária e, consequentemente, será obrigatório mudar a cultura interna da empresa para incluir o digital em tudo. (LLORENTE, 2016).

Originalmente, big data é um termo usado para se referir a técnicas de armazenamento de grandes volumes de dados (Jacobs, 2009). Atualmente, processar e oferecer acesso rápido a informações geradas a partir desses dados é o desafio tecnológico ao qual se fundamenta.

Mainframes eram a principal forma de armazenagem e processamento de dados para muitas empresas, e eram acessados através de terminais com baixo poder de processamento e acessíveis apenas a funcionários. (REINSEL et al. 2009)

Inovações trouxeram técnicas mais rápidas de leitura desses mainframes e foram um passo importante para a migração de arquitetura em muitos bancos. Evoluíram para serviços em nuvem com leitura desse processamento em bases mais dinâmicas e escrita dos dados em em lotes, geralmente na madrugada.

O fato mais importante foi a troca de informação entre usuário e empresa. Com formas mais apuradas de leitura e escrita dos dados, houve um interesse maior e mais vantagens em outros tipos de dados, acompanhar transações e gerar estatísticas mais precisas das oscilações dos valores, padrões, entregando um processamento com mais valor agregado ao cliente.

Esse conjunto de técnicas gerou um tipo novo de profissional, o engenheiro de dados que coordena e administra bases de dados dos mais diferentes tipos; analistas de negócio com perfis mais técnicos e uma exigência maior de estratégias para criar produtos competitivos para os perfis analisados pelo processamento de grandes volumes de dados.

**2.3 Lei Geral de Proteção de Dados**

Tendo em vista grandes mudanças ocorridas na tecnologia da informação, movida com mais intensidade pelo tratamento automatizado de dados, o Tribunal Constitucional Federal (TCF) alemão de 1983 levou argumentos oriundos do direito constitucional e dos direitos humanos à proteção de dados para uma importante decisão.

Percebeu que era anti-democrático um indivíduo ter mais conhecimento sobre um do que o outro tem de si, e não ter a possibilidade de saber ou tomar ação sobre o que sabe. Ele poderia abrir mão do exercício de liberdades protegidas por direitos fundamentais, como a liberdade de reunião e de associação, pela insegurança do conhecimento que a outra parte poderia ou não possuir de si.

Em 25 de maio de 2018 o tema voltou à tona com mais rigor, entrou em vigor o Regulamento Geral de Proteção de Dados da União Europeia (RGPD-UE), que dá o direito de proteção de dados aos seus cidadãos, agora com maior efetividade e efetivação visível. O fato foi tão importante, que fez surgir o maior mercado euro-asiático para dados após a reforma do direito à proteção de dados do Japão, relatado por Döhmann (2020).

Além disso, em agosto de 2018, o Brasil aprovou uma lei geral de proteção de dados, também moldada por muitos conceitos de proteção de dados defendidas na Europa. Historicamente, esse direito já foi clamado há pelo menos 50 anos. Hoje ele volta à tona se tornando um dos critérios mais importantes para que o país possa continuar fazendo negócios com países que possuem políticas de proteção de dados.

Por meio de artifícios jurídicos, a lei 13.709/2018 regulariza o uso dos dados pessoais, e a forma como são administrados, protegendo assim os direitos dos indivíduos, tratando os dados com transparência e responsabilidade. No art. 7ª. existe uma previsão dos requisitos para que se realize o tratamento dos dados pessoais, sendo melhor detalhado no art. 11º.

