Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Ciência da Computação

Aplicações de software desenvolvidas no contexto da Inteligência Artificial (IA), Machine Learning e Big Data e o direito dos cidadãos de acordo com a Lei Geral de Proteção de Dados (LGPD)

Iuri Sousa Vieira

Monografia apresentada como requisito parcial
para conclusão do Curso de Engenharia da Computação

Orientadora
Prof.a Dr.a Edna Dias Canedo

Brasília
2021

Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Ciência da Computação

Aplicações de software desenvolvidas no contexto da Inteligência Artificial (IA), Machine Learning e Big Data e o direito dos cidadãos de acordo com a Lei Geral de Proteção de Dados (LGPD)

Iuri Sousa Vieira

Monografia apresentada como requisito parcial
para conclusão do Curso de Engenharia da Computação

Prof.a Dr.a Edna Dias Canedo (Orientadora)
CIC/UnB

Prof. Dr. Vinicius Ruela Pereira Borges
CIC/UnB

Prof.a Dr.a Germana Menezes da Nobrega
CIC/UnB

Prof. Dr. João José Costa Gondim
Coordenador do Curso de Engenharia da Computação

Brasília, 09 de novembro de 2021

Dedicatória

Dedico este trabalho a minha família, que sempre me incentivou e apoiou eu todas as decisões, aos meus amigos que conheci nesse tempo, e também a professora Edna Dias Canedo que me ajudou a elaborar este trabalho e assim contribuir a terminar um ciclo da minha vida.

Agradecimentos

Agradeço a Deus por, ao longo deste processo complicado e desgastante, me ter feito ver o caminho, nos momentos em que pensei em desistir.

Agradeço também aos meus pais, irmãos e familiares que sempre me apoiaram ou me ajudaram de algum modo a chegar até aqui.

Agradeço ainda aos meus amigos "jogadores", que apesar das dificuldades, conseguimos com que todos chegassem ao fim desse ciclo, um apoiando o outro e ajudando no que fosse possível.

Por fim, agradeço aos amigos que fiz nessas árduas e demoradas viagens diárias a faculdade, pois vocês fizeram isso se tornar menos cansativo.

Resumo

Cada vez mais as empresas estão interessadas nos dados de seus usuários, e com isso, a proteção de dados se tornou uma pauta importante em todo o ecossistema tecnológico. Com tamanha evolução, a necessidade de haver regulações que limitam e controlam o acesso, o armazenamento, o processamento e compartilhamento destes dados. À vista disso, foi criado o Regulamento Geral Europeu de Proteção de Dados (GDPR), e o tomando por base, foi criado posteriormente a Lei Geral de Proteção de Dados Pessoais (LGPD). Para a adequação das organizações, surgiu então na literatura métodos e processos que as auxiliam nessa implementação. Este trabalho tem o objetivo de encontrar e investigar alguns destes métodos/processos e verificar sua conformidade com a LGPD, e em seguida, apresentar um checklist que respeite e atenda as normas da regulação brasileira. Para alcançar a meta, foi então proposto um checklist embasado a partir destes métodos/processos encontrados na literatura, sempre com a finalidade de alcançar a ética, a privacidade e a anuência a LGPD. Para a validação do checklist foi elaborado um formulário a ser respondido pelos profissionais da área de Tecnologia da Informação (TI) com o intuito de coletar a importância de cada passo do checklist, e as suas possíveis limitações. Nos resultados obtidos, encontramos respostas satisfatórias que sanaram as dúvidas levantadas para o seu processo de implementação, percebendo uma deficiência dos colaboradores em relação a importância da transparência de comunicar todos os passos ao usuário afetado ou interessado, o que quando feito garantirá a integridade do sistema. Devido a grande maioria das respostas serem de desenvolvedores foi possível perceber que existe uma grande preocupação em relação aos processos de armazenamento e anonimização dos dados, sendo a etapa em que esses profissionais mais participam, e isso é um ponto positivo pois significa que eles estão preocupados com os princípios impostos pela lei e querem estar em conformidade com as regras e reconhecem a sua importância.

Palavras-chave: Ética, Privacidade, LGPD, GDPR, Inteligência Artificial, Machine Learning, Big Data

Abstract

Companies are increasingly providing their users' data, and as a result, data protection has become an important agenda across the entire technological ecosystem. With such evolution, the need for regulations that limit and control the access, storage, processing and sharing of these data. In view of this, the European General Data Protection Regulation (GDPR) was created, taking it as base, the General Personal Data Protection Law (LGPD) was subsequently created. For the adequacy of the associations, methods and processes that help in this implementation appeared in the literature. This work aims to counter and investigate some of these methods/processes and verify their compliance with LGPD, and then present a checklist that respects and meets the standards of Brazilian regulations. To achieve the goal, a checklist based on methods/processes found in the literature was then proposed, always with the significance of achieving ethics, privacy and compliance with the LGPD. For the validation of the checklist, a form was prepared to be answered by professionals in the Information Technology (IT) area, in order to collect the importance of each step of the checklist, and its possible limitations. In the results obtained, satisfactory answers were found that resolved the doubts raised for its implementation process, realizing a deficiency of employees in relation to the importance of transparency in communicating all the steps to the affected or interested user, which when done will guarantee the integrity of the system. The vast majority of responses are from developers, it was possible to see that there is a great concern in relation to data storage and anonymization processes, being the stage in which these professionals participate the most, and this is a positive point because it means that they are concerned with the imposed principles by law and want to comply with the rules and recognize their importance.

Keywords: Ethics, Privacy, LGPD, GDPR, Artificial Inteligencie, Machine Learning, Big Data

Sumário

Lista de Figuras

Lista de Tabelas

Capítulo 1

Introdução

A proteção de dados pessoais se tornou uma preocupação sem precedentes, devido ao que se pode nominar de sociedade tecnológica, e notório que depois da introdução do uso da tecnologia, da informática, e da ampla digitalização que já se tornou indispensável, afetando diretamente a vida social, política, cultural e econômica no Mundo [1].

Com essa evolução crescente inevitável, surgem dúvidas em relação a privacidade dos dados dos seus usuários e o que as empresas podem ou não ter acesso. No Brasil, foi criada a Lei nº 13.709/2018 chamada de Lei Geral de Proteção de Dados (LGPD) [6], que entrou em vigor no dia 18 de Agosto de 2018, inspirada na lei de privacidade de dados Europeia, General Data Protection Regulation (GDPR) [7], cujo objetivo é deixar claro os parâmetros a serem implicados quando se trata de dados pessoais e até quando isso ultrapassa a sua privacidade.

De acordo com o Artigo 5º da LGPD deve-se considerar os itens do I ao XIX, quando é exposto o que significa cada tipo de dado, o que pode ser feito com os dados pessoais, quem será o responsável pelos dados pessoais armazenados, e quem poderá visualizar, utilizar, e/ou compartilhar os dados com o consentimento dos donos dos dados. Além disso, a LGPD impõe medidas para o não cumprimento das regras, explicitadas no Artigo 52, sendo a mais leve uma advertência e estabelecimento de um prazo para se adequar, e nos casos mais graves uma multa simples, de até 2% (dois por cento) do faturamento da pessoa jurídica de direito privado limitada em R$ 50.000.000,00 (cinquenta milhões de reais) por infração, ou até divulgação da infração cometida pela empresa [6].

Existe uma grande preocupação no desenvolvimento e no uso de aplicações no contexto da Inteligência Artificial (IA), sabendo que um sistema de IA é um software projetado por humanos que funciona em uma dimensão física ou digital adquirindo e interpretando os dados coletados, infere conhecimentos ou processa as informações geradas por esses dados e decide realizar a melhor ação para chegar a um determinado objetivo [8]. Portanto, existem diversas diretrizes éticas que caracterizam a privacidade de dados e a proteção

de dados como as chaves de suas recomendações, ainda mais com a grande abundância de dados que uma aplicação de IA utiliza, é extremamente importante que a privacidade individual não seja exposta como um resultado [9].

A General Data Protection Regulation (GDPR) [7] procura acompanhar as mudanças tecnológicas e socioeconômicas, proteger os direitos básicos das pessoas e controlar seus dados. A GDPR define direitos importantes para os usuários em relação a qualquer processamento de seus dados pessoais também como obrigações dos processadores que irão moldar a maneira como as aplicações no contexto de IA será desenvolvida e aplicada [10].

A IA não é mencionada explicitamente na GDPR, mas muitas disposições estão relacionadas à ela. Na verdade, algumas delas são desafiadas por novas formas de processamento de dados pessoais apoiadas pela IA, são elas disposições sobre o âmbito de aplicação, os fundamentos jurídicos, a proteção de dados princípios e tomada de decisão automatizada [11] [10]. Entretanto, a LGPD, não aborda o processo automatizado de utilização desses dados, apenas reitera o direito do titular dos dados de solicitar a revisão das decisões tomadas exclusivamente com base no processamento automático de dados pessoais que afetem os seus interesses, incluindo as decisões destinadas a definir o seu estatuto pessoal, profissional, de consumidor e de crédito ou personalidade [6] [12].

Big Data são dados muito volumosos para serem tratados e analisados por protocolos de banco de dados tradicionais, e pra conseguir um resultado satisfatório a partir destes dados escolher um jeito alternativo para o processamento, termo em bastante evidência atualmente, seus dados são gerados em quase todas as atividades humanas, devido ao uso crescente de dispositivos eletrônicos [13].

Os dispositivos eletrônicos usados deixam rastros de dados onde quer que vão. As informações de localização geradas por um smartphone por exemplo, podem ser capturadas pela operadora móvel e, por fim, capturadas pelo aplicativo instalado no aparelho, para que seja traçado um mapa do deslocamento do usuário. A utilização das informações fornecidas pelo usuário está protegida pelos termos de uso assinados, até podem proteger os direitos do usuário claramente estipulados no conteúdo dos documentos alcançados entre as partes, mas isso nunca pode resolver os problemas relacionados à privacidade dos dados [14].

Em caso de o termo assinado pelo usuário não ter vinculo com uma organização que tenha representação em território nacional, o que poderia lhe garantir que seus dados sejam usados indevidamente? Como regular esse acesso aos dados que sejam efetivamente usados por essa organização e que não estão relacionados diretamente ao termo?[14] E dúvidas criadas por Guilherme e Américo [14] em 2012, vieram a ser respondidas e regularizadas somente em 2018 com a LGPD [6] e a GDPR [7], que em especial na LGPD, finalmente

dispõe sobre o tratamento de de dados pessoais, seja de pessoa física ou jurídica, com o intuito de proteger os direitos fundamentais de liberdade e privacidade.

O Machine Learning foi introduzido em vários contextos, seja prevenção de fraudes, diagnósticos médicos, e até o desenvolvimento de carros autônomos. E novas tecnologias estão cada vez mais acessíveis aos controladores de dados, com grandes nuvens computacionais de grandes empresas, como Amazon, IBM, Google e Microsoft, que oferecem um baixo custo, escalonabilidade, serviços e ferramentas para a aprendizagem de máquina na nuvem, com o foco em mineração de dados e outros tipos de ánalise de dados [15].

A segurança do Machine Learning não é um assunto novo, mas nos últimos anos tem atraído uma grande atenção [16] [17]. Existe uma grande quantidade de trabalhos de pesquisa relacionados a segurança de algoritmos de deep learning desde que Szegedy [17] destacou varias ameaças nesses algoritmos [18]. Motivados por esse e por vários outros problemas, a GDPR [7] estendeu a proteção contra decisões tomadas por processamento automatizado para proteger os titulares dos dados. Todos os princípios de proteção de dados se aplicam a tal processamento, mas talvez o mais importante seja o que estipula que este processamento deve ser legal, justo e transparente. Embora possa parecer simples, a aplicação prática a cada elemento do machine learning será um desafio difícil [15]. A LGPD define a proteção contra processos automatizados como no GDPR, e não a algoritmos específicos [6].

Como deve-se então, implementar tais medidas e como explicar ao usuário o que, para que, por quem, e quantas vezes foram ou serão usados os seus dados, afim do usuário ter o controle de sua privacidade. Mas como fazer isso em uma área tão automatizada [11] e que já possui aplicações que nem sequer se preocupavam, antes da LGPD, com o tratamento destes dados?

1.1 Problema de pesquisa

É necessário adaptar as organizações sempre verificando se a solução proposta no contexto da Inteligência Artificial (IA), Machine Learning e Big Data podem infringir os direitos dos cidadãos, instituídos na Lei Geral de proteção de dados (LGPD), pois o não cumprimento dos princípios da LGPD pode acarretar em punições severas expressas na Lei nº 13.709/2018, como por exemplo: multa, publicização das infrações cometidas, bloqueio do banco de dados, entre outros [6].

Em razão disso, varias entidades, sejam elas parlamentares ou não, e pesquisadores da área, elaboraram métodos e soluções para a adequação das organizações e softwares a LGPD [6] ou a GDPR [7], surgindo processos de desenvolvimento como o ECCOLA [5] desenvolvido pela Comissão Européia, metodologias como a BEST [19], e checklists

como o proposto por Marinho [20]. Assim, é preciso investigar se as soluções propostas e desenvolvidas no contexto de Inteligência Artificial, Machine Learning e Big Data estão em conformidade com os princípios da LGPD, e subsequentemente avaliar se a implementação dessas soluções na prática é viável ou irá impactar fortemente no processo de desenvolvimento de software.

1.2 Justificativa

O não cumprimento da Lei Geral de Proteção de Dados (LGPD) pode acarretar em diversos problemas para a empresa, sendo que no caso mais grave uma multa de até R$ 50.000.000,00 (cinquenta milhões de reais) por infração ou 2% (dois por cento) do faturamento da pessoa jurídica de direito privado, suspensão parcial ou total do banco de dados e suspensão do tratamento de dados a que se refere a infração, e podendo até ocorrer a divulgação das infrações cometidas pela empresa [6].

De acordo com Sarlet [1], a facilidade de acesso aos dados pessoais aumenta a velocidade de acesso, transmissão e cruzamento de tais dados e aumenta a possibilidade de afetar os direitos básicos dos usuários ao compreender e controlar as informações sobre seus dados pessoais, privadas e sociais, por isso deve se aplicar a melhor solução possível em sua implementação, o que consequentemente irá impactar no desenvolvimento de software e nas tecnologias que utilizam dados como o pilar de suas finalidades, e sempre visar o bem maior que é a privacidade dos titulares dos dados.

1.3 Objetivos

1.3.1 Objetivo Geral

O objetivo geral deste trabalho é investigar na literatura se os métodos já existentes no contexto da Inteligência Artificial, Machine Learning e Big Data garantem a ética e os direitos dos cidadãos conforme os princípios da Lei Geral de Proteção de Dados (LGPD) [6], e propor um checklist que se encaixe nos requisitos éticos e de conformidade com a LGPD para o desenvolvimento destes sistemas.

1.3.2 Objetivos Específicos

Visando atingir o objetivo geral deste trabalho, foram definidos os seguintes objetivos específicos:

- Realizar uma revisão de literatura em relação à Lei Geral de Proteção de Dados (LGPD);
- Realizar uma revisão de literatura em relação à General Data Protection Regulation (GDPR);
- Explicitar as diferenças e similaridades entre a LGPD e a GDPR;
- Realizar uma revisão de literatura em relação à Inteligência Artificial, ao Big Data, e ao Machine Learning;
- Investigar se os métodos, frameworks e/ou processos escolhidos garantem a ética no desenvolvimento das aplicações no contexto de Inteligência Artificial, Machine Learning e Big Data;
- Investigar se os métodos, frameworks e/ou processos apresentados na literatura estão em conformidade com a LGPD.
- Desenvolver um checklist que tenha a conformidade e a ética como seu objetivo.

1.4 Resultados Esperados

Este trabalho irá apresentar para as equipes de desenvolvimento de software se os métodos, frameworks e/ou processos identificados na literatura garantem a ética no desenvolvimento das aplicações no contexto de Inteligência Artificial, Machine Learning e Big Data. Além disso, irá verificar quais desses métodos estão em conformidade com a LGPD a partir de métodos encontrados na literatura e se são capazes de apoiar as equipes durante as fases de desenvolvimento de software. Por fim, propor o checklist desenvolvido para implementar sistemas que respeitam a ética e a LGPD.

1.5 Metodologia de Pesquisa

Neste trabalho será usado o método científico de abordagem dedutivo, e de procedimento os métodos comparativo e monográfico, conforme proposto por Prodanov [21]. Também seria aplicado um questionário para avaliar o checklist proposto. A metodologia de pesquisa utilizada segue as etapas descritas na Figura 1.1.

Em relação a natureza da pesquisa, a metodologia escolhida foi a Pesquisa Aplicada, conforme apresentado na Figura 1.1, pois o objetivo deste trabalho é fornecer um corpo de conhecimento contendo diretrizes para as equipes interessadas em adaptar sua tecnologia à LGPD [21] [6].

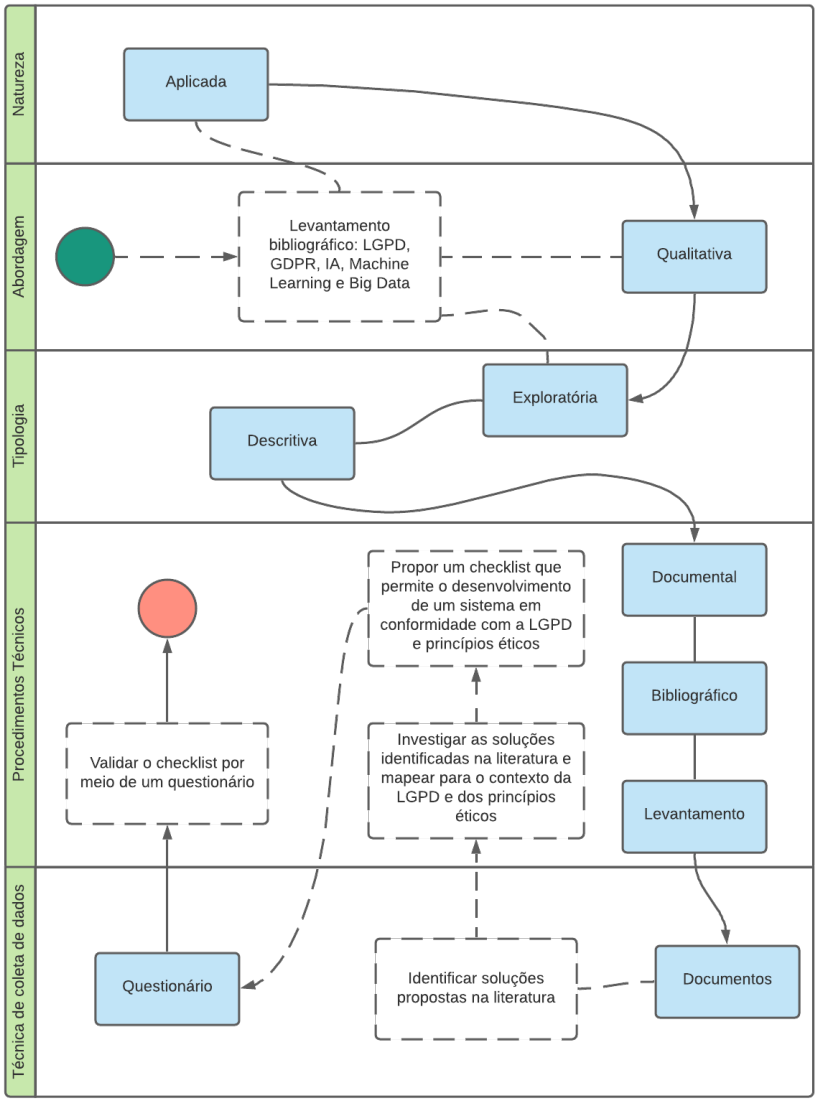

Figura 1.1: Metodologia de Pesquisa Adotada

A abordagem é qualitativa (Figura 1.1), o ambiente natural é a fonte direta de coleta de dados, e os pesquisadores são a ferramenta principal. Essa pesquisa é descritiva. Os pesquisadores tendem a analisar seus dados em geral, e o processo e seu significado são o foco principal deste método [21].

Na tipologia, a pesquisa será exploratória [21], conforme apresentado na Figura 1.1, e tem como finalidade fornecer mais informações sobre a implementação da LGPD [6] em softwares no contexto de IA, Machine Learning e Big Data, estabelecer metas e propor ou identificar um novo método. Sendo também descritivo porque inclui uma pesquisa para classificar e interpretar as opiniões das pessoas sobre as diretrizes propostas. Neste trabalho, assumirá a forma de estudos bibliográficos e será conduzido por:1. Levantamento bibliográfico; 2. Identificar exemplos que incentivem a compreensão; 3. Propor um

checklist embasado nos exemplos encontrados.

Quanto aos procedimentos técnicos (Figura 1.1), que é a forma como será explicado os dados necessários exigidos para a pesquisa, foi traçado um modelo conceitual e operacional, denominado design, que expressa a ideia de modelo, esboço e plano, para que possa ser traduzido em esboço. Os itens escolhidos para compor este trabalho foram a pesquisa bibliográfica que tem como base materiais já elaborados presentes em fontes literárias, a pesquisa documental que é a utilização de materiais que ainda não receberam tratamentos analíticos ou que podem ser reelaborados, e o levantamento afim de validar a proposta [21].

Os dados foram identificados a partir de artigos científicos, livros ou relatórios, ou seja, através de Documentos, conforme apresentado na Figura 1.1, que foram identificados e analisados. Por fim, um checklist foi proposto com o intuito de guiar o interessado na conformidade com a LGPD no desenvolvimento de aplicações no contexto de IA, Machine Learning e Big Data, e sua validação foi realizada através do questionário.

1.6 Estrutura do Trabalho

Este trabalho está organizado da seguinte maneira. No Capitulo 2 é apresentado o embasamento teórico necessário para o entendimento desse trabalho, constituído por conceitos da LGPD, GDPR, Ética no contexto de aplicações de Inteligência Artificial, Machine Learning e Big Data.

No Capítulo 3 será descrito as metodologias a serem analisadas para alcançar os objetivos definidos. Detalhando seus passos e etapas, e todo seu processo de implementação. Além da analise dos métodos escolhidos, e as discussões em relação a esses métodos.

O Capítulo 4 apresenta o checklist desenvolvido, como foi seu processo de validação, os resultados obtidos, e também são apresentadas as ameaças para a validação deste trabalho.

Finalmente, no Capítulo 5, serão apresentadas as conclusões, incluindo um resumo da contribuição deste trabalho e uma discussão de trabalhos futuros relacionados.

Capítulo 2

Embasamento Teórico

Neste Capítulo serão apresentados todos os conceitos teóricos e técnicos necessários para a contextualização e o entendimento deste trabalho.

2.1 Lei Geral de Proteção de Dados (LGPD)

Desde quando a Lei 13.709/2018 [6] foi proposta, ficou evidente a urgência e importância de abordar a privacidade e o controle de dados pessoais. Em seu Capítulo I, a LGPD fala sobre as disposições gerais da lei. Neste capítulo, estão os fundamentos e a apresentação do escopo dela, as definições de cada um dos novos termos e os princípios aplicáveis, e antes mesmo de apresentar os conceitos, a lei se preocupa em definir claramente seu objetivo e escopo de atuação [19].

Em seu Art. 2° [6] o principal fundamento abordado é a privacidade. É importante observar que a proteção de dados e a privacidade são questões diferentes. Por exemplo, se alguém postar dados em sua página pessoal em sua rede social, esses dados se tornarão públicos. No entanto, isso não significa que os dados podem ser usados arbitrariamente. Qualquer pessoa que o utilize deve respeitar os direitos do titular dos dados especificados na LGPD. Portanto, tais dados não pertencem à proteção constitucional da privacidade, mas sim ao âmbito da proteção de dados [19].

2.1.1 Caracterização dos tipos de dados

Para a lei, dado pessoal é uma "informação relacionada à pessoa natural identificada ou identificável", ou seja, dados como nome, endereço, sexo, RG e CPF [19]. Portanto, dado pessoal é toda e qualquer informação, seja ela direta ou indireta, que possa permitir a identificação de uma pessoa [1].

A lei explicita ainda mais dois tipos de dados, que são os dados pessoais sensíveis e os dados anonimizados. Sendo os dados pessoais sensíveis um dado pessoal sobre origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou a organização de caráter religioso, filosófico ou político, dado referente à saúde ou à vida sexual, dado genético ou biométrico, quando vinculado a uma pessoa natural, e os dados anonimizados um dado relativo ao titular que não possa ser identificado, considerando a utilização de meios técnicos razoáveis e disponíveis na ocasião de seu tratamento, onde quando os dados passam por anonimização eles não são considerados mais dados pessoais [6].

Diante do cuidado com o tema, foi estabelecido como regra geral (Art. 1º [6]) que quem processa dados públicos ou privados (sejam dados naturais ou jurídicos), incluindo atividades realizadas em meios digitais, deve ter uma base jurídica que justifique os dados pessoais que trata. Portanto, o processamento realizado não é um pressuposto exclusivo. Deve cumprir pelo menos um pressuposto legal para ser considerado legítimo e lícito, podendo mesmo ser acumulado e agregado como na GDPR [7]. Estes fundamentos são de forma comum e diversa, e devem ser detalhados e ajustados especificamente pela Agência Nacional de Proteção de Dados (ANPD), pelo Legislativo e pelo Judiciário [1].

2.1.2 O tratamento de dados pessoais

Em relação ao Art. 5° da Lei 13.709/2018 LGPD [6], é importante citar alguns conceitos relevantes para o entendimento deste trabalho. São eles:

- **(Art. 5°, V) Titular:** pessoa natural a quem se referem os dados pessoais que são objeto de tratamento;
- **(Art. 5°, VI) Controlador:** pessoa natural ou jurídica, de direito público ou privado, a quem competem as decisões referentes ao tratamento de dados pessoais;
- **(Art. 5°, VII) Operador:** pessoa natural ou jurídica, de direito público ou privado, que realiza o tratamento de dados pessoais em nome do controlador;
- **(Art. 5°, VIII) Encarregado:** pessoa indicada pelo controlador e operador para atuar como canal de comunicação entre o controlador, os titulares dos dados e a Autoridade Nacional de Proteção de Dados (ANPD);
- **(Art. 5°, IX) Agentes de tratamento:** o controlador e o operador;
- **(Art. 5°, XIV) Eliminação:** exclusão de dado ou de conjunto de dados armazenados em banco de dados, independentemente do procedimento empregado;
- **(Art. 5°, XVII) Relatório de impacto à proteção de dados pessoais:** documentação do controlador que contém a descrição dos processos de tratamento de

dados pessoais que podem gerar riscos às liberdades civis e aos direitos fundamentais, bem como medidas, salvaguardas e mecanismos de mitigação de risco.

A LGPD [6] fornece notas específicas sobre o responsável por cada atividade em suas regras. Portanto, esta é uma etapa necessária para que todas as organizações definam os participantes envolvidos em cada processo. De forma geral, é necessário ter bem delineadas as atividades do: controlador, operador e do encarregado [22].

Esses agentes são essenciais para o desenvolvimento de todas as atividades que envolvem o tratamento de dados pessoais, e suas funções e ações devem ser definidas no início do projeto. Deve-se lembrar que mesmo que cada agente tenha atividades próprias, em caso de incidente no processamento destes dados, todos responderão em unidade. Portanto, é muito importante adotar uma atitude transparente e boas práticas comuns a todos esses atores [22]. É dever então, dos agentes e da empresa, entender que a LGPD [6] traz direitos específicos em relação aos titulares dos dados e ao tratamento de dados pessoais realizado com suas informações [22], sendo eles: Acesso; Alteração; Eliminação; Revogação de consentimento; Não discriminação no uso dos dados; Revisão de decisões automatizadas.

Isso significa que a organização deve estar preparada para garantir o exercício dos direitos do titular dos dados durante todo o processo de execução do processamento dos dados. Portanto, é necessário verificar se todos os procedimentos e processos podem atender aos direitos e requisitos dos titulares de dados pessoais. Caso contrário, a empresa pode enfrentar penalidades severas [22].

É importante destacar que, tratando-se de um pedido, não há informação LGPD [6] que indique o prazo que deve ser entregue ao titular - a lei utiliza o termo "prazo razoável". No entanto, a General Data Protection Regulation (GDPR) [7] define o tempo de resposta em no máximo 72 horas; se sua empresa deseja seguir os padrões internacionais, é interessante adotar os padrões ditados pelo GDPR [22].

Em seu Art. 5° [6], a LGPD considera Tratamento de Dados: qualquer operação realizada em dados pessoais, como a coleta, produção, recepção, classificação, utilização, acesso, reprodução, transmissão, distribuição, processamento, arquivamento, armazenamento, eliminação, avaliação ou controle da informação, modificação, comunicação, transferência, difusão ou extração. **Os 10 princípios que devem ser levados em consideração no tratamento de dados pessoais (Art. 6°, LGPD [6]) são:**

1. **Finalidade:** tratamento para fins legais, específicos, claros e informados, não sendo possível realizar o processamento posterior de forma inconsistente com esses fins;
2. **Adequação:** estar de acordo com os antecedentes do tratamento, a compatibilidade do tratamento com a finalidade de informar o titular;

3. **Necessidade:** limitar o processamento ao mínimo necessário para atingir o seu objetivo, e abranger os dados relevantes, proporcionais à finalidade do processamento de dados e não excessivos;
4. **Livre acesso:** Assegurar aos titulares uma consulta gratuita e conveniente sobre a forma e duração do tratamento e a integridade dos seus dados pessoais;
5. **Qualidade dos dados:** conforme necessário e de forma a atingir o seu objetivo de processamento, para garantir a exatidão, clareza, relevância e atualização dos dados para o titular;
6. **Transparência:** garantir ao titular informações claras, precisas e de fácil acesso sobre a forma de tratamento e o desempenho do agente terapêutico correspondente, e cumprir os segredos comerciais e industriais;
7. **Segurança:** utilizar medidas técnicas e de gestão para proteger os dados pessoais para prevenir o acesso não autorizado e a destruição, perda, alteração, comunicação ou disseminação acidental ou ilegal;
8. **Prevenção:** tomar medidas para prevenir danos causados pelo tratamento de dados pessoais; Impossível tratar para fins discriminatórios ilegais ou abusivos;
9. **Não discriminação:** impossível tratar para fins discriminatórios ilegais ou abusivos;
10. **Responsabilização e prestação de contas:** demonstrar as medidas eficazes tomadas, que podem comprovar o cumprimento das normas de proteção de dados pessoais, e mesmo comprovar a eficácia dessas medidas.

Dessa forma, pode-se perceber o impacto do GDPR nos requisitos de processamento de dados pessoais na criação de documentos brasileiros. A LGPD enfatiza que o tratamento de dados pessoais deve obedecer ao princípio da boa fé, e ter finalidade, restrição, responsabilidade, tecnologia de segurança e medidas que garantam a segurança, a transparência e a possibilidade de negociação com o titular [23].

O artigo 5° do GDPR [7] sobre os princípios e restrições que devem ser seguidos no tratamento de dados pessoais, deve seguir de maneira que a licitude, a lealdade, a transparência, a finalidade, o limite, a proporcionalidade, a exatidão, a integridade e a confidencialidade são algumas das principais características que o tratamento de dados deve seguir. Os artigos 12, 13 e 14 do regulamento europeu indicam mais claramente os requisitos, restrições e regras para melhorar a transparência dos assuntos e atividades de informação [23].

O Capítulo II da LGPD [6] Requisitos necessários dedicados ao processamento de dados, especialmente aqueles relacionados ao consentimento. Embora a obtenção de consentimento não seja a única suposição sob a qual os dados podem ser obtidos e processados, embora esta seja a suposição mais comum. Os interesses legítimos também são base legal, ou seja, se a organização (ou terceiro) precisar fornecer produtos / serviços, ou mesmo melhorá-los, ou mesmo inovar, a premissa abrangerá os interesses [19]. Pensando então, no lado do usuário, aqueles à quem o dado se diz respeito, a LGPD deixa claro em seu Artigo 7° [6], que o tratamento de tais dados somente poderá ser realizado nas seguintes hipóteses:

1. mediante o fornecimento de consentimento pelo titular;
2. para o cumprimento de obrigação legal ou regulatória pelo controlador;
3. pela administração pública, para o tratamento e uso compartilhado de dados necessários à execução de políticas públicas previstas em leis e regulamentos ou respaldadas em contratos, convênios ou instrumentos congêneres, observadas as disposições do Capítulo IV desta Lei;
4. para a realização de estudos por órgão de pesquisa, garantida, sempre que possível, a anonimização dos dados pessoais;
5. quando necessário para a execução de contrato ou de procedimentos preliminares relacionados a contrato do qual seja parte o titular, a pedido do titular dos dados;
6. para o exercício regular de direitos em processo judicial, administrativo ou arbitral, esse último nos termos da Lei nº 9.307 (Lei de Arbitragem) [24];
7. para a proteção da vida ou da incolumidade física do titular ou de terceiro;
8. para a tutela da saúde, exclusivamente, em procedimento realizado por profissionais de saúde, serviços de saúde ou autoridade sanitária;
9. quando necessário para atender aos interesses legítimos do controlador ou de terceiro, exceto no caso de prevalecerem direitos e liberdades fundamentais do titular que exijam a proteção dos dados pessoais;
10. para a proteção do crédito, inclusive quanto ao disposto na legislação pertinente.

Um ponto de atenção bastante relevante é retratado pelo Art. 8° [6], que é a questão do consentimento. Com o passar dos anos, devido à sensibilidade e vulnerabilidade das informações pessoais proporcionadas pelo desenvolvimento tecnológico, a necessidade de coleta de dados, principalmente para obtenção de consentimento em ambiente virtual,

tem se tornado cada vez mais importante. Nesse sentido, certifique-se de que as pessoas/usuários estejam cientes de que devem concordar com a utilização dos dados e ter o direito de saber a finalidade da coleta e o direito de acessar seu conteúdo a qualquer momento, portanto, garantir a liberdade e a privacidade é muito importante [23].

O objetivo do consentimento nada mais é do que indicar que deve haver um processo de tomada de decisão, sem a cooperação da outra parte que tratam dos dados, o detentor dos dados sozinho não pode alcançá-lo. Deste ponto de vista obrigatório, os agentes de processamento de dados (especialmente os controladores) devem assumir uma série de responsabilidades [1]. Portanto, mesmo que o titular divulgue claramente suas dados, o responsável pelo tratamento e o operador não podem ficar isentos das suas responsabilidades, nomeadamente no que se refere à forma e duração do tratamento efetuado pelo titular com base nos seus dados, bem como à eventual partilha do controlador e a eventual quota de acesso do operador à informação [19].