Para implementação da lei, um sistema computacional precisa avaliar se os dados de forma que não estejam inseridos em nem um dos dez critérios estabelecidos nas bases legais:

1. **Consentimento**: O usuário autoriza o uso das informações que forneceu.
2. **Cumprimento de obrigações legais ou regulatórias pelo controlador**: para tratar de condições impostas por leis ou determinações do poder legislativo, o controlador dos dados pode manter e utilizar os dados.
3. **Pela Administração Pública**: o Estado fará o tratamento e compartilhará dados que são essenciais na realização de políticas públicas ou no cumprimento de contratos públicos.
4. **Para a realização de estudos por órgão de pesquisa**: organizações que cuidam do mapeamento de índices e precisam desses dados para alguma análise de interesse público e legítimo podem manter os dados para esses fins.
5. **Para execução de contrato ou de procedimentos preliminares relacionados a contrato do qual seja parte o titular, a pedido do titular dos dados**: caso exista algum vínculo contratual ou acordo, o titular pode requisitar a utilização de seus dados.
6. **Para exercício regular de direitos em processo judicial, administrativo ou arbitral**: para processos em trâmite no Poder Judiciário ou com a Administração Pública, será requerido ao titular o compartilhamento de seus dados.
7. **Para proteção da vida ou da incolumidade física do titular ou de terceiro**: no caso de alguma ameaça à integridade física do titular ou de pessoa relacionada, é possível que seja requerido o fornecimento de certos dados.
8. **Para tutela da saúde, exclusivamente, realizado por profissionais de saúde, serviços de saúde ou autoridade sanitária**: é possível que seja requisitado ao titular a tutela de seus dados em situações envolvendo profissionais da saúde ou autoridades sanitárias.

**Para atender interesses do controlador ou terceiro**: o artigo 10 da lei ressalta a mais polêmica das bases legais, o de legítimo interesse.

***Segundo a LGPD (INCLUIR REFERÊNCIA):***

***“Art. 10.*** *O legítimo interesse do controlador somente poderá fundamentar tratamento de dados pessoais para finalidades legítimas, consideradas a partir de situações concretas, que incluem, mas não se limitam a:*

***I –*** *apoio e promoção de atividades do controlador; e*

***II –*** *proteção, em relação ao titular, do exercício regular de seus direitos ou prestação de serviços que o beneficiem, respeitadas as legítimas expectativas dele e os direitos e liberdades fundamentais, nos termos desta Lei.*

***§ 1º*** *Quando o tratamento for baseado no legítimo interesse do controlador, somente os dados pessoais estritamente necessários para a finalidade pretendida poderão ser tratados.*

***§ 2º*** *O controlador deverá adotar medidas para garantir a transparência do tratamento de dados baseado em seu legítimo interesse.*

***§ 3º*** *A autoridade nacional poderá solicitar ao controlador relatório de impacto à proteção de dados pessoais, quando o tratamento tiver como fundamento seu interesse legítimo, observados os segredos comercial e industrial.”*

1. **Para proteção do crédito**: previne que o devedor se utilize do direito à proteção de dados para se esquivar do pagamento de dívidas.

**2.5 Trabalhos Relacionados**

Uma técnica de micro-dados conhecida como K-Anonimização foi desenvolvida por (SWEENEY, 2002), onde uma estrutura é construída com dados de uma pessoa de forma que produza novos dados que os originais não sejam identificados, permitindo que continuem sendo utilizados, substituindo dados por algum caractere ou generalizando valores individuais por categorias definidas.

[Dwork et al. 2014] propôs uma técnica que consiste em adicionar ruídos aos dados de forma que eles possam ser menos identificáveis e ainda utilizáveis, chamada Privacidade Diferencial. Para certos contextos, onde há lacunas que podem ser exploradas para fazer valer a lei e ainda manter a integridade de alguns dados, ela pode ser útil para mascarar alguns dados para um tipo de busca, e estar íntegro para outro.

Celidonio, Neves e Doná (2020) se basearam na ISO/IEC 27701/2019 para estabelecer uma comissão do SGPI aos principais stakeholders da organização; executar o *gap analysis* completo do ambiente, que consiste em um método de avaliação do desempenho de uma unidade de negócios para determinar se os requisitos ou objetivos de negócios estão sendo atendidos e, se não, quais medidas devem ser tomadas para atendê-los, considerando aspectos técnicos e jurídicos; preparar, qualificar e envolver toda a equipe para a adequação à nova legislação; implementar correções necessárias considerando pessoas, processos e produtos da Governança de TI.