Portanto, o direito de saber, deve fornecer aos cidadãos os elementos necessários para iniciar o processo de tomada de decisão sobre o seu fluxo de dados. Fornecer informações claras e suficientes é a porta de entrada para ambos [1]. Fica ainda mais claro no Art. 8° [6], explicitando que o usuário deverá fornecer o consentimento para o uso de seus dados, caso seja de seu interesse, sendo ele por escrito ou por outro meio que demonstre a manifestação de vontade do titular. **Artigo 8°, LGPD [6]:**

- § 1º Caso o consentimento seja fornecido por escrito, esse deverá constar de cláusula destacada das demais cláusulas contratuais.
- § 2º Cabe ao controlador o ônus da prova de que o consentimento foi obtido em conformidade com o disposto nesta Lei.
- § 3º É vedado o tratamento de dados pessoais mediante vício de consentimento.
- § 4º O consentimento deverá referir-se a finalidades determinadas, e as autorizações genéricas para o tratamento de dados pessoais serão nulas.
- § 5º O consentimento pode ser revogado a qualquer momento mediante manifestação expressa do titular, por procedimento gratuito e facilitado, ratificados os tratamentos realizados sob amparo do consentimento anteriormente manifestado enquanto não houver requerimento de eliminação, nos termos do inciso VI do caput do Art. 18 desta Lei.
- § 6º Em caso de alteração de informação referida nos incisos I, II, III ou V do Art. 9º desta Lei, o controlador deverá informar ao titular, com destaque de forma específica do teor das alterações, podendo o titular, nos casos em que o seu consentimento é exigido, revogá-lo caso discorde da alteração.

O GDPR também pontua a importância primordial do consentimento até mesmo para a garantia da licitude do tratamento de dados a ser realizado ou em realização, como explicitam os artigos 6° e 7° [7]. No mesmo sentido, o regulamento aponta a necessidade de confirmação do consentimento, tendo em vista que o silêncio ou a omissão não é considerado uma forma de consentimento. Da mesma forma, o GDPR enfatiza a liberdade de escolha do titular, portanto, sua recusa ou revogação não lhe causará nenhum dano [23].

O artigo 9°[6] prende-se à questão da transparência das informações no processamento de dados e aponta características relacionadas ao livre acesso às informações. Nesse sentido, uma declaração clara e um acesso conveniente quanto à finalidade e forma do tratamento, sua duração e informações sobre os medicamentos utilizados para o tratamento são elementos essenciais [23].

Em seu artigo 10º [6], define que quando o tratamento de dados for baseado no legitimo interesse do controlador, somente os dados pessoais específicos necessários poderão ser tratados. Ainda, o controlador deverá adotar regras para garantir a transparência destes dados tratados e ANPD poderá solicitar a qualquer momento ao controlador um relatório de impacto a proteção dos dados pessoais.

2.1.3 Tratamento de dados pessoais sensíveis

Os dados pessoais qualificados como sensíveis são encontrado em todas as partes informacionais do ser humano. Na LGPD – assim como no GDPR –, entendeu-se que a melhor forma de proteger os titulares seria trazendo exemplos claros de dados assim considerados. Portanto, segundo o artigo 5.º, inciso II, da LGPD [6], dados sensíveis são toda e qualquer informação referente a origem racial ou étnica, convicção religiosa, opinião política e filiação a sindicato ou a organização de caráter religioso, filosófico ou político. São também considerados dados sensíveis os dados que se referem à saúde ou à vida sexual, e dados genéticos ou biométricos [1].

Deve-se ter cuidado ao se tratar dados sensíveis, no ponto de vista dos direitos e liberdades fundamentais, que em certos contextos podem ter riscos significativos aos seus titulares. Nesse sentido, é essencial então determinar se um dado é sensível ou não, ou seja, verificar o contexto em que ele será utilizado, além das relações que podem ser estabelecidas com as demais informações disponíveis, e se seu tratamento pode transformar estes dados em um instrumento de estigmatização ou discriminação. Por isso, deve-se admitir que certos dados, mesmo que não tenham a natureza de um dado sensível, sejam considerados como tal, considerando a que fim levará o seu tratamento [1].

Estes dados somente podem ser tratados com o consentimento de seus titulares, excepcionalmente quando usado por órgãos de pesquisa e saúde, desde que se responsabilizem por total segurança e que estes dados não sejam compartilhados [25].

2.1.4 Direitos do titular dos dados

Em seu Capítulo III, a LGPD foca nos direitos que toda pessoa natural tem em relação a seus dados pessoais, garantindo os direitos fundamentais de liberdade, intimidade e privacidade [6]. A jurisdição brasileira, no entanto, não dá toda a atenção em vista das decisões automatizadas aplicadas em dados. Antes mesmo da aprovação da LGPD, já existiam normas que protegiam os individuos, entretanto, elas tratavam de casos e grupos específicos [1].

Por isso, houve certas divergências durante seu processo de aprovação, entre outras partes, o direito a explicação. Existiram pelo menos quatro versões para o artigo 20° [6] que especificamente trata de decisões automatizadas na LGPD. O texto original aprovado em 2018 dizia que o titular dos dados teria direito a reivindicar a revisão, por pessoa natural, das decisões automatizadas em cima de seus dados [1].

Em dezembro de 2018, o Presidente da República expediu uma medida provisória (MP 869/2018) que, além de outras, suprime a referencia à pessoa natural na revisão de processos automatizados [1]. No processo de aprovação da medida provisória, o Congresso acrescentou uma emenda para incluir o § 3°, que esclarecia novamente que a revisão era exclusivamente a pessoa natural. Mas em sua sanção presidencial, o Presidente vetou o § 3° como pode ser visto no infográfico na Figura 2.1.

Texto Original 2018	MP 869	Texto Aprovado Congresso	Aprovado com vetos
Art. 20. O titular dos dados tem direito a solicitar revisão, por **pessoa natural**, de decisões tomadas por processos automatizados de dados pessoais que afetem seus interesses.	Art. 20. O titular dos dados tem direito a solicitar revisão de decisões tomadas por processos automatizados de dados pessoais que afetem seus interesses.	**Adição:** § 3° A revisão de que trata o caput deste artigo deverá ser realizada por pessoa natural, conforme previsto em regulamentação da autoridade nacional, que levará em consideração a natureza e o porte da entidade ou o volume de operações de tratamento de dados	**VETADO** o § 3°

Figura 2.1: Infográfico referente ao § 3° do Capitulo III da LGPD [1]

Houveram novos questionamentos, tanto para a inclusão de uma definição para "decisão automatizada", quanto para a derrubada do veto ao ao § 3.º do artigo 20 da LGPD [6]. Tais reviravoltas levam a uma aparente incerteza quanto aos direitos dos titulares

ante ao uso de decisões automatizadas em seus dados. Apesar disso, existe um regime que protege o indivíduo e lhe garante uma explicação, e caso aconteça tratamento de dados pessoais para uma decisão automatizada, a LGPD será acionada [1].

Resta entender então (i) se o sistema em questão utiliza-se de decisões automatizadas e quais; (ii) se esclarece o direito a explicação; e (iii) se garante os direitos do cidadão para o caso de seu uso [1]. Então, neste caso, o LGPD requer que os controladores e operadores gerenciem estritamente todas as coisas feitas com os dados. Exige também que o titular receba, a qualquer momento, um extrato contendo detalhes dos dados e do seu processamento [25].

Os direitos dos usuários incluem também: confirmar se existe um método de processamento de consentimento, revogar seu consentimento para acessar os dados e corrigir, anonimizar, bloquear ou eliminar adequadamente o conteúdo com o qual eles não concordam, a portabilidade de seu terceiro designado, e informações sobre possíveis compartilhamentos [25].

Da mesma forma, a liberdade de retirar o consentimento e solicitar o apagamento de dados é reiterada, refletindo a liberdade de escolha do indivíduo, portanto, a retirada deve ser expressa como consentimento. Novamente, as disposições legais reiteram que os dados anônimos não têm direito ao mesmo tratamento que os dados pessoais [23].

2.1.5 Os agentes do tratamento de dados pessoais

Para que haja um controle e a transparência destas ações, e respeito aos propósitos do tratamento de dados, são responsabilidades do controlador e do operador documentar as operações e etapas realizadas durante todo o processo de tratamento de dados pessoais [23]. O Capítulo VI [6] da LGPD é dedicado a apresentar os deveres e responsabilidades do Controlador, do Operador e do Encarregado [25].

É dever do operador realizar o tratamento de acordo com as orientações fornecidas pelo controlador. O controlador deve, também, indicar o encarregado pelo tratamento de dados pessoais, sua identidade e informações de contato devem então ser divulgadas publicamente. As atividades atribuídas ao encarregado são [6]:

- aceitar reclamações e comunicações dos titulares, prestar esclarecimentos e adotar providências;
- receber comunicações da autoridade nacional e adotar providências;
- orientar os funcionários e os contratados da entidade a respeito das práticas a serem tomadas em relação à proteção de dados pessoais;

Se ocorrer o descumprimento da lei, cabe então indenização e multa [6], onde o Operador e Controlador podem ser ambos indiciados, mas lembrando que é de responsabilidade

dos dois garantir os processos e decisões tomadas estejam de acordo com a LGPD. Há a possibilidade também de regresso, ou seja, aquele que pagar a indenização ao Titular poderá cobrar ao outro. Além disso, os Titulares podem também processar coletivamente os dois [25].

É importante lembrar que vale a inversão do ônus da prova. A LGPD permite que o promotor não forneça provas, mas o réu pode fornecer provas de defesa, quando ela sabe que as alegações são credíveis e o fornecimento do titular é escasso, ou seja, quando uma das partes não tem uma boa situação financeira [25].

2.1.6 Autoridade Nacional de Proteção de Dados (ANPD)

A Autoridade Nacional de Proteção de Dados é uma parte importante para garantir a eficácia da Lei Geral de Proteção de Dados. Na verdade, o uso de agências administrativas para proteger os dados tornou-se um recurso, falando francamente, a maioria dos marcos regulatórios sobre o assunto são os mesmos [1]. E ela foi criada para trazer mais segurança e estabilidade à aplicação da LGPD, nas circunstâncias específicas do Brasil, existem disposições muito amplas na lei, dependendo da regulamentação da Autoridade, que dela depende os ajustes necessários para que a legislação esteja em conformidade com as normas sociais e jurídicas [23].

A existência de uma autoridade reguladora para supervisionar a aplicação da estrutura regulatória usada para proteger os dados pessoais. Diversos aspectos básicos da proteção de dados pessoais, como a dificuldade para os cidadãos monitorarem efetivamente o processamento de dados e seu impacto, e a necessidade de uma disciplina continuamente atualizada devido ao desenvolvimento tecnológico, que justificam o uso de informações pessoais. Hoje, tais instituições aparecem na grande maioria dos marcos regulatórios dessa área e quase sempre são um de seus pilares [1].

2.2 General Data Protection Regulamentation (GDPR)

O GDPR é um projeto para proteger os dados e as identidades dos cidadãos da UE. O projeto foi concebido em 2012 e aprovado em 2016. A União Europeia acredita que a proteção dos dados pessoais é um direito dos cidadãos da UE. Portanto, todas as empresas e organizações, independentemente da escala ou ramo de atividade, devem seguir regras rígidas para coletar, processar, compartilhar e proteger os dados pessoais . Na verdade, este regulamento aplica-se a qualquer tipo de serviço para cidadãos que chegam a um dos países da UE. Por exemplo, isso significa que se uma loja online no Brasil ou em qualquer outro país quiser enviar produtos para clientes da UE sem violar a lei, ela deve se adaptar ao GDPR [7].

2.2.1 Legalidade, Justiça e Transparência

Os três componentes deste princípio estão claramente ligados: o titular dos dados deve ser informado sobre o processamento que irá ocorrer (transparência), o processamento deve corresponder a esta descrição (justiça) e o processamento deve ser para um dos fins especificados no Regulamento (legalidade). O titular dos dados também deve ser informado da existência da operação de processamento [26].

2.2.2 Limitação

O regulamento afirma que os dados pessoais só podem ser recolhidos para fins específicos, explícitos e legítimos. Ou seja, para cumprir o princípio de limitação da finalidade, você deve definir antecipadamente para que os dados serão usados e limitar o processamento apenas ao necessário para atender a essa finalidade [7].

Avisos de privacidade, termos e condições e formulários de consentimento devem fornecer ao titular dos dados informações inequívocas sobre a extensão do processamento envolvido. Essas declarações públicas devem ser refletidas no processamento real e na documentação desse processamento [11].

Por exemplo, muitos supermercados coletam informações pessoais para que possam fornecer aos clientes ofertas direcionadas que correspondam a seus hábitos de consumo habituais. Seria uma violação deste princípio que esses supermercados entregassem estes dados a uma empresa irmã que comercializa férias, uma vez que foge ao âmbito da finalidade para a qual os dados foram recolhidos [27].

O regulamento permite algum tratamento posterior para fins de arquivo de interesse público, para fins de investigação científica ou histórica, ou para fins estatísticos [40]. As salvaguardas para o tratamento deste tipo de informação são estabelecidas no artigo 89.º e terá de examinar as opções técnicas e organizacionais para cumprir o regulamento. A pseudonimização e a criptografia, por exemplo, seriam medidas válidas, assim como restringir o acesso a tais informações com base na função e nos requisitos de um determinado conjunto de procedimentos [28].

2.2.3 Minimização dos Dados

O Regulamento estabelece que os dados pessoais que você coleta e/ou processa devem ser adequados, relevantes e limitados ao que é necessário em relação aos fins para os quais são processados. Isso significa que você não deve armazenar mais dados além do que é estritamente necessário. Afinal, é difícil perder informações que você não tem [7].

O cumprimento deste princípio de proteção de dados será facilitado pelo mapeamento de dados e pela revisão de seus procedimentos. Garantir que você sabe como os dados

são usados é fundamental para minimizar os dados que você coleta e processa, e deve ser integrado à maneira como sua organização trabalha como parte de uma abordagem de privacidade desde o projeto [29].

A minimização de dados também deve ser levada em consideração em acordos com os fornecedores e processadores de dados. Isso pode incluir retirar certos dados antes de passar as informações para processamento externo e, em seguida, reconectar os dados quando eles retornarem do processador [30].

2.2.4 Precisão

O regulamento exige que os dados pessoais sejam exatos e, quando necessário, atualizados. Além de ser uma boa prática para qualquer negócio, protege o titular dos dados de uma série de ameaças, como roubo de identidade. Ele também garante que quaisquer decisões automatizadas de criação de perfil feitas em relação ao titular dos dados usem dados precisos [26].

O regulamento visa claramente regulamentar quando, como e em que condições a definição de perfis pode ser conduzida. Se a sua organização se dedica à criação de perfis de qualquer tipo - e especialmente se houver impactos materiais para o titular dos dados - você precisa garantir que tenha processos em vigor para manter todos os dados pessoais precisos e atualizados [11].

A consequência deste princípio é o direito do titular dos dados à retificação. Isso concedeu ao titular dos dados o direito de retificação de dados pessoais inexatos e o direito de ter dados pessoais incompletos preenchidos. Você deve garantir não apenas que os dados pessoais sejam precisos, mas também que tenha um processo pelo qual os titulares dos dados possam solicitar a correção ou o preenchimento de seus dados pessoais [31].

2.2.5 Limitação de Armazenamento

O regulamento exige que os dados pessoais sejam mantidos de forma a permitir a identificação dos titulares dos dados apenas durante o tempo necessário para os fins para os quais os dados pessoais são tratados [7].

Em termos simples: se você não precisa mais dos dados, livre-se deles. Como você deve estar definindo um propósito para toda coleta de dados, deve ser bastante simples determinar quando os dados não são mais necessários. Algumas organizações, no entanto, podem precisar reter dados pessoais para fins de longo prazo com processamento intermitente e, nesses casos, a exclusão sumarizada de dados pode não ser possível [26].

A pseudonimização - divisão de dados pessoais em conjuntos que não permitem individualmente a identificação do titular dos dados - é uma solução para garantir o armazenamento de dados pessoais, mas apresenta seus próprios problemas no que diz respeito à usabilidade. Se esses dados pessoais precisarem ser processados regularmente, o tempo gasto para reverter a pseudonimização pode ser oneroso [28].

2.2.6 Integridade e Confidencialidade

Este princípio é talvez o mais importante do ponto de vista financeiro. Embora as violações dos outros princípios de proteção de dados possam ser prejudiciais aos titulares dos dados, o impacto geralmente é limitado. Violações deste princípio, no entanto, tendem a resultar em violações de dados, o que torna muito fácil para as autoridades de supervisão provar que os dados não foram mantidos de forma segura [11].

Isso exige que as organizações processem os dados pessoais de uma maneira que garanta a segurança adequada dos dados pessoais, incluindo proteção contra processamento não autorizado ou ilegal e contra perda, destruição ou dano acidental [29].

Integridade é propriedade de precisão, isso está claramente vinculado ao princípio da exatidão e é necessário pelas mesmas razões: o titular dos dados não deve ser prejudicado por informações imprecisas. Isso também inclui garantir que os dados pessoais estejam corretamente vinculados e garantir que os dados não sejam corrompidos com o tempo ou por práticas inadequadas de armazenamento [27].

2.2.7 Responsabilidade e Conformidade

A Clausula 2 do Artigo 5 é breve mas extremamente importante: o controlador deve ser responsável por, e ser capaz de demonstrar conformidade com a LGPD. Que afirma que o controlador de dados é responsável por garantir o cumprimento dos seis princípios anteriores de processamento de dados e por ser capaz de demonstrar esse cumprimento [26]. Como tal, o controlador de dados precisa garantir que os princípios de processamento de dados sejam atendidos onde quer que os dados pessoais vão: processadores externos e organizações / divisões internas devem ser exigidos por contrato. Deve haver processos adicionais incorporados aos contratos de serviço para demonstrar que os dados pessoais são processados em conformidade com esses princípios em todas as fases [30].

Deixar de garantir que seus fornecedores conheçam as requisitos dos princípios pode ter um impacto considerável. Como o GDPR responsabiliza o controlador pelo cumprimento de seus requisitos, eles certamente sofrerão o impacto das ações de fiscalização, multas e danos à reputação [7].

Uma cultura de responsabilidade deve ser alimentada de cima para baixo. É muito simples para um funcionário não ter nenhum senso de responsabilidade se os gerentes seniores e o gerente de conformidade não mostrarem o mesmo nível de dedicação. Os programas de treinamento e conscientização da equipe devem garantir que toda a equipe compreenda suas diversas funções e responsabilidades em relação à privacidade e à proteção de dados [31].

2.3 Diferenças entre a LGPD e a GDPR

O GDPR marca uma reviravolta na regulamentação de qualquer empresa ou organização sobre a utilização de dados pessoais, sendo o primeiro a proporcionar aos utilizadores amplos direitos de privacidade no mundo virtual e a exigir dos processadores de dados uma atitude responsável e segura ao manipular informações dos titulares dos dados [20].

O primeiro ponto que merece ser descrito é a aplicabilidade extraterritorial das diretrizes da lei, que cobre o tratamento de dados de indivíduos pertencentes à UE ou dados localizados na UE. Isso significa que, independentemente da fonte do controlador ou processador (que pode ser público ou privado), se ele executa processos ou fornece serviços nas condições acima, eles devem estar em conformidade com a GDPR. Outro fator importante frisar que a GDPR não fornece aos indivíduos (titulares dos dados) a propriedade dos dados, mas sim controle sob o que acontecerá com eles - como são armazenados, para que finalidade são usados e com quem são compartilhados. Além disso, em casos de vazamentos de dados, o responsável pelo tratamento dos dados deve informar o proprietário e a autoridade responsável, no prazo de 72 horas, o fato ocorrido, caso não cumpra esta obrigação será aplicada uma multa [20] [7].

Além de estabelecer um DPO, a norma européia também autoriza as funções mais importantes para a aplicação da lei - desde aconselhamento, vigilância e multas quando necessário - a Data Protection Authorities (DPA, Autoridade de Proteção de Dados, em português) de cada país membro. Assim, cada país membro da UE tem a responsabilidade de escolher o conselho de administração que constituirá a autoridade de fiscalização e especificar os poderes que lhe são conferidos, ou seja, as responsabilidades específicas de cada DPA dependem das legislações nacionais de cada estado membro. Por fim, é importante destacar que o GDPR impõe parâmetros aos países que desejam manter relações comerciais com a UE, ou seja, continuar a fazer negócios (trata-se principalmente de processamento de dados, mesmo que apenas para funcionários de empresas), é necessário que o país em questão possua uma regulamentação completa e abrangente [20] [7].

Já a lei brasileira, o ponto chave da LGPD é que é necessário entender que a privacidade está relacionada ao consentimento individual, portanto, os prestadores de serviços devem

sempre fazer os requisitos de forma clara e inequívoca para que o portador de dados possa controlar totalmente a forma como suas informações. Será usado em serviços para alcançar um melhor equilíbrio entre oferta e demanda [6].Por exemplo, em relação aos encarregados da proteção de dados, enquanto o GDPR estabelece,em seu Artigo 37, os casos específicos em que é necessário contratar um encarregado, como no caso em que o processamento dos dados é feito por um órgão ou autoridade pública, na LGPD, a redação atual do Artigo 41 leva ao entendimento de que toda organização que faz o processamento de dados precisaria contratar um encarregado da proteção de dados, embora ainda seja necessária maior clareza a respeito [32] [6].

Além disso, no que diz respeito à notificação aos indivíduos sobre violação de dados, a LGPD não fixou um prazo, determinando apenas que a notificação seja realizada em um período razoável de tempo, conforme definido pela autoridade nacional. Portanto, a Agência Nacional de Proteção de Dados determina o prazo no momento da operação [6] [32] [20].

Na Tabela 2.1 é apresentado as principais diferenças entre a GDPR [7] e a LGPD [6], para ilustrar mais claramente suas divergências.

Item da Lei	**GDPR**	**LGPD**
Registro de atividades de processamento	Não obrigatório para empresas com menos de 250 funcionários	Obrigatório para todas as empresas
Multas	Até 4% do faturamento (€ 20M)	Até 2% do faturamento (R$ 50M)
Requisição de direitos	Em até 30 dias, gratuidade opcional	Em tempo razoável, gratuita
Notificação obrigatória de incidentes	72 horas	Tempo razoável (a ser definido)
Agência reguladora	Definida	ANPD (MP PROCON, outros)
Data Protection Officer /Encarregado de Dados	Pessoa natural ou jurídica	Pessoa natural ou jurídica
Legítimo interesse	Mais restrito	Mais flexível
Dados anonimizados	Não são considerados pessoais em perfis	Podem ser considerados pessoais em perfis
Perfis comportamentais	Necessário causar impacto no titular dos dados	Sempre se considera causar impacto no titular dos dados

Transferências internacionais	Possível, com base no legítimo interesse, caso não seja frequente	Com consentimento específico, mesmo sem legítimo interesse
Dados de saúde	Não podem ser tratados mediante contrato	Podem ser tratados mediante contrato de prestação de serviço

Tabela 2.1: Comparativo entre a LGPD e a GDPR

2.4 Inteligencia Artificial

É possível definir inteligência por meio das características que ela exibe: capacidade de responder a novas situações; capacidade de resolver problemas, responder perguntas, fazer planos, entre vários outros [33]. Normalmente, pode estar relacionado ao processo de pensamento, raciocínio e comportamento. Mesmo para o processo de medir o sucesso pela lealdade ao comportamento humano, e compará-lo com o conceito de inteligência ideal denominado racionalidade para medir o processo de sucesso. Se o sistema "entende o que é certo", então é razoável "fazer a coisa certa" [34].

Algumas pesquisas acreditam que a IA ainda está longe de ser comparada aos humanos, e consequentemente, não acham que deve se preocupar com a ética nesse setor. Mas ele já se mostrou muito eficaz quando combinado com outras tecnologias como a robótica, mostrando o seu potencial no mercado ou em estudos científicos [35].

O escopo da IA é controverso: conforme as máquinas se tornam mais poderosas, as tarefas que são consideradas como requerendo "inteligência" geralmente precisam ser removidas da definição. Esse fenômeno é chamado de efeito IA. Por exemplo, o reconhecimento óptico de caracteres geralmente se tornou uma tecnologia convencional e foi excluído da Inteligência Artificial [36].

A IA era principalmente dividida entre uma IA forte e IA fraca. Na IA forte acreditava que se você tiver um computador com poder de processamento suficiente e fornecer-lhe inteligência suficiente, as pessoas podem criar um computador que pode literalmente pensar e pensar conscientemente como humanos. Por outro lado, IA fraca é apenas a visão de que o comportamento inteligente pode ser modelado e usado por computadores para resolver problemas complexos. Essa visão sustenta que o fato de um computador estar operando de maneira inteligente não prova que seja realmente inteligente no sentido humano [33].

Os sistemas de Inteligência Artificial podem não apenas armazenar e manipular dados, mas também adquirir, expressar e manipular conhecimento. O principal problema que o projetista de um sistema de Inteligência Artificial deve evitar é a aquisição, representação e manipulação do conhecimento, geralmente a estratégia de controle ou motor de raciocínio que determina os itens de conhecimento acessados, as inferências feitas e a sequência das etapas utilizadas. Na Figura 2.2, você pode ver esses problemas e as inter-relações entre os componentes do sistema de Inteligência Artificial clássico [2].

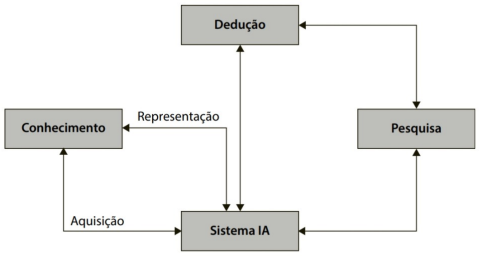

Figura 2.2: Visão conceitual de sistemas de Inteligência Artificial [2]

O teste de Turing proposto por Alan Turing (1950), que foi um dos pioneiros da Inteligência Artificial, visa fornecer uma definição satisfatória de operação inteligente. Se o interrogador humano não puder determinar se a resposta escrita veio de um pessoal ou de um computador depois de fazer algumas perguntas por escrito, o computador passará no teste. E para conseguir passar no teste, o computador precisará ter os seguintes recursos [34]:

- Processamento de linguagem natural, para que ele possa se comunicar com sucesso em linguagem natural;
- Representação do conhecimento, usada para armazenar o que você sabe ou ouve;
- Raciocínio automático para usar as informações armazenadas para esclarecer dúvidas e tirar novas conclusões;
- Aprendizado de máquina para se adaptar a novas situações e detectar e inferir padrões.
- Visão computacional para perceber objetos;
- Robótica para manipular objetos e movimentar-se.

As seis disciplinas: Processamento de linguagem natural, Representação do conhecimento, Raciocínio automático, Aprendizado de máquina, Visão computacional, e Robótica, compõem a maior parte da IA, e Turing merece crédito por projetar um teste que

permanece relevante depois de 60 anos [34]. Porém, a crescente complexidade dos problemas a serem computacionalmente resolvidos, onde o volume de dados é um dos grandes empecilhos para que se possa analisar os dados em tempo hábil e encontrar uma solução válida [3], têm levado ao desenvolvimento de ferramentas computacionais mais complexas e autônomas, mais independentes da intervenção humana para aquisição de conhecimento [37].

Em Inteligência Artificial, podemos definir quatro tecnologias, que são os principais pilares que impulsionam seu grande progresso [36] [38]:

1. Categorização: a Inteligência Artificial precisa de muitos dados relevantes para resolver problemas. A primeira etapa na construção de uma solução de Inteligência Artificial é criar uma medida de intenção de projeto que é usada para classificar problemas. Independentemente de o usuário estar tentando construir um sistema que possa, por exemplo, ajudar os médicos a diagnosticar câncer ou ajudar os administradores de TI a diagnosticar problemas de rede sem fio, os usuários precisam definir métricas para dividir o problema em partes secundárias;

2. Classificação: uma vez que o usuário tenha classificado o problema em diferentes áreas, a próxima etapa é ter um classificador para cada categoria, o que levará o usuário a conclusões significativas. Em uma rede sem fio, uma vez que o usuário conhece a categoria do problema (por exemplo, o problema antes ou depois da conexão), o usuário precisa começar a classificar a causa do problema: associação, autenticação, DHCP ou outro, com fio e fatores do dispositivo;

3. Aprendizado de máquina: depois de ter dividido em metadados específicos do domínio, pode-se então essas informações ao algoritmo do aprendizado de máquina. Existem muitos algoritmos e técnicas de aprendizado de máquina, como o aprendizado de máquina supervisionado usando redes neurais (ou seja, aprendizado profundo), que se tornou um dos métodos mais populares.

4. Filtragem colaborativa: muitas pessoas passam pela filtragem colaborativa quando selecionam filmes no Netflix ou compram produtos da Amazon, que recebem recomendações de outros filmes ou itens de que possam gostar. Além de recomendadores, a filtragem colaborativa também é usada para categorizar grandes conjuntos de dados e fazer a diferença nas soluções de IA. Aqui, toda a coleta e análise de dados serão transformadas em percepções ou ações significativas. É como um assistente virtual que pode ajudá-lo a resolver problemas complexos.

2.4.1 Ética no contexto de IA

A Inteligência Artificial terá um impacto significativo na sociedade e sobre isto não restam dúvidas. Em vez disso, o debate atual gira em torno de até que ponto o impacto é positivo ou negativo, para quem, de que forma, onde e em que escala de tempo [39]. No entanto, tem havido muitas pesquisas sobre ética em IA, destacando as principais questões éticas em sistemas de IA. Muitos princípios foram propostos e discutidos, e alguns são amplamente aceitos. Com base em extensa análise Código de Ética da AI de Jobin [40], temos que os princípios centrais são:

1. Transparência: recomenda-se a maior divulgação possível de informações. Incentivar o fornecimento de explicações "em termos não técnicos" ou que possam ser revisadas por humanos. Medidas alternativos concentradas no monitoramento, interação e mediação com as partes interessadas e o público, e na facilitação de reclamações;

2. Justiça e equidade: manifesta-se principalmente na justiça, prevenção, monitoramento ou mitigação do preconceito e da discriminação. Também tem sido citado como respeito à diversidade, inclusão e igualdade, o direito de se opor à tomada de decisões ou de obter compensação e reparação. Por fim, é enfatizado a importância do acesso justo à Inteligência Artificial, dados e benefícios;

3. Privacidade: geralmente relacionado à proteção e segurança de dados. Alguns associam privacidade com liberdade ou confiança. Os modelos de privacidade de IA são divididos em três categorias: privacidade diferenciada, privacidade de design, minimização de dados e controle de acesso e outras soluções técnicas, que exigem mais pesquisa, conscientização e métodos de supervisão

4. Não maleficência: elaboração de medidas gerais de segurança, considerando que a IA nunca deve causar danos previsíveis ou acidentais;

5. Beneficência: promover o bem-estar, a paz, o desenvolvimento, e a felicidade. Criando oportunidades socioeconômicas e prosperidade econômica;

6. Responsabilidade e prestação de contas: agir com "boa fé" e esclarecer a propriedade das responsabilidades legais no contrato ou nas soluções. Também é recomendável focar no que pode causar danos potenciais, e enfatizar a responsabilidade de relatar estes possíveis casos;

7. Confiança: estabelecer ou manter a confiança incluem educação, confiabilidade, responsabilidade, o processo de monitoramento e avaliação da integridade dos sistemas de IA ao longo do tempo, ferramentas e técnicas para garantir a conformidade com especificações e padrões;

8. Sustentabilidade: desenvolvimento e implementação de Inteligência Artificial que leve em consideração a proteção ambiental, melhore o ecossistema e a biodiversidade da Terra, contribua para uma sociedade mais justa e igualitária e que promova a paz;

9. Dignidade: respeitar, manter e até melhorar a dignidade. Se os desenvolvedores respeitarem a dignidade por meio de planos de governança ou diretrizes técnicas e metodológicas emitidas pelo governo, eles podem manter a dignidade;

10. Liberdade e autonomia: Liberdade de expressão ou informação, o direito de tomar decisões e o direito de controlar a privacidade, a liberdade, o empoderamento ou a autonomia;

11. Solidariedade: redistribuir os benefícios da IA para não ameaçar a coesão social e respeitar os grupos potencialmente desfavorecidos.

Essas etapas deixam claro que para se ter um bom software de IA, deve haver o mútuo desenvolvimento e colaboração nestes projetos por engenheiros, cientistas sociais, advogados, filósofos, eticistas, e com isso obter uma implementação transparente desde a ideologia do projeto até o seu final [41].

2.5 Machine Learning

A maioria das ferramentas utilizadas em IA é baseada em Machine Learning (ML), ou Aprendizagem de Máquina (AM) em português, como mostrado na Figura 2.3, que é uma subárea da IA e faz parte de várias tecnologias em uso atualmente [37]. No aprendizado de máquina, os recursos de entrada e saída são de acordo com suas funções no sistema, e são definidos como atributos preditivos e destinos ou atributos de destino [42]. A forma como a máquina os armazena não mudará, no último caso, todos os dados são números binários, dependendo da natureza do circuito que os compõe, mas o significado que lhes é atribuído pode variar dependendo do contexto em que se encontram [3].

Em tarefas de descrição, o algoritmo de ML não prevê o valor, mas extrai o padrão do valor previsto do conjunto de dados. Esses algoritmos não utilizam o conhecimento de "supervisores externos", eles usam um paradigma de aprendizagem não supervisionado. Suas tarefas são agrupar dados e localizar grupos de objetos semelhantes no conjunto de dados. Além de encontrar regras de associação que relacionam valores de um subconjunto de atributos previstos com os valores de outro subconjunto [37]. A Figura 2.4 ilustra os tipos de aprendizagem e tarefas relacionadas hierarquicamente. O AM indutivo aparece na parte superior, que aprende a generalizar a partir de um conjunto de dados. De-

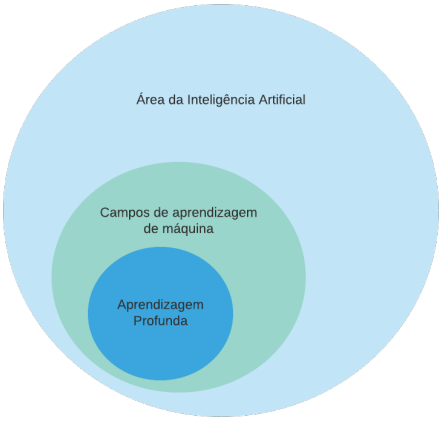

Figura 2.3: Áreas da Inteligência Artificial [2]

pois, há as categorias de algoritmos de aprendizagem supervisionada (tarefas preditivas) e aprendizagem não supervisionada (tarefas descritivas) [3].

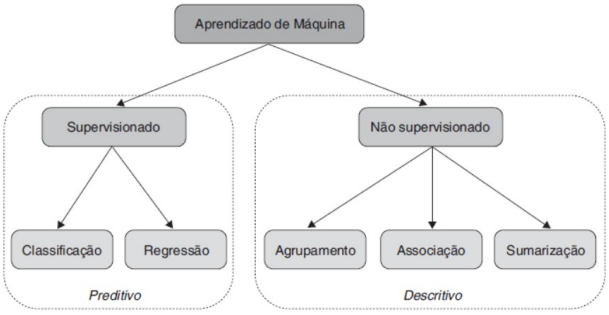

Figura 2.4: Hierarquia de aprendizado [3]

Tarefas preditivas são distinguidas pelo valor do rótulo a ser previsto: para tarefas de classificação, é cauteloso; para tarefas de classificação, é cauteloso; para tarefas de classificação, é cauteloso. Para tarefas de classificação, é cauteloso; para tarefas de regressão, é contínuo. Tarefas descritivas são geralmente divididas em: agrupamento, dividindo os dados em vários grupos de acordo com a similaridade dos dados; abstrato; e associação, para encontrar padrões de associação frequentes entre os atributos do conjunto de dados [43]. ML tem como premissa o acúmulo de conhecimento, portanto, extrair significado de um determinado conjunto de dados é a premissa. Esse processo pode ser dividido em sete etapas: coleta de dados, preparação dos dados, seleção do modelo, treinamento, avaliação, ajuste dos parâmetros e aplicação (Figura 2.5) [37].