No trabalho de [Liao et al. 2017] é possível preservar a privacidade da localização para uma única consulta e a privacidade da trajetória para consultas contínuas, usando um algoritmo baseado na técnica de *k-anonymity* para proteger a privacidade da trajetória do usuário em serviços baseados em localização.

[Agostinelli et al., 2019] afirmou que para garantir a aplicabilidade do GPDR, as empresas precisam repensar sua notação de modelagem de processos de negócios (BPMN) e como gerenciam os dados pessoais dos usuários dentro do negócio. O autor utilizou o BPMN em uma empresa do setor de telefonia, com o objetivo de aplicar o GPDR para garantir a privacidade dos dados dos usuários, no processo de credenciamento de novos usuários e a responsabilidade dos controladores de dados sobre o processo. Para garantir que não haja violação do princípio de privacidade de dados, o autor propôs que contramedidas ad hoc sejam implementadas durante a etapa de automação do BPMN de forma preventiva.

Além disso, concluiu que o desenho da modelagem de processos é importante para a implementação bem-sucedida da lei de privacidade de dados. Na análise, levantou os pontos críticos do GDPR em relação às restrições de privacidade e propôs um conjunto de padrões de design para capturar e integrar essas restrições nos modelos representados no BPMN.

3 PROBLEMA

A princípio, a lei trouxe um dilema: abrir mão de dados que potencialmente poderiam ser explorados para fins comerciais em um momento onde o banco havia acabado de alcançar um nível tecnológico desejado para extrair algo deles. Isso implicaria que a instituição agora deveria desviar esforço, tempo e capital em uma implementação que mudaria muitos de seus negócios.

Houve insistência em seguir planos arquiteturais que escondam os dados ao invés de anonimizá-los, seja separando as bases ou filtrando a exposição deles a partir de um critério. Poucos foram os esforços para de fato entender os critérios da lei e analisar o cenário após a anonimização dos dados.

Técnicas de Big Data e análises de grandes volumes de dados ainda eram bem recentes no banco, e poucas áreas trabalhavam com eles fora dos mainframes. Anonimizar dentro deles era impensado, pois eles são ótimos com processamentos de números, e não com análises de critérios, vínculos e avaliações de estado.

O banco precisaria dar um salto tecnológico em uma direção oposta a seus interesses e pagar caro por isso. A solução encontrada foi formar um time dedicado para avaliar possíveis alternativas, explicando para as áreas que detêm os dados como seriam afetadas e negociar uma forma de cooperação.

4 METODOLOGIA

Para dar um senso de evolução, faremos uma análise inicial do contexto tecnológico e organizacional de umas das áreas mais lucrativas dentro do banco: área de investimentos. Nela os clientes aderem ao que o banco chama de “produtos”, realizam uma contratação e o dinheiro da “compra” se torna um ativo dentro do banco, seguindo um fluxo de interesse dentro do que seria “rentável” do ponto de vista da área de negócio. Isso leva a se atingir um valor de rendimento que faça sentido para o contratante e para o próprio banco, que ganha um valor de comissão por cada ganho que o cliente receba.

Do ponto de vista tecnológico, aquele “número” que representa o valor investido passa por uma série de “análises de encaixe”. Para isso, o sistema define, baseado em históricos, regras e padrões, qual seria o melhor lugar para somar aquele valor de forma que o montante atual se torne ainda maior após um ou mais dias de processamentos relacionais, entre movimentações de números e diferentes origens.

Para o sistema, não importa se uma empresa faliu ou se obteve um grande sucesso no fim do dia. O sistema apenas avalia os históricos de dinheiro que entrou e dinheiro que saiu; origem com número maior, origem com número menor. Entretanto, para que seja resumido dessa forma simplista, uma construção altamente complexa de regras foi feita. Os dados são organizados de forma que esse entendimento por parte do sistema, que “apenas” processa as regras, possa fazer isso da forma mais rápida e precisa que puder.

No contexto de análise, a inserção de dados era feita via terminais de atendimento, onde um gerente preenche a tela com alguns valores, escolhe as opções de produto e valores de investimento. Esses dados trafegam de maneira segura via micro serviços até uma base de dados Oracle na plataforma como serviço do banco (PAAS). Ao anoitecer, quando essas bases de dados não serão mais “escritas”, um outro sistema, um “robô”, faz as leituras dessas bases, coloca em um diretório “neutro”, onde outro robô lê e transfere para uma linguagem e formato que faça sentido para os mainframes processarem juntamente com dados já presentes.

Nessa época, em meados de dezembro de 2019, essa arquitetura ainda predominava no banco com pequenas iniciativas de usar técnicas que permitissem atualizações dos dados registrados e processadas em tempo real. Tais iniciativas aconteciam de forma bem discreta e experimental, com 1\10 dos times focado em elaborar essas propostas e mostrar resultados, enquanto outros times cuidavam de colocar mais regras, novos produtos e atender a área de negócios, corrigir bugs e ampliar o sistema.

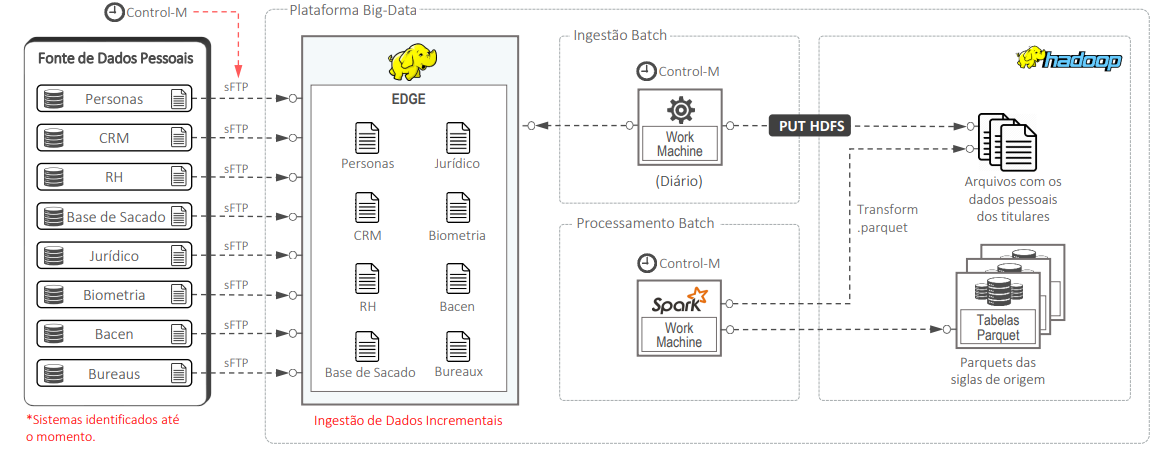


Figura 1 - Representação da arquitetura de extração de dados da instituição / FONTE: Macro Arquitetura de Solução (Cabral, Thiago, 2019)

Em janeiro de 2020, quando a sigla LPD foi criada no banco e todas as áreas foram avisadas de que um *Minimum Viable Product* (MVP) seria entregue. O MVP seria finalizado até a data prevista para entrada em vigor da lei geral de proteção de dados, 18 de agosto de 2020. Essa decisão trouxe inúmeros questionamentos, reuniões, palestras, apresentações de propostas e mudanças arquiteturais para todas as áreas do banco, que até então trabalhavam isoladas.