1. **Seleção:** Dados em quantidade e qualidade suficientes devem ser selecionados para extrair o conhecimento necessário. Quanto mais exemplos/amostras você conseguir, melhor será o aprendizado. Em contrapartida, a qualidade está diretamente relaci-

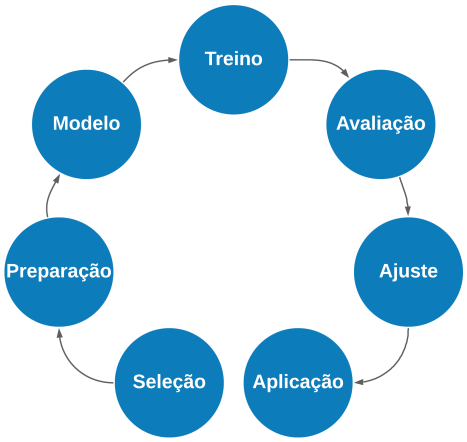

Figura 2.5: Processo do aprendizado de máquina [3]

onada ao atributo selecionado. Deve ser dada prioridade aos mais relevantes para o modelo. Dependendo do tipo de aprendizado, os dados obtidos podem incluir atributos previstos e atributos de destino [44].

2. **Preparação:** Os dados devem ser preparados e adaptados ao modelo utilizado. Nesta etapa, ele incluirá conversão de unidade, conversão de escala, normalização, discretização e alterações de representação de dados. É importante avaliar o equilíbrio dos dados, ou seja, se os dados coletados para diferentes faixas ou categorias de predições existem em quantidades iguais. Dados incompletos, inconsistentes, repetitivos ou ruidosos devem ser processados ou eliminados de forma a não afetar o processo de treinamento e, assim, distorcer o aprendizado. Os dados também podem ser divididos em dois grupos, um para a fase de treinamento e outro para teste. Essa medida evita a confusão entre um modelo dedicado no conjunto de dados de treinamento e um modelo que pode generalizar o aprendizado e fazer previsões corretas sobre dados até então desconhecidos [36].

3. **Modelo:** Deve-se atentar para as vantagens fornecidas para os tipos de dados envolvidos e suas respectivas complexidades. Os modelos podem incluir regressão linear, regressão logística, classificação, agrupamento, aprendizado profundo, etc [43].

4. **Treino:** O objetivo do treinamento é aprimorar o modelo a partir da avaliação de cada nova amostra de treinamento. Um problema linear simples representado por apenas um recurso pode ser determinado de acordo com a seguinte equação [3]:

$$y = b + ax$$

- y = saída
- x = entrada
- b = viés
- a = peso

O treinamento altera o modelo especificado pela equação, modificando os parâmetros de viés (b) e o peso (a), já que os outros não podem ser alterados. Conforme o número de atributos previstos aumenta, o número de parâmetros pode se tornar maior. O algoritmo deve ter meios para avaliar ou comparar os resultados obtidos com o modelo a cada etapa do teste realizada, o que é fundamental. O treinamento consiste na inferência de hipóteses que mais se aproximem do resultado adequado para a saída a partir de determinado conjunto de dados de entrada [42]. A máquina deve ser programada para buscar a constante redução do conjunto de hipóteses durante o treinamento.

5. **Avaliação:** Utiliza-se os dados inicialmente separados para teste e determina se o modelo obtido após o treinamento tem a precisão esperada ao prever o destino com base no conjunto de dados anteriormente desconhecido [37].
6. **Ajuste:** Após a fase de avaliação, novos hiper-parâmetros (controlam o próprio processo de aprendizagem) podem ser definidos, incluindo: o número de vezes que todo o processo é repetido no conjunto de dados de treinamento; a taxa de aprendizagem, que é um fator que afeta as mudanças nos parâmetros do modelo; e uns aos outros [3].
7. **Aplicação:** Envolve o uso do modelo para fazer previsões a partir de uma máquina que foi treinada [37].

Vários algoritmos foram desenvolvidos para implementar o aprendizado supervisionado e o aprendizado não supervisionado. Abaixo estão listados alguns desses algoritmos [36]:

Modelo: Supervisionado

Tipo: Regressão

- Regressão linear: recebe o valor de algumas variáveis, obtém um valor estimado por meio de uma equação e o aplica a outras variáveis.

Tipo: Classificação

- Regressão logística: Permite definir características semelhantes a determinados grupos de variáveis.

- Árvores de decisão: Realizam uma pesquisa de cima para baixo nos dados calculando todas as árvores possíveis. Quando a árvore é muito complexa, seu tamanho será reduzido para torná-la o mais versátil possível. Para classificar os elementos, a árvore precisa ser percorrida. Quando uma instância encontra uma folha, sua classe é a classe folha correspondente
- Redes neurais artificiais: É baseado em um sistema de aprendizado biológico formado pela interconexão de neurônios. Como nos neurônios, o sistema conecta várias unidades simples que recebem informações de outros elementos (entradas) e enviam as informações processadas para outros elementos. Esses algoritmos aprendem exemplos e generalizam conceitos.
- K-Vizinhos mais próximos: A classificação dos itens é realizada comparando a similaridade dos itens a serem classificados com os dados de treinamento.

Modelo: Não supervisionado
Tipo: Agrupamento

- K-Means: É um algoritmo de particionamento que divide os dados em grupos separados (clusters), onde os objetos são apenas parte do grupo. O algoritmo encontra uma maneira de dividir melhor os dados X em K grupos, agrupando assim dados semelhantes. Cada grupo é representado por seu centro e cada dado está contido no grupo mais próximo a ele.
- Hierárquicos: A sequência de partições aninhadas é gerada com base na matriz de proximidade. Os resultados desses algoritmos dependem da ordem de entrada dos dados.
- Grafos: O algoritmo usa grafos vizinhos para realizar o agrupamento.

2.6 Big Data

A quantidade de dados no mundo está crescendo exponencialmente, e a análise desses conjuntos de big data chamados de big data se tornou uma parte fundamental da sobrevivência em um ambiente competitivo. O advento da era do big data trouxe diferentes definições e perspectivas, principalmente para quem trabalha em áreas afins (como gestão do conhecimento e inteligência competitiva) [45]. Mais precisamente, o termo big data se refere à capacidade de processar grandes quantidades de repositórios de informações, grandes quantidades de dados estruturados, semiestruturados ou mesmo não estruturados e completamente diferentes em quase tempo real. Por meio desse processamento, é possível encontrar novos dados, caso não haja tal processo, esses dados serão desconhecidos

[46]. Além do problema da grande capacidade, outras características do big data também são mencionadas, como a alta velocidade de processamento e a grande variedade de dados envolvidos. De acordo com a visão da IBM (2014), para a empresa, o big data é baseado em 4 V's: volume, diversidade, velocidade e precisão [45]:

1. Volume: grandes quantidades de dados estão sendo geradas;
2. Variedade: Indica que os analistas agora estão processando dados contendo vários formatos, incluindo dados estruturados e não estruturados, dados de e-mail, dados de mídia social, dados de sensores, etc;
3. Velocidade: como bilhões de sensores continuam a coletar dados, o fluxo de dados não para. E com a melhoria das redes de comunicação atual, sua velocidade de chegada é mais rápida do que antes, o que também significa que a velocidade de processamento também é mais rápida;
4. Veracidade: a maioria desses dados pode não ser confiável ou ser incompleta; portanto, apesar da incerteza inicial, novas tecnologias que possam fornecer uma visão consistente ainda são necessárias.

Com o aumento massivo da quantidade de dados gerados pela Internet Com o surgimento das mídias sociais, é necessário gerenciar e armazenar informações de forma organizada. De acordo com o gerenciamento e armazenamento desses dados, eles podem ser divididos em estruturados, não estruturados e semiestruturados [36].

- Dados estruturados: são dados com forma e comprimento definidos, como números, datas e grupos de palavras. Na maioria dos casos, os dados estruturados são o resultado do processo de gerar dados inerentes ao sistema de negociação ou dados gerados pelo processo de observação e medição;
- Dados não-estruturados: são dados que não seguem um formato específico, como imagens de satélite, dados científicos, fotos e vídeos, texto específico da empresa e dados de mídia social. Esses dados requerem equipamentos de armazenamento e processamento que suportem seu formato e garantam uma melhor eficiência de análise;
- Dados semiestruturados: podem ser definidos como o meio-termo entre a não-estruturada e a estruturada.

2.7 Trabalhos Correlatos

Em seu trabalho, Ribeiro [47] estudou a existência de discriminação em algoritmos de Inteligência Artificial e a possibilidade da LGPD como ferramenta normativa para amenizar o viés discriminatório nessas decisões do sistema. Após o conceito inicial, explicou o que é um algoritmo e, em seguida, apresentou diferentes exemplos de discriminação em um sistema de decisão automática. Depois, é conceituado qual é a distinção entre algoritmos e as maneiras pelas quais algoritmos podem se tornar enviesados. Concluindo então, que algoritmos de Inteligência Artificial não são neutros porque podem tomar decisões discriminatórias por meio de vieses em seus sistemas, e dizendo que a LGPD levará as instruções impostas por lei aos desenvolvedores e empresas, que são obrigados a realizar investigações e auditorias mais aprofundadas a fim de analisar o verdadeiro potencial discriminatório de seus sistemas e garantir que os resultados cumpram os requisitos legais [47].

Nesta pesquisa, a União Europeia abordou a relação entre dados gerais das Regulamentações de proteção (GDPR) e Inteligência Artificial (IA). Apresentou alguns conceitos básicos de IA e estudou a tecnologia mais recente com foco na aplicação de IA em dados pessoais. Em seguida, o estudo fornece uma análise sobre como ajustar IA respeitando a GDPR, verificando até que ponto a IA é aplicável ao GDPR em sua estrutura conceitual. Depois, discutiu a tensão e a proximidade da IA e os princípios de proteção de dados (por exemplo, restrições de propósito e minimização de dados). Ele verificou a aplicação de IA quando se processa a base de dados pessoais, especialmente sobre a tomada de decisão automática. Procedeu então, a uma análise completa da tomada de decisão automática, tendo em vista a aceitabilidade, e as medidas de proteção a serem tomadas caso o titular dos dados requerer o direito de receber uma explicação pessoal. Por isso, então, mostra até que ponto o GDPR fornece uma abordagem preventiva baseada em risco, com foco em conceitos de prevenção e proteção de dados por padrão. A possibilidade de usar IA para fins estatísticos de forma consistente com o GDPR também é considerada. A conclusão do estudo é que a IA pode ser implantada de maneira consistente com o GDPR, mas o GDPR não pode fornecer orientação adequada para controladores e precisa expandir e implementar seus regulamentos. Algumas sugestões foram feitas a esse respeito [11].

De acordo com Gruschka et al. [48], Big data aumentou o acesso a informações confidenciais, que, após o processamento, irão prejudicar diretamente a privacidade pessoal e violar as leis de proteção de dados. Como resultado, o controlador e o processador de dados podem ser severamente punidos por não conformidade e podem até levar à falência. Por isso, neste artigo, é discutido o estado atual das leis e regulamentos e analisa-se diferentes tecnologias de proteção de dados e privacidade no contexto da big data. Além disso, é apresentado e analisado dois projetos de pesquisa da vida real como estudos de

caso envolvendo dados confidenciais e conformidade com os regulamentos de dados. Por fim, os autores explicaram quais tipos de informações podem se tornar riscos à privacidade, as tecnologias de proteção da privacidade usadas de acordo com os requisitos legais e o impacto dessas tecnologias no estágio de processamento de dados e nos resultados de pesquisa [48].

O novo Regulamento Geral Europeu de Proteção de Dados impõe restrições estritas ao processamento de dados de identificação pessoal. Pensando nisso, Bonatti e Kirrane [49], mostraram que a GDPR não afeta apenas as empresas européias, porque o regulamento se aplica a todas as organizações que monitoram ou prestam serviços aos cidadãos europeus. Só é permitida a análise exploratória gratuita de dados anônimos, o que traz alguns riscos jurídicos. E eles acreditam que para outros tipos de processamento de dados pessoais, a base legal mais flexível e segura é o consentimento explícito. Explicam então os métodos de gerenciamento de consentimento em conformidade com o GDPR sendo desenvolvidos pelo projeto europeu H2020 SPECIAL e também é apresentado alguns aspectos relevantes de big data [49].

O Regulamento Geral de Proteção de Dados da União Europeia (GDPR) exige o princípio de minimização de dados, que usam apenas dados da coleção necessária para atingir um propósito específico. No entanto, muitas vezes é difícil determinar a quantidade mínima de dados necessária, especialmente em modelos complexos de aprendizado de máquina, como redes neurais. Goldsteen et al. [4] propuseram um método inovador e bastante prático para reduzir a quantidade de dados pessoais necessários para fazer previsões usando modelos de aprendizado de máquina, excluindo ou resumindo alguns recursos de entrada. Esse método usa o conhecimento codificado no modelo para gerar generalizações que têm pouco efeito em sua precisão. Permitindo que os criadores e usuários de modelos de aprendizado de máquina minimizem os dados de maneira comprovável [4].

As questões legais e éticas constituem um componente importante da pesquisa moderna, relacionadas ao sujeito e ao pesquisador. Yip et al. [50] revisaram brevemente as várias diretrizes e regulamentações internacionais que existem sobre questões relacionadas ao consentimento informado, confidencialidade, fornecimento de incentivos e várias formas de má conduta de pesquisa. Os pesquisadores devem observar as principais diretrizes internacionais e as diferenças regionais na legislação. Portanto, aconselhamento ético específico deve ser procurado nos Comitês de Revisão de Ética locais.

A Inteligência Artificial é uma tecnologia emergente, mas o sistema de política atual não é perfeito e o mecanismo de supervisão não está em vigor. Essa tecnologia inevitavelmente traz riscos como vazamento de privacidade pessoal, aumentando a lacuna entre ricos e pobres e poluição ambiental. Também levanta questões éticas, como ética dos direitos humanos, ética da informação e ética da responsabilidade. A tecnologia de Inte-

ligência Artificial está se desenvolvendo rapidamente a fim de salvaguardar os interesses fundamentais do ser humano e promover o desenvolvimento saudável da sociedade. Li et al. [51] reforçaram que é preciso fortalecer a cooperação internacional, estabelecer políticas públicas sólidas e promover o estabelecimento da ética da Inteligência Artificial e outras soluções.

Kwan et al. [52] utilizaram os princípios da teoria fundamentada para eliciar conhecimentos relacionados à Confiança, Ética e Transparência. São abordadas essas qualidades como Requisitos Não Funcionais (NFRs), com o objetivo de construir catálogos para subsidiar a construção do *Socially Software* responsável. O corpus usado foi construído em uma coleção selecionada da literatura sobre Responsabilidade Social Corporativa, com ênfase em Ética Empresarial. O desafio é como codificar o conhecimento da perspectiva social, principalmente através da visão da Responsabilidade Social Corporativa, sobre como as organizações ou instituições alcançam a confiabilidade.

Cerqueira [53] propôs um guia para apoiar os Product Owners e desenvolvedores de sistemas baseados em IA na elicitação de requisitos éticos. Foi utilizado a metodologia Design Science Research e na etapa de compreensão do problema realizaram uma revisão sistemática de literatura. Além disso, desenvolveram um guia online e realizaram a sua validação através de um survey aplicado em um grupo focal com profissionais da área. Os resultados preliminares revelaram que o Guia contribui para preencher a lacuna entre princípios de alto nível e abstratos e a prática, auxiliando os desenvolvedores e Product Owners, principalmente em projetos de desenvolvimento ágil, a elicitar requisitos éticos e operacionalizar a ética em IA.

Capítulo 3

Metodologias Identificadas na Literatura

3.1 Metodologias

3.1.1 Minimização dos Dados para conformidade com a GDPR em modelos de Machine Learning

No método proposto por Goldsteen et al. [4], os autores escolheram a métrica *Normalized Certainty Penalty*(NCP) [54] para medir a qualidade dos resultados da generalização. Essa métrica basicamente compara os tamanhos dos intervalos generalizados com o original, calculando a média dessa pontuação em todos os registros do conjunto de dados. Como o objetivo principal era priorizar a privacidade dentro das restrições de precisão, por isso, a alta perda de informação que existe usando o NCP é a qualidade desejada.

Treino Generalizado

A metodologia é composta por várias etapas, a primeira é **treinar uma base de dados generalizada**, isto é, treinar um modelo para prever as previsões do modelo original, com o objetivo de aprender os limites de decisão deste modelo. Por isso, foi escolhido o algoritmo arvore de decisão uni-variada [55] para fazer essa generalização, uma vez que as divisões que a árvore cria em cada nó interno podem ser usados como base para determinar os intervalos generalizados.

Como o objetivo é encontrar a melhor generalização sem prejudicar a precisão do modelo, os autores começam gerando uma árvore de decisão com folhas homogêneas. Cada folha contém apenas entradas que geram a mesma previsão no modelo original. Em seguida, **deriva-se o conjunto inicial de generalizações** combinando todos os valores da divisão de cada característica dos nós internos da árvore [4].

Depois que o conjunto inicial de recursos generalizados é obtido, eles aplicam as generalizações aos dados de teste e verificam a precisão do modelo original sobre ele. Medindo a **precisão relativa**, ou seja, qual porcentagem das previsões originais são retidas ao aplicar o modelo aos dados generalizados. Com base na precisão relativa medida, é tomada a decisão de continuar ou não o processo. Se o limite de precisão for alcançado, as generalizações derivadas diretamente do modelo generalizador são usadas. Se a precisão for inferior ou superior ao limite desejado, prossegue com a execução de etapas adicionais para melhorar a precisão ou a generalização. No final de cada iteração, a precisão resultante é novamente medida aplicando a generalização atual aos dados de teste e verificando a precisão [56].

Se a precisão alcançada for superior ao limite, emprega-se uma etapa para **melhorar a generalização**. Isso é feito podando iterativamente a árvore de decisão, ou seja, subindo das folhas para os nós mais altos da árvore. Cada poda remove efetivamente (pelo menos) um valor de divisão para um dos recursos, combinando dois intervalos de nível inferior em um único intervalo e reduzindo o número geral de intervalos para aquele recurso. Na implementação apresentada, não é utilizado nenhum meio sofisticado para escolher quais nós podar. O algoritmo simplesmente sobe um nível em toda a árvore simultaneamente. Continua a subir na árvore, um nível de cada vez, até alcançar o nó raiz ou até que o limite de precisão seja alcançado [57].

Se a precisão alcançada for inferior ao limite, é empregado uma etapa projetada para **melhorar a precisão**, removendo recursos da generalização. Isso significa que, em vez de generalizá-lo, esse recurso não será alterado. Pra esse passo, eles definiram uma métrica adicional chamada ILAG, baseada na função de pontuação de Fung et al. [58], Goldsteen et al. [4] a adaptaram para maximizar a perda de informação (pontuação NCP) para cada unidade de ganho de acurácia. Então, o ILAG [4] ficou definido como:

$$\begin{cases} \frac{NCP(f)}{GanhoAcuracia(f)}, & \text{se } GanhoAcuracia(f) \neq 0 \\ NCP(f), & \text{caso contrário} \end{cases} \tag{3.1}$$

O processo de minimização eventualmente produz um conjunto de dados mínimo necessário para atingir o nível de precisão exigido. A saída é um conjunto de intervalos de recursos generalizados ajustados para um modelo de ML específico [57]. Este conjunto de recursos generalizados podem ser usados sempre que coletar novos dados para análise. O processo de minimização completo [4] é apresentado na Figura 3.1.

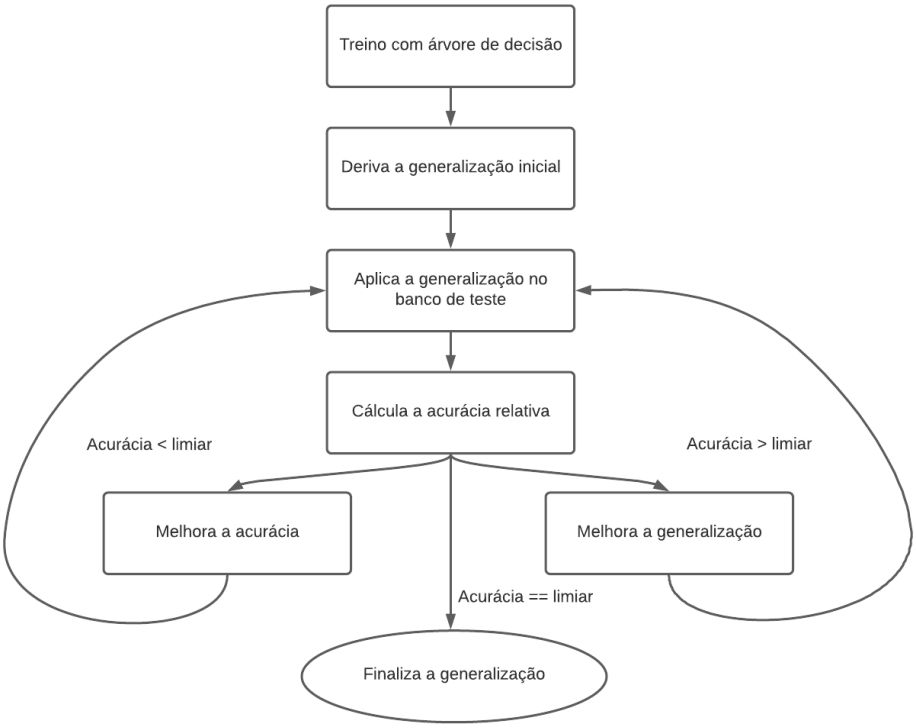

Figura 3.1: Processo de minimização [4]

Aplicando as generalizações em dados recentemente coletados

Depois que o conjunto de recursos generalizados é determinado, há várias maneiras de coletar novos dados para classificação. A primeira opção apresentada, é usar os intervalos de recursos calculados combinando todos os valores de divisão para cada recurso. Isso resulta em uma recodificação global dos dados, ou seja, cada característica tem um conjunto predeterminado de intervalos. Ao usar essa abordagem, o usuário cujos dados são coletados nunca realmente divulgam seus dados exatos, apenas os intervalos relevantes [4] [56].

Uma segunda opção abordada, é ter um procedimento de minimização, no qual uma parte do código mapeia pontos de dados originais para pontos de dados generalizados, dependendo do cluster ao qual o ponto de dados pertence. Esse trecho de código pode ser executado no terminal onde os dados são coletados e generalizar imediatamente os dados brutos antes de enviá-los para análise; por exemplo, em um servidor ou nuvem. Tal procedimento de minimização pode resultar em uma recodificação local, potencialmente permitindo melhores generalizações [4] [57].

Outra opção é determinar dinamicamente os intervalos apresentados ao usuário, com base em suas escolhas para recursos anteriores. Cada vez que um valor generalizado é selecionado, esta informação pode ser usada para melhorar dinamicamente as genera-

lizações para outros recursos. Isso é possível porque as generalizações de um recurso podem ter sido restritas por domínios que não são mais relevantes uma vez que o valor de outro recurso é conhecido. A ordem de preenchimento dos valores do recurso pode ser determinada pelo usuário [4] [56].

3.1.2 Desenvolvimento de IA Sustentável (*Sustainable AI Development*, SAID)

Três camadas que devem ser abordadas pelo SAID proposto por Djeffal [59]: a camada de tecnologia, a camada social e a camada de governança. Juntas, essas camadas tem o objetivo de oferecer um bom design de Inteligência Artificial. As camadas não devem ser percebidas como existindo em isolamento clínico uns dos outros, mas como nós diferentes em uma interação ativa. Deve haver um feedback constante e ajustes a fim de melhorar as escolhas de design, configurações técnico-sociais e o direito de governança. No entanto, para conseguir atingir seu objetivo, é necessário compreender as diferentes camadas e as escolhas associadas.

Camada de Tecnologia

A camada de tecnologia traduz questões de desenvolvimento sustentável para o nível de aplicações específicas. Em consonância com isso, as escolhas de design tecnológico devem ser identificadas, destacadas e analisadas. O importante é vincular o design da tecnologia aos objetivos perseguidos pelo Desenvolvimento Sustentável (DS) [60]. Nesse sentido, os Objetivos do Desenvolvimento Sustentável (ODS's) podem desempenhar diferentes papéis: o primeiro papel é usar a tecnologia para a realização dos ODS's. No caso da IA, seria o uso de diferentes tecnologias de IA para atingir os objetivos de desenvolvimento sustentável. Caso não haja impacto positivo nos ODS, a camada de tecnologia orienta as escolhas de design. Existem várias iniciativas que analisam o que podem significar boas escolhas de design e essas escolhas de design têm um efeito sobre como o sistema opera e como ele pode ser compreendido [59].

Camada Social

A camada social analisa as consequências do uso de sistemas de IA na esfera social, ou seja, a realidade sociotécnica de um sistema de IA. O foco da análise de sustentabilidade aqui olha para os impactos dos sistemas nos indivíduos, grupos e na sociedade como um todo. O duplo efeito do desenvolvimento sustentável também atua neste campo. No que diz respeito ao acesso à justiça, existem aplicações que permitem aos cidadãos fazer

reivindicações específicas, no entanto, a microssegmentação, por exemplo, pode ter como objetivo barrar o acesso à justiça em certas situações [59] [61].

Camada de Governança

A camada de governança examina todas as formas de influenciar os sistemas de sistemas artificiais, independentemente do nível (nacional, internacional, transnacional). Junto com os objetivos de desenvolvimento sustentável, foi elaborada uma estrutura de governança específica para a realização das prisões. Conforme mencionado anteriormente, o ODS aborda exatamente essa questão ao examinar e questionar a governança e a implementação. Naturalmente, ele vê a governança de uma perspectiva de vários níveis, permitindo diferenças no terreno, ao mesmo tempo em que enfatiza a comparabilidade entre as diferentes camadas da governança [59] [60].

3.1.3 ECCOLA

A ECCOLA foi desenvolvida na forma de um "baralho de cartas", esse método foi baseado na Teoria da Essência da Engenharia de Software (Essence Theory of Software Engineering) [62], que foi usado para descrever sua primeira versão. Os métodos descritos na linguagem Essence são utilizados por meio de cartões [5].

Em seu desenvolvimento, foi estipulado três objetivos principais para o método: ajudar a criar consciência sobre a ética da IA e sua importância; fazer um método modular adaptável adequado para uma ampla variedade de contextos na Engenharia de Software, e; tornar a ECCOLA adequada para o desenvolvimento ágil e também fazer da ética uma parte do desenvolvimento ágil em geral [5].

Existem 21 cartões no total na ECCOLA (Figura 3.2, Figura 3.3, Figura 3.4, Figura 3.5, Figura 3.6, Figura 3.7, Figura 3.8, Figura 3.9), e esses cartões são divididos em 8 temas, com cada tema consistindo de 1 a 6 cartões. Esses temas de ética da IA são encontrados em várias diretrizes éticas [63], como transparência ou dados. Cada cartão individual, então, trata de um aspecto mais atômico daquele tema, como, no caso de dados, privacidade e qualidade de dados [5].

Cada cartão é dividido em três partes: (1) motivação (ou seja, o por quê isso é importante), (2) o que fazer (para resolver este problema) e (3) um exemplo prático do tópico (para tornar as questões mais tangíveis). Como os cartões são geralmente utilizados como cartões físicos, o cartão é dividido em dois, com a metade esquerda de cada cartão contendo o conteúdo textual e a metade direita contendo um espaço em branco para anotações. Este espaço para fazer anotações foi incluído para tornar o uso dos cartões mais convenientes na prática [5].

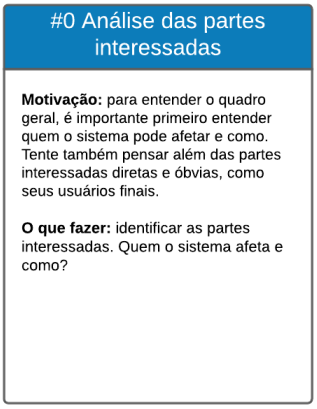

Figura 3.2: Card do processo de analise [5]

Figura 3.3: Cards do processo de transparência [5]

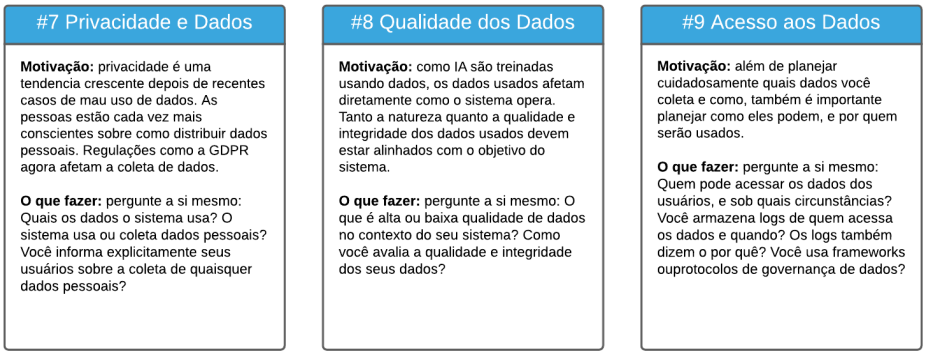

Figura 3.4: Cards do processo de dados [5]

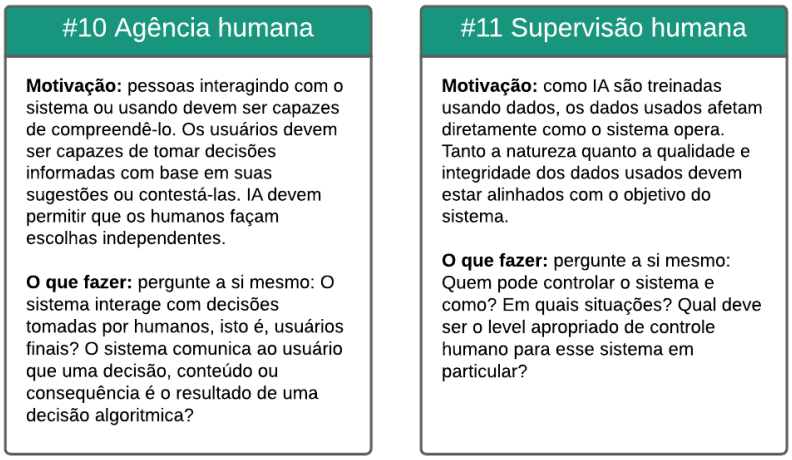

Figura 3.5: Cards do processo de agência e supervisão [5]

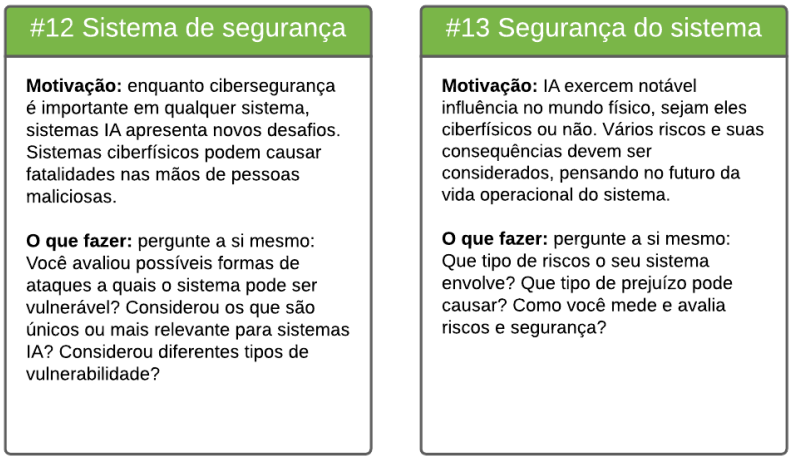

Figura 3.6: Cards do processo de segurança [5]

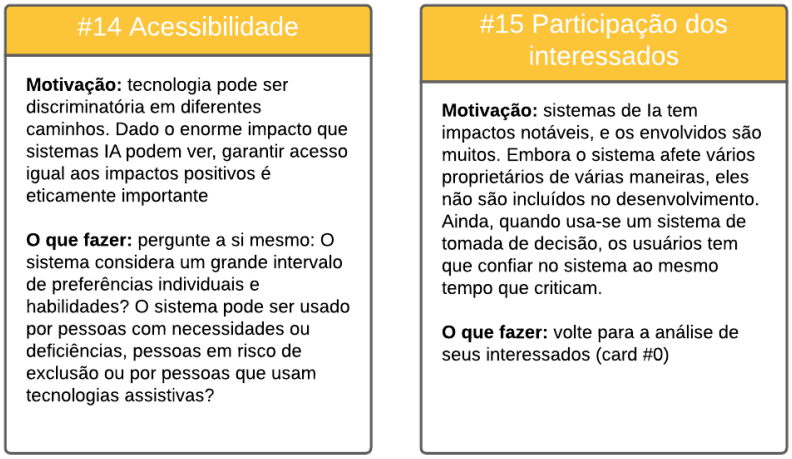

Figura 3.7: Cards do processo de justiça [5]

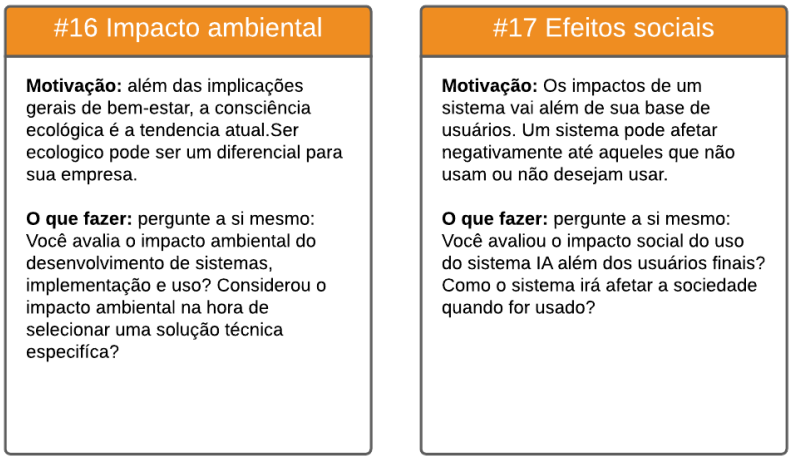

Figura 3.8: Cards do processo de bem-estar [5]

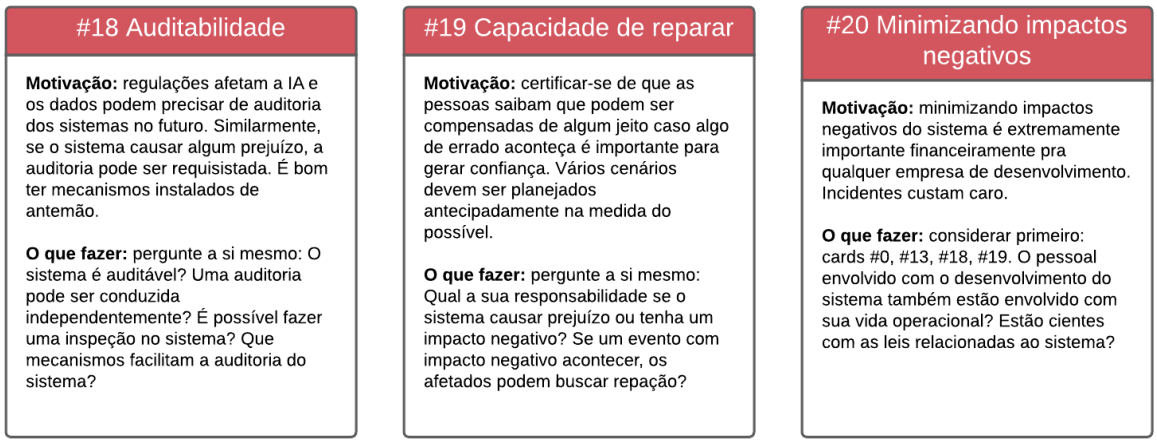

Figura 3.9: Cards do processo de prestação de contas [5]

3.2 Metodologias/Frameworks no conceito da ética

A ética deve ser observada nas metodologias/frameworks para o desenvolvimento de softwares de acordo com Peterson [64], abordando 5 princípios: custo-benefício, precaução, sustentabilidade [65], autonomia e igualdade. Nessa análise não será abordado o custo-benefício, pois o intuito é verificar a integridade em relação a ética, independente do custo que pode gerar essa implementação.