Áreas que trabalham com dados sensíveis de clientes foram avisadas que precisam dispor desses dados para análise sistêmica da sigla LPD, bem como informações atreladas como produtos contratados, ordens judiciais e outras informações que pudessem sustentar um legítimo interesse.

A discussão se estendeu, inúmeras discordâncias, conflito de interesses pela falta de entendimento do que seria a lei, se a multa de fato se aplicaria, se ao invés de anonimizar, não seria possível segregar as informações para um ambiente *offline* e seguir o modelo atual de trabalho, baseado em dados brutos e completos, ao invés de percebê-los como estatística.

O banco havia acabado de finalmente conseguir técnicas boas para analisar os dados de seus clientes e ter bons relatórios para propostas de novos produtos, gerar mais capital de negócio e investir mais em Big Data, segurança e arquitetura de nuvem, mas precisou investir isso em um processo que anonimizaria muitos desses dados, e esse foi o maior dilema.

A arquitetura do sistema que analisaria as regras para os dados seguiu assumindo que em algum momento uma ferramenta faria a leitura de todos os dados nas diferentes áreas de diferentes tecnologias do banco e as unificaria em um grande volume de dados, podendo ser lidos pela tecnologia Apache HBase a qualquer momento.

Usando a tecnologia Spring Boot para criação de micro serviços e leitura das bases em HBase, armazenando informações chave em bases Oracle e usando Apache Camel como um orquestrador do fluxo de chamadas aos serviços e roteador de informações, o primeiro MVP do time de Confirmação de Existência e Dossiê foi criado.

Engenheiros de dados foram contratados para substituir a ferramenta de unificação dos dados que nunca surgiu. Conseguiram povoar a base para atender uma homologação mínima enquanto o MVP seria aprimorado.

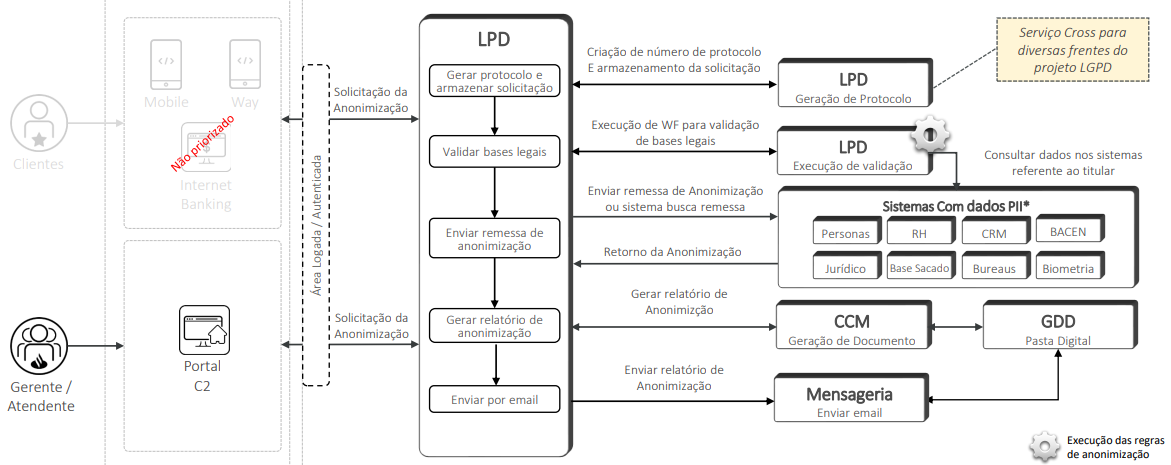


Figura 2 - Versão final da arquitetura do projeto / FONTE: Macro Arquitetura de Solução (Cabral, Thiago, 2019)

Nessa etapa, formava-se um time de desenvolvimento para discutir uma arquitetura de solução para o problema efetivo da anonimização de dados. Após inúmeras discussões e palestras, ficou decidido que o time construiria uma esteira analítica da lei. Ou seja, o sistema que anonimizaria os dados no banco, na verdade seria um grande funil de dados para filtrar os CPFs ou CNPJs elegíveis para serem anonimizados segundo a lei.