3.2.1 Minimização dos Dados para conformidade com a GDPR em modelos de Machine Learning

Precaução: Se encaixa no processo de generalização dos dados, onde os dados quando generalizados não são possíveis mais serem relacionados a seus donos, com isso, preservando a privacidade destes [64] [4].

Sustentabilidade: A metodologia é sustentável, pois não irá gerar mais impactos naturais além dos atuais, já que não será necessário adquirir novos equipamentos ou aumentar o gasto com energia elétrica para sua implementação [65],[4].

Autonomia: Não há especificação de como devem ser coletados os dados na metodologia, pois é utilizado uma base pública de dados, por isso o desenvolvedor tem a autonomia de trabalhar com esses dados. O usuário por sua vez, pode não estar ciente disso, não tendo então o conhecimento de que ele tem a autonomia de pedir explicações de o por quê seus dados foram usados [64] [4].

Igualdade: A generalização gera uma igualdade justa entre os dados, pois não leva em conta diferenças sociais ou culturais em sua implementação, e tanto quanto o resultado independe destes fatores [64] [4].

3.2.2 Processo de desenvolvimento ECCOLA

Precaução: A metodologia é precavida em todos os aspectos, indica primeiro entender quem o sistema pode afetar e como, além ter cards sobre transparência e prestação de contas. Ainda recomenda a cada card implementado, realizar 3 ações, sendo elas: (1) Se preparar, ou seja, escolher os cards que precisam ser implementados; (2) Revisar, sempre ter os cards em mãos, anotar as ações tomadas e as discussões éticas levantadas durante o processo; (3) Avaliar, revisar para que todas as ações preparadas foram efetuadas, e se necessário, repetir todo o processo [5] [64].

Sustentabilidade: É explorado o impacto ecológico que o desenvolvimento de IA tem sobre o meio ambiente e faz perguntas aos desenvolvedores nesse sentido, fazendo-os examinar as fontes de energia dos data centers e seu impacto. O bem-estar é analisado da perspectiva do efeito social e faz perguntas aos desenvolvedores para avaliar o impacto mais amplo do uso de IA além dos usuários alvos imediatos e o efeito sistêmico que isso poderia ter na sociedade em geral [66] [64].

Autonomia: O ECCOLA explora esse princípio por meio do conceito de agência humana e faz perguntas aos desenvolvedores na interação homem-máquina. Demonstra a necessidade de compreensão pelos usuários humanos, o trabalho e as decisões tomadas para apoiar a tomada de decisão humana e permitir que os humanos tomem suas próprias decisões. Examina também a supervisão humana levantando questões para desenvolvedores sobre a medida de suporte que a IA oferece aos usuários humanos e se o sistema mina a autonomia humana ao ignorar suas decisões [66] [64].

Igualdade: A justiça é abordada a partir da diretriz de diversidade da União Européia, ela permite que os desenvolvedores façam perguntas sobre se a IA permite acesso igual em termos de uma ampla gama de usuários em relação a deficiências e grupos diversos, incluindo os grupos de usuários não tecnicamente experientes. Analisa também a participação das partes interessadas no desenvolvimento e pergunta aos desenvolvedores se essas diferentes partes, especialmente os usuários finais, estão incluídos como parte do desenvolvimento do sistema [66] [64].

3.2.3 Desenvolvimento de IA Sustentável (*Sustainable AI Development*, SAID)

Precaução: É reiterado a todo tempo na metodologia a importância de se ter uma aplicação sustentável e ética, mas não é explicitado um tipo, ou um passo a passo para se precaver de possíveis erros ou tomadas de decisões erradas [59] [64].

Sustentabilidade: Inclui direitos humanos, mas também uma preocupação mais coletiva. Como estrutura, a SAID também poderia ajudar a atualizar o discurso sobre o desenvolvimento sustentável e a trazer novas considerações que são importantes para o desenvolvimento sustentável na era da tecnologia digital [59] [64].

Autonomia: O objetivo é construir tecnologias sustentáveis, em linha com isso, as escolhas de design tecnológico devem ser identificadas, destacadas e analisadas. O importante é vincular o design da tecnologia aos objetivos perseguidos pelo DS. Nesse sentido, os ODS's podem desempenhar diferentes papéis: o primeiro papel é usar a tecnologia para a realização do ODS. No caso da IA, seria usar diferentes tecnologias de IA para atingir os objetivos de desenvolvimento sustentável. O desenvolvedor tem a autonomia de escolher, mas pode ter um grande impacto no acesso à justiça, especialmente quando a base de uma decisão é confusa. Essas escolhas de design afetam como o sistema opera e como pode ser compreendido [59] [64].

Igualdade: Considerando as diversas formas de se atingir ou não determinados objetivos, uma mera avaliação de impacto não será suficiente. Se abordarmos a implementação de um sistema de IA de uma perspectiva de equidade, a avaliação de impacto só pode ser feita levando-se em conta também o estado atual das coisas. A proposta aqui feita é fazer uma comparação sociotécnica que não focalize apenas os impactos positivos e negativos, mas também avalie a situação atual. A SAID é uma estrutura eficaz para equilibrar diferentes elementos, considerando que o desenvolvimento sustentável passa por solução de considerações conflitantes. Por meio do SAID, a governança da IA está inserida em um processo internacional inclusivo que abrange todos os países e diversas partes interessadas. Estabelece certos objetivos comuns e uma estrutura para medir o sucesso, mas também deixa liberdade para a implementação em vários níveis [59] [64].

3.3 Metodologias/Frameworks no conceito da LGPD

A LGPD [6] estabelece normas e regras rigorosas para a proteção de dados pessoais, regulamentando seu tratamento, definido como qualquer ação realizada desde a coleta,

cópia, edição, armazenamento, publicação, impressão, transmissão, processamento e compartilhamento de dados pessoais, e o objetivo desta seção é verificar a conformidade com a LGPD das metodologias escolhidas [20].

Para verificar se as metodologias respeitam as normas da LGPD, será considerado o checklist de conformidade à LGPD proposto por Marinho [20], que é composto por uma lista com 8 itens essenciais, sendo eles:

1. **Estabelecer uma estrutura de prestação de contas e governança**, pois para se ter conformidade é necessário ter o apoio da gestão, é essencial que a diretoria entenda as possíveis implicações da Lei e garanta os recursos necessários para alcançá-la.

2. **Escopo e planejamento do projeto**, é necessário saber quais áreas se encaixam no escopo da LGPD e considerar os processos existentes que podem ser afetados.

3. **Realizar um inventário de dados e uma auditoria de fluxo de dados**, sem entender os quais dados são processados e como eles são processados se torna impossível cumprir os requisitos de processamento de dados da LGPD.

4. **Realizar uma análise detalhada de brechas**, a abordagem correta da conformidade estabelece identificar as lacunas que precisa preencher, isto é, avaliar seus fluxos de trabalho, processos e procedimentos atuais.

5. **Desenvolver políticas, procedimentos e processos operacionais**, promover uma avaliação de suas práticas de gerenciamento de privacidade e processamento de dados, e então obter um relatório resumido das suas lacunas de conformidade juntamente com as possíveis correções.

6. **Proteger os dados pessoais por meio de medidas processuais e técnicas**, é exigido que as organizações implementem medidas técnicas e organizacionais para que os dados pessoais sejam processados corretamente.

7. **Comunicações**, todos os envolvidos no processamento de dados devem ser capacitados e treinados para seguir processos e procedimentos internos, e com isso estabelecer comunicações internas eficazes.

8. **Monitorar e auditar a conformidade**, a conformidade com LGPD é um projeto dinâmico, deve-se realizar auditorias internas regularmente e atualizar seu processo de proteção de dados, incluindo a verificação de seus registros de atividades de processamento (logs), mecanismos de consentimento, testes de controles de segurança de informações e análise de impacto de privacidade (PIA).

3.3.1 Minimização dos Dados para conformidade com a GDPR em modelos de Machine Learning

Estrutura de prestação de contas e governança

Não foi constatado na metodologia.

Escopo e planejamento do projeto

É escolhido um conjunto de dados e treinados um ou mais modelos originais nele (primeiro aplicando a seleção de recursos). Depois, considera o modelo resultante e sua precisão como a linha de base. Mas não é apresentado um planejamento de como deve ser, ou como vai acontecer o projeto, entidades que estarão no escopo, ou padrões de sistema a serem usados [4] [20].

Realizar um inventário de dados e uma auditoria de fluxo de dados

Não foi constatado na metodologia.

Realizar uma análise detalhada de brechas

A preocupação da metodologia foi diretamente com os dados e a necessidade de generalizar essas informações para preservar a privacidade de seus donos, mas não há a discussão de em quais pontos poderá haver uma falha ou lacunas em relação as normas [4] [20].

Desenvolver políticas, procedimentos e processos operacionais

A politica empregada para definir a privacidade, a proteção dos dados, e a segurança da informação são simplesmente duas métricas, a NCP e o ILAG, onde NCP é definido entre 0 e 1, onde 1 significa perda total dos dados, e 0 a base original de dados. O ILAG nada mais é que a divisão do NCP pelo ganho de acurácia da ultima iteração, o que é bastante perigoso e falho, pois toda a privacidade do usuário depende de um cálculo, gerando desconfiança e insegurança. Faltando várias etapas que garantem uma boa conformidade, como planejar e lidar com as solicitações de acesso de sujeitos dos dados, políticas para requerer o consentimento de acesso aos dados, entre outros [4] [25].

Proteger os dados pessoais por meio de medidas processuais e técnicas

A anonimização é implementada através da generalização dos dados, mas é necessário ainda explicitar e informar o dever de se ter uma política de segurança e privacidade entendida e assimilada por todos da organização [4] [67].

Comunicações

Não foi constatado na metodologia.

Monitorar e auditar a conformidade

Não é explicitado um registro das tomadas de decisões, apesar do algoritmo usado ser uma árvore o que passa uma ideia de progressão e um caminho a ser percorrido, mas isso não é registrado oficialmente, e é mencionado ainda a dificuldade de usar esse tipo de algoritmo em uma base de dados mais complexa [20].

3.3.2 Processo de desenvolvimento ECCOLA

Estrutura de prestação de contas e governança

As três práticas de governança (Governança dos Dados, Governança das Informações, e Governança da Empresa) estão presentes nas atividades dos cartões ECCOLA. Além disso, cada cartão promove uma ou mais práticas de governança nas ações recomendadas. Por exemplo, o card #8 Qualidade de Dados (Figura 3.4) faz perguntas como: Quais são os dados de boa ou má qualidade em seu sistema? Como é avaliada a qualidade e integridade dos dados? Fazer essas perguntas ajuda a garantir a adesão dos desenvolvedores às práticas de governança [68].

A prestação de contas também é um tema abordado na metodologia, no card #2 Explicação (Figura 3.3) por exemplo, ele explica a importância de se ter um sistema confiável, onde as razões por trás das ações são entendíveis ao usuário, e no card #18 Auditabilidade (Figura 3.9) é mostrado a notoriedade de ter mecanismos de antemão para criar um sistema auditável [20] [68].

Escopo e planejamento do projeto

Apresentado no card #0 Análise das partes interessadas (Figura 3.2), onde é destacado que para entender o quadro geral, é importante primeiro entender quem o sistema pode afetar e como. E ainda atenta a também pensar além das partes interessadas diretas e óbvias [5].

Realizar um inventário de dados e uma auditoria de fluxo de dados

Os cards #4 Documentar Trade-offs (Figura 3.3) e #9 Acesso aos Dados (Figura 3.4) englobam esta seção, onde no card #4 destaca a importância de documentar todas as decisões tomadas incluindo as alternativas que foram descartadas, e o card #9 reitera que

além de planejar cuidadosamente quais dados você coleta e como, é também importante planejar como eles podem, e por quem serão usados [66].

Realizar uma análise detalhada de brechas

As análises de possíveis brechas são especificadas no card #20 Minimizando impactos negativos (Figura 3.9), que relata a proeminência de minimizar estes impactos, que não se deve somente ao foto de proteger os dados dos proprietários, mas também pelas consequências financeiras cabíveis. Ainda pede a revisão dos cards #0 Análise das partes interessadas (Figura 3.2), #13 Segurança do sistema (Figura 3.6), #18 Auditabilidade e #19 Capacidade de reparar (Figura 3.9), pois a implementação bem feita e revisada destes pontos resultam em impactos mínimos ao sistema [20] [68].

Desenvolver políticas, procedimentos e processos operacionais

É retratado nos cards #9 Acesso aos Dados (Figura 3.4), #2 Explicação e #5 Rastreabilidade (Figura 3.3), onde o card #9 relata a importância de decidir quais dados serão usados e como, se podem e por quem serão usados. O card #2 esclarece que as razões das ações tomadas tem que ser claro e conciso para ser confiável. Por fim, o card #5 define que a rastreabilidade oferece uma explicabilidade, ou seja, ajuda a entender as decisões e os caminhos percorridos pelo sistema. E esses três em conjunto resultará em práticas de gerenciamento e privacidade, garantindo que esta seção foi levada em consideração [69].

Proteger os dados pessoais por meio de medidas processuais e técnicas

A privacidade dos usuários é o principal objetivo desta metodologia, com isso é apresentado os 3 principais cards relacionada a segurança dos dados, sendo eles: #7 Privacidade e Dados (Figura 3.4), #11 Supervisão humana (Figura 3.5) e #12 Sistema de segurança (Figura 3.6). O card #7 garante a privacidade dos dados respeitando as normas estabelecidas em Lei, o card #11 confirma a necessidade de se ter um responsável/supervisor das ações e decisões tomadas, e o card #12 assegura de que foram avaliadas todas as possíveis formas de ataque a quais o sistema pode ser vulnerável, e com isso tornando o sistema seguro e confiável [20] [68].

Comunicações

Existe um card específico para essa seção no ECCOLA, que é o card #3 Comunicação (Figura 3.3), na prática, comunicação é uma parte essencial da transparência, comunicar os passos dados pelo sistema, pedir o consentimento sempre que possível, entre outras ações garantem um sistema transparente, e um sistema transparente gera confiança [5].

Monitorar e auditar a conformidade

O card #18 Auditabilidade (Figura 3.9) garante que todas as ações foram documentadas e monitoradas para caso seja necessário uma auditoria ou que algum usuário deseje saber para quê e como seus dados foram usados, e o card #19 Capacidade de reparar (Figura 3.9) atende os direitos do proprietário em casos de erros ou impactos negativos, e assim assegurar a confiança dos usuários [20] [66].

3.3.3 Desenvolvimento de IA Sustentável (*Sustainable AI Development*, SAID)

Estrutura de prestação de contas e governança

Não foi constatado na metodologia.

Escopo e planejamento do projeto

Na metodologia o único ponto citado em respeito a essa seção é a necessidade de se ter um sistema ético e justo, em que as 3 camadas, Camada de Tecnologia, Camada Social e a Camada de Governança, juntas dão conta do que pode significar um bom design de Inteligência Artificial. As camadas não devem ser percebidas como existindo em isolamento clínico umas das outras, mas como nós diferentes em uma interação ativa [59].

Realizar um inventário de dados e uma auditoria de fluxo de dados

Apenas é mencionado no SAID que a escolha de algoritmos específicos para a coleta de dados pode fazer uma grande diferença no que diz respeito à sua funcionalidade, mas também à sua transparência. Isso pode ter um grande impacto no acesso à justiça, especialmente quando a base de uma decisão é confusa, e essas escolhas de design afetam como o sistema opera e como pode ser compreendido. Mas em nenhum momento é levado em consideração a importância de se ter documentado essas decisões para que, caso no futuro necessite, realizar uma auditoria do fluxo de dados [59] [20].

Realizar uma análise detalhada de brechas

Não foi constatado na metodologia.

Desenvolver políticas, procedimentos e processos operacionais

Não foi constatado na metodologia.