O guia DAMA-DMBoK® diz que a Governança de Dados deve exercer autoridade e controle sobre o gerenciamento de ativos de dados, planejando, monitorando e executando as particularidades da norma sobre todas as questões. Ela deve se sobressair à todas as decisões relativas ao tratamento de dados das áreas, e não se adequar ao que é dito individualmente por cada uma delas.

O time seguiu com uma arquitetura ainda em constante mudança. Inicialmente, um framework de integração (Apache Camel) orquestraria chamadas a micro serviços, salvando informações específicas em uma base de dados isolada das demais bases do banco. Um deles faria o registro da solicitação de anonimização, outro confirmaria a existência do indivíduo na base unificada com todos os dados do banco e o outro seria o motor que verifica se os dados atendem ou não a critérios estabelecidos na lei.

Esse motor, feito na Plataforma de automação de workflow e decisão Camunda, cria threads de chamadas a serviços que avaliam os dados quanto a sua legibilidade perante uma base legal específica. Produtos ativos ou inativos a pelo menos 15 anos, ordens judiciais em aberto contra ou em nome do banco e legítimo interesse segundo critérios estabelecidos no artigo 10 da lei compõe o BPMN estruturado para operar dentro do framework, semelhante a proposta de [Agostinelli et al., 2019], porém, de maneira automatizada.

O relatório gerado é enviado para o e-mail do usuário que registrou a solicitação de anonimização. Um lote contendo a identificação de todos os solicitantes do dia, elegíveis à anonimização, é enviado para suas respectivas áreas dentro do banco, transferindo assim a responsabilidade de anonimizar o dado, ao contrário do que foi proposto por [Liao et al. 2017] que usa um algoritmo para agrupar os dados a partir de um critério no lugar que eles estão e aplica a anonimização já na fonte.

Como o motor depende da atualização dos dados unificados, rotinas de leitura tentam extrair esses dados, e a expectativa é que uma nova solicitação de anonimização não possa ser feita para alguém que já deveria ter sido anonimizado. Muitas pessoas tentam forçar esse tipo de sistema a fim de ganhar alguma ação judicial por não ter seus dados anonimizados.

Quando uma área tem que se responsabilizar pela anonimização, ela é quem precisa pagar pela multa, e não o banco em si. A área LGPD dentro do banco foi criada para proteger a instituição contra todos estes interesses conflitantes. A governança permite muita autonomia entre as áreas dentro da instituição, mas busca mais que tudo preservar-se de dados unilaterais causados por legislações e acordos que possam comprometer seus interesses.

As técnicas usadas para anonimização ficaram a critério de cada área. A maioria optou por simples substituição parcial do dado por “####”, tentando preservar alguma informação que ainda pudesse ser usada como estatística. Na verdade, equipes foram criadas para tentar maquiar esses dados de forma que não percam seu valor original.

**4 CONCLUSÃO**

**4.1 Resultados**

O interesse do banco era preservar os dados acima de tudo. Criar uma área que unificou os dados em uma base “isolada”, que passariam por critérios de avaliação e classificadas quanto a sua elegibilidade foi uma manobra conveniente e eficaz para seus objetivos. Cada área ficou responsável por fazer o que quisesse com a classificação, até mesmo isolar aquela informação ao invés de anonimizar.

Estatisticamente falando, o interesse pelo pedido de anonimização foi muito baixo, e o banco também se reservou o direito a apenas “estar pronto” caso fosse solicitado. Os pedidos são atendidos em até 15 dias, sendo a confirmação de solicitação e dossiê recebidas na hora.

O aprendizado adquirido pela área e a base unificada fomentaram novos planos de negócios e possibilidades que já eram percebidas, mas não priorizadas. Pelo grande avanço nas técnicas de Big Data e uma base unificada bem construída, o banco pôde expandir sua área de marketing como um efeito colateral da corrida pelo emprego da lei.

**4.2 Discussão**

Por existirem áreas com origens de dados e tecnologias muito diferentes, com gestores e governanças defendendo interesses próprios, seria inviável impor uma técnica padronizada de anonimização que forçasse a aplicação da lei. As políticas e a forma como o banco defende seus processos dificulta o exercício de regras mais gerais.

Mesmo que um ganho mais do que favorável em termos do que foi investido versus o que foi conquistado, a mentalidade do banco não mudou. O exercício da lei mudou alguns processos, como a forma com que os dados são preservados, consentimentos e cláusulas mais bem esclarecidos. Porém, o banco ainda considera a aquisição e permanência de dados o ativo mais precioso de todos, e isso é bem perigoso.

Empresas mais novas tem maturidade suficiente para acreditar que podem aplicar a lei e preservar apenas os dados que foram consentidas a elas, e fazer um bom trabalho para que mais pessoas acreditem e consintam da mesma forma. A disposição do banco em investir dois anos nesse projeto se deu por conta da conveniência de ainda poder manter uma base refletida dos dados que pode ou não nunca ser apagada.

**4.3 Considerações finais**

Não há como afirmar se a arquitetura escolhida foi a melhor ou pior, mas ela certamente foi a mais coerente para os interesses propostos. A tecnologia usada poderia ter sido bem menor se a mentalidade do banco fosse diferente, mas não teriam evoluído frentes desconhecidas em tão pouco tempo pela pressão da multa imposta pelo não emprego da lei.

Com o conhecimento e análise feitas nesse artigo, um diretor executivo conseguiria ter a percepção de que embora seja exaustivo aplicar a lei, diferentes arquiteturas podem trazer resultados satisfatórios independente da escolha. Firmar um objetivo e evitar que conflitos de interesses distorçam a visão que se tem é sempre a melhor estratégia.

Desistir dos dados e aplicar a lei com uma tecnologia mais simples que simplesmente lei a base e aplique a anonimização pode ser mais barata, mas unificar a base em uma mais poderosa e usar nelas os conceitos da lei trouxe um ganho surpreendente que compensou todo o esforço.

**REFERÊNCIAS**

Elemento obrigatório.

Devem ser apresentadas em uma única lista em ordem alfabética no final do trabalho. A margem deve ser alinha à esquerda. As referências devem ser digitadas, com espaçamento simples e para separar uma referência da outra, usar dois espaços simples em branco.

Modelos de referência:



ACCIOLY, F. **Publicação eletrônica [mensagem pessoal]**. Mensagem recebida por mtmendes@uol.com.br em 26 nov. 2003.



ASSIS, A.P.; HERNANDEZ, H.M.; COLMANETTI, J.P. **Curso de barragens**: publicação G.AP–AA006/02. Brasília: Departamento de Engenharia Civil e Ambiental, Universidade de Brasília, 2006.



ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 5739**: Concreto - ensaios de compressão de corpos-de-prova cilíndricos. Rio de Janeiro, 2007.



BEZERRA, L.M. **Estudo teórico-experimental da ligação entre pilares mistos preenchidos e vigas pré-moldadas de concreto**. 2011. 123f. Tese (Doutorado) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2011.



BOUAZAOUI, L. et al. Static behaviour of a full-scale steel–concrete beam with epoxy-bonding connection. **Engineering Structures**, v.30, n.7, p.1981-1990, July 2008.



BRASIL. Lei no 7.000, de 20 de dezembro de 1990. Dispõe sobre a proibição da pesca. **Diário Oficial da União**, Brasília, DF, 21 jan. 1991. Seção 1, p.51.



CAMAPUM DE CARVALHO, J. et al. La Recosntituion dês éprovettes em laboratoire: théorie et pratique opératoire. Paris: Laboratorie Central des Ponts et Chaussées, 1987. (Rapport de Recherche LPC N° 145).



EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. **Unidade de Apoio, Pesquisa e desenvolvimento de Instrumentação Agropecuária (São Carlos, SP)**. Paulo Estevão Cruvinel. Medidor de temperatura para solos. BR n. PI 8903105-9, 26 jun. 1989, 30 maio 1995.



ENGELMAN, D.M. Membranes are more mosaic than fluid. **Nature**, n.438, p.578-580. Doi: 10.1038/nature04394.



HAYT JUNIOR, W.H.; BUCK, J.A. **Eletromagnetismo**. Tradução de Antonio Romeiro Sapienza. 6.ed. Rio de Janeiro: McGraw-Hill, 2003.



JÚLIO, E.N.B.; BRANCO, F.A.; SILVA, V.D. Concrete-to-concrete bond strength. Influence of the roughness of the substrate surface. **Construction and Building Materials**, v.18, n.9, p. 675-681, Nov. 2004.



KATZENBACH, J.R.; SMITH, D.K. The Discipline of teams (cover story). **Harvard Business Review**, v.83, n.7/8, p.162-171, July/Aug. 2005a.

\_\_\_\_\_\_. The Discipline of teams. **Harvard Business Review**, Best of HBR 1993, 2-11, 2005b.



KRAUS, J.D. **Electromagnetics:** with applications. 5.ed. Boston: WCB/McGraw-Hill, 1999.



MEHTA, P.K.; MONTEIRO, P.J.M. **Concreto**: estrutura, propriedades e materiais. São Paulo: Pini, 1994.



NÚCLEO DE INFORMAÇÃO TECNOLÓGICA. **Manual de inteligência competitiva**. São Carlos: UFSCar, 2004.



RAMALHO, J.A. Variáveis. In: \_\_\_\_\_\_. **Clipper 5.0**: básico. São Paulo: Makron Books, 1991. Cap.4, p.67-92.



RAY, W.F.; HEWSON, C.R. High performance rogowski current transducers. In: CONFERENCE RECORD OF THE 2000 IEEE INDUSTRY APPLICATIONS, 2000, Rome. **Proceedings…** New York, IEEE, 2000. p.3083-3090.



STINE, K.J. Brewster angle microscopy: techniques. In: STEED, J.W.; GALE, P.A. (Ed.). **Supramolecular chemistry**: from molecules to nanomaterials. New York: John Wiley, 2012. p.58-63.



YAO, C. et al. Contactless measurement of lightning current using self-integrating B-dot probe. **IEEE Transactions on Dielectrics and Electrical Insulation**, v.18, n.4, p.1323-1327, Aug. 2011. Disponível em: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=5976134&queryText%3Dcontactless+measurement+of+lightning. Acesso em: 12 Jan. 2015.

YAO, C. et al. A Novel lightning current monitoring system based on the differential-integral loop. **IEEE Transactions on Dielectrics and Electrical Insulation**, v.20, n.4, p.1247-1255, Aug. 2013. Disponível em: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=6571441&queryText%3Da+novel+lightning+current+monitoring. Acesso em: 12 Jan. 2015.

**GLOSSÁRIO**

Elemento opcional que consiste em uma lista em ordem alfabética das palavras ou expressões técnicas ou pouco conhecidas utilizadas no texto.

**Apêndice A – Digitar o título do apêndice A**

Elemento opcional, texto ou documento elaborado pelo autor, que serve de fundamentação, comprovação e ilustração.

**Apêndice B – Digitar o título do apêndice B**

**ANEXO A – Digitar o título do anexo A**

Elemento opcional, texto ou documento não elaborado pelo autor, que serve de fundamentação, comprovação e ilustração.

**ÍNDICE**

Elemento opcional que consiste em uma lista de autor, título ou assunto em ordem alfabética ou sistemática (por classes, numérica ou cronológica) que remete para as informações contidas no texto.