Proteger os dados pessoais por meio de medidas processuais e técnicas

É assimilada a importância de proteger os proprietários ou os usuários que podem ser afetados com o sistema, mas a única medida de segurança assegurada pela metodologia é a anonimização dos dados, e a garantia de que o usuário tem o controle sobre o sistema, em que as decisões tomadas pelo sistema são influenciadas pelas ações do usuário.

Comunicações

Não foi constatado na metodologia.

Monitorar e auditar a conformidade

Não foi constatado na metodologia.

3.4 Síntese do Capítulo

A partir das análises e pesquisas feitas neste capítulo, foi possível observar que das três metodologias escolhidas, apenas uma se enquadrava em todos os requisitos necessários para se ter um sistema em conformidade com a LGPD, em razão disso, o Capitulo 4 irá apresentar um checklist para que a empresa e seus desenvolvedores se adaptem e atinjam o objetivo de se ter um sistema que respeite a LGPD.

Capítulo 4

Proposta de um Checklist para verificar a conformidade à LGPD

O intuito do checklist (Tabela 4.1) proposto, é auxiliar as organizações e seus colaboradores a implementar a LGPD em seus sistemas com mais facilidade, utilizando-se de perguntas diretas e objetivas para que a prioridade seja sempre a privacidade de seus usuários.

O checklist proposto foi originado das normas da LGPD [6] e das metodologias já existentes na literatura, como os propostos por Marinho et al. [20], Lima et al. [67], Pinheiro [23] e Garcia et al. [19]. Além de adaptações de metodologias relacionadas a GDPR, como os desenvolvidos por Truond et al. [28], Chatzipoulidis et al. [31], e o método ECCOLA [5].

Pensando ainda na documentação de registros e auditorialidade da implementação, foram adicionados algumas colunas relevantes para registrar as atividades e decisões tomadas em relação ao checklist, não só para a consulta da equipe de desenvolvedores, supervisores, mas também para o usuário final saber quais normas em relação a LGPD aquele sistema implementou, e se caso não tenha implementado, o porquê dessa decisão [20]. A coluna "Foi implementado?" (Tabela 4.1) foi idealizada para informar se aquela seção foi implementada ou não. A coluna "Justificativa" (Tabela 4.1) foi inserida com o intuito de informar o por quê aquela seção foi implementada ou não, pois a explicabilidade é uma questão indispensável quando se deseja um sistema transparente, confiável e seguro. E por fim, a coluna "Responsável" (Tabela 4.1), que tem como objetivo registrar quem foi o responsável por implementar ou não aquela seção, já que isso pode ser questionado em uma futura auditoria, ou caso algum usuário se sinta lesado por algum impacto negativo criado pelo sistema desenvolvido, onde o responsável registrado junto a empresa serão os encarregados de sanar todas as dúvidas perante a lei ou não [19].

Questões a serem consideradas	**Foi implementado?**	**Justificativa**	**Responsável**
Foi definido um responsável pela empresa para a implementação da LGPD e responder por possíveis consequências?			
Foi definido a finalidade do sistema?			
Foi definido quem o sistema pode afetar e como?			
É pedido o consentimento do usuário para o uso de seus dados?			
Cada decisão do seu sistema pode ser entendida pelo usuário final?			
Foram avaliados os dados que serão mantidos, sua origem e a base legal para seu processamento?			
As decisões tomadas no pro cessamento dos dados é registrada para uma possível auditoria no futuro?			
As decisões tomadas pelo sistema são registradas afim de uma possível auditoria no futuro?			
O sistema foi testado em diferentes cenários e seus resultados foram documentados?			
A qualidade/integridade dos dados é garantida pelo sistema?			

É definido quem, como, e em quais circunstâncias os dados podem ser acessados?			
É registrado/documentado quem acessou esses dados?			
É comunicado ao usuário que uma decisão, conteúdo ou consequência é o resultado de uma decisão algorítmica?			
Foi realizado uma análise para detectar possíveis brechas no sistema?			
Foi avaliado o impacto social do uso do sistema além dos usuários finais?			
Os dados são protegidos por meio de medidas processuais e técnicas, tais como criptografia, anonimização e pseudonimização?			
Os funcionários da empresa receberam treinamento sobre os princípios básicos da LGPD?			
Foi definido datas regulares de auditoria do processamento de dados e controles de segurança?			
Se um imprevisto negativo acontecer os afetados podem buscar reparação?			

O usuário é informado de que pode pedir uma revisão de como, quando e por quem seus dados foram usados a qualquer momento?			

Tabela 4.1: Checklist para desenvolver um sistema em conformidade com a LGPD

4.1 Validação do checklist

Para devidamente validar o checklist proposto, foi aplicado um *survey* (Tabela 4.2) de forma online, com as questões do checklist em que as respostas são relacionadas ao quão importante ele acredita que é a questão levantada para o desenvolvimento de sistemas em conformidade com a LGPD. Após isso, são feitas as perguntas abaixo para levantar se o método criado juntamente com suas especialidades realmente fazem ele se tornar confiável, seguro e auditado.

1. Com esse checklist para implementar as normas em seu sistema, você confiaria de que a pessoa responsável preencheria todos os campos de forma verdadeira e justa? Justifique.
2. Você se sentiria mais seguro caso o responsável precisasse assinar um termo de responsabilidade e compromisso para o preenchimento do checklist? Justifique.
3. Caso uma questão do checklist seja de difícil implementação no sentido financeiro, a sua empresa ainda empregaria? Justifique.
4. Você, como usuário, se sentiria mais seguro caso tivesse acesso a informações tais como: processo de implementação da LGPD? quando foi implementado? Quem foi o responsável por implementar? Caso não tenha sido implementado, o porquê dessa decisão? Justifique.
5. Pensando ainda no lado do usuário, você confiaria de que as respostas do checklist condiz com a realidade de implementação da empresa? Caso não, o que lhe deixaria mais seguro em relação às respostas?
6. Você acha necessário a organização ter o costume de realizar registros de Logs dos tratamentos executados nos dados pessoais? Justifique.

A primeira pergunta foi elaborada pensando na veracidade dos fatos descritos pela pessoa responsável por preencher o checklist, pois se a realidade não condiz com o que foi registrado, o método então perde o seu objetivo, porque não será mais confiável e muito menos seguro.

A segunda pergunta é recorrente da primeira, onde é perguntado se caso fosse elaborado um documento em forma de termo de responsabilidade, em que o responsável deveria assinar e assim de forma legal registrar de que tudo que for descrito no checklist convém da realidade dos fatos. Por isso, é perguntado se o entrevistado se sentiria mais seguro em relação as respostas e por quê.

A pergunta 3 tem o objetivo de saber o quão compromissado a empresa está com a LGPD, e se estão por dentro das consequências em caso de não cumprimento, pois a multa [6] provavelmente será bem acima dos custos envolvidos em sua implementação.

A quarta pergunta envolve a questão da divulgação do checklist por parte da empresa, onde o usuário poderia facilmente ter acesso a esse processo de desenvolvimento e ter conhecimento de como seus dados foram ou estão sendo tratados pela empresa, além de saber quem foi o responsável por isso.

Reforçando a questão da confiança e da transparência, a quinta pergunta novamente traz a questão da veracidade dos registros, seu principal intuito é saber como o usuário se sentiria seguro em relação as respostas do checklist.

Por fim, a auditabilidade, a última pergunta tem o objetivo de saber se o usuário ou a empresa sabem a importância de se ter todos os passos registrados, pois isso gera explicabilidade, que tem como consequência a transparência [66], o que deixa o usuário mais confiante em relação ao sistema, pois ele sabe que se caso algo aconteça a empresa terá todas decisões tomadas documentadas.

4.2 Resultados do Survey

O formulário para validar a proposta do checklist apresentado na Tabela 4.2 ficou disponível por 25 dias de forma online, e seu tempo médio de resposta foi de 16 minutos. 31 profissionais da área de tecnologia no total responderam ao questionário, sendo a maioria trabalhando em organizações privadas, e quase a metade, 48,4%, ocupando o cargo de Desenvolvedor.

Com os resultados obtidos, é possível perceber que 100% dos participantes acreditam que definir a finalidade do sistema (Questão 2 (Q02) da Figura 4.1), quem ele pode afetar e como (Questão 3 (Q03) da Figura 4.1), e que deve-se sempre pedir o consentimento do usuário para o uso de seus dados (Questão 4 (Q04) da Figura 4.2), o que reforça a importância da preparação, você ter bem definido como será o sistema final de antemão

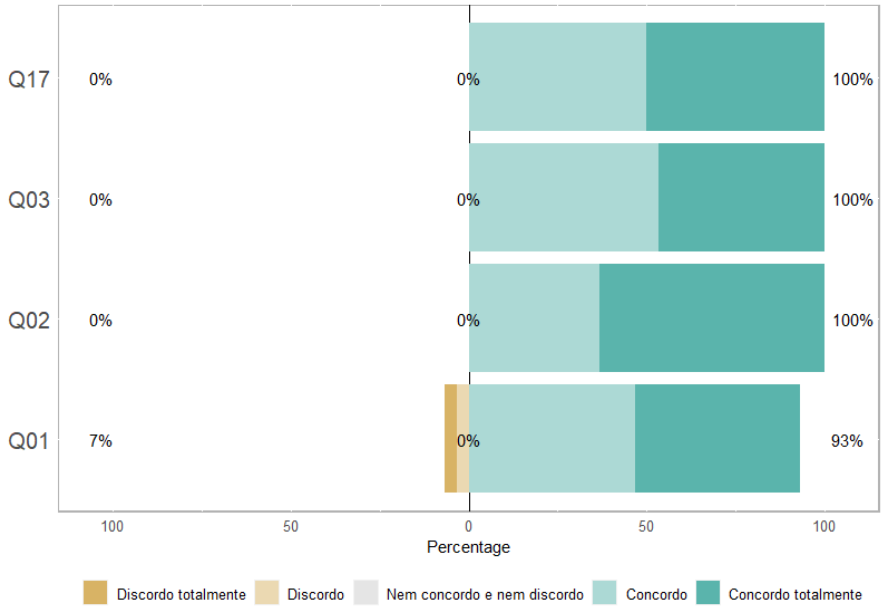

Figura 4.1: Questões relacionadas a etapa de preparação para a receber a LGPD

lhe dá o conhecimento de saber exatamente onde é necessário ter mais atenção, onde acontecerá o uso desses dados, e com isso, criar um termo de consentimento bem elaborado e detalhado para que o usuário saiba precisamente a finalidade do uso dessas informações, e consequentemente, construir uma transparência que traz segurança e confiabilidade.

Em relação a proteção dos dados, compostas pelas Questões Q04, Q06, Q10, Q11 e Q16 da Figura 4.2, foi possível enxergar que em quase unanimidade, aproximadamente 97% (realizando a média aritmética entre elas) concordam que os dados devem receber total atenção em um sistema que deseja ter conformidade com a LGPD, desde o consentimento para o seu uso, definir quais dados serão usados, manter uma qualidade/integridade destas informações, quem e em quais circunstâncias poderão ser acessados, além da aplicação de medidas processuais e técnicas para garantir que estão e serão armazenados e processados de forma segura, e como mencionado por Marinho et al. [20], estes realmente são passos muito importantes para respeitar tanto a LGPD, quanto a ética e a privacidade.

A respeito da documentação e auditoria dos processos, os resultados obtidos foram bem otimistas, em que 100% concordam que realizar testes e documentar seus resultados (Questão 9 (Q9) da Figura 4.3) é extremamente importante, e também documentar as decisões tomadas pelo no processamento de dados e no decorrer do sistema (Q07 e Q08 da Figura 4.3). O fato de nem todos compreenderem que realizar registros de quem ou quando os dados foram acessados (Q12 da Figura 4.3) é preocupante, já que deveria ser unanimidade nas respostas, pois essa informação garante que os dados não tenham sido acessados indevidamente ou por pessoas não autorizadas, e em caso de necessidade de auditoria, essas informações podem esclarecer questionamentos relacionados a privacidade

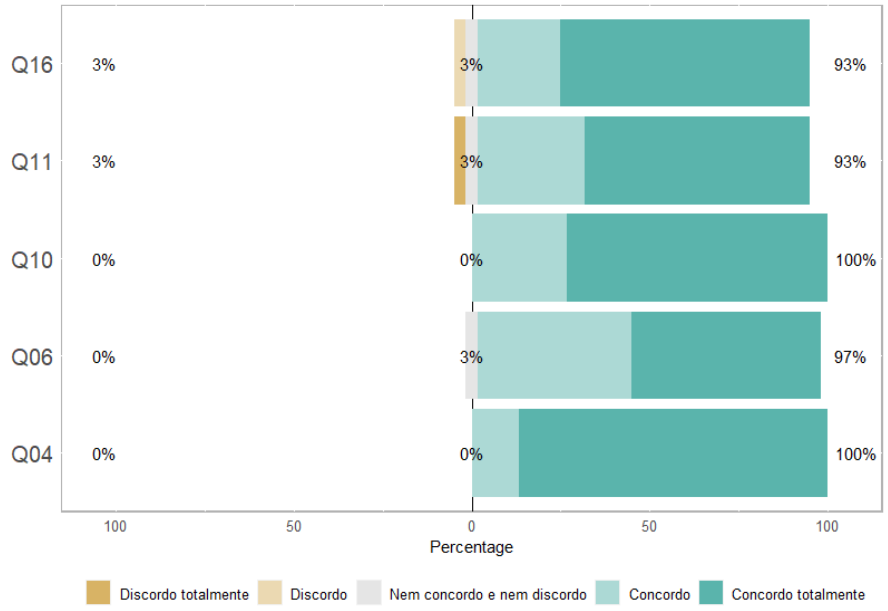

Figura 4.2: Questões relacionadas as etapas de proteção dos dados

e segurança do usuário.

Os resultados obtidos relacionados a comunicação ao usuário relatam um ponto negativo, já que apenas 70% dos participantes reconhecem como importante que as decisões do sistema podem ser entendidas pelo usuário (Q05 da Figura 4.4), um sistema que se auto explica e de fácil entendimento remete a confiança e transparência, dois fatores imprescindíveis para se ter software ético. E ainda, somente 57% entendem a notoriedade de se comunicar ao usuário que uma consequência, decisão ou conteúdo é resultado de uma decisão algorítmica (Q13 da Figura 4.4), a não exposição destes fatos pode gerar dúvidas e desconfianças ao usuário, por exemplo, caso a empresa tenha acesso a dados sensíveis tais como étnicos, uma decisão baseada nestes dados e que não seja comunicado que é o resultado de um processo algorítmicos pode levar o usuário a entender que alguém julgou estes dados unicamente, gerando desconforto e até se sentir distinguido, e como consequência podendo ser autuado com base no Artigo 6° inciso IX da LGPD [6], que retrata a não discriminação, isto é, a impossibilidade de realização do tratamento para fins discriminatórios, ilícitos ou abusivos.

Nas questões atribuídas a etapa de reparação, com os resultados apresentados na Figura 4.5, observou-se que 87% concordam que que é importante o usuário poder buscar reparação caso algum imprevisto negativo aconteça (Q19 da Figura 4.5), a reparação é uma peça chave para garantir a confiança do usuário, deixar claro que ele será amparado pela empresa caso algo aconteça. E ainda, 93% confirmam de que atender o desejo do usuário de saber o que, pra quê e como seus dados foram usados/processados (Q20 da Figura 4.5) é essencial, como é explicitado por Lima et al. [67] e Pinheiro [23].

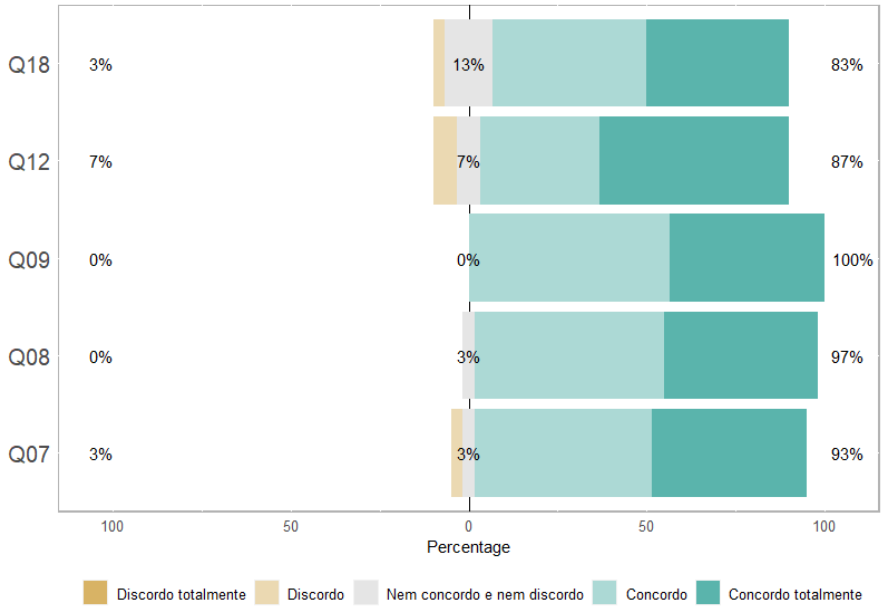

Figura 4.3: Questões relacionadas a documentação/auditoria dos processos e decisões tomadas

Com o intuito direto de validar o checklist, foram ainda feitas 6 perguntas para validar seu processo de implementação. A primeira pergunta (Questão 21 (Q21) da Tabela 4.2) é relacionada a confiabilidade das respostas preenchidas pelo responsável no checklist, em que de 31 respostas, 20 foram positivas, e os principais argumentos levantados para tal resposta foram de que a pessoa responsável por isso com certeza teria o conhecimento necessário para saber as consequências cabíveis, e então saberia a importância de seguir em conformidade com a ética e a LGPD.

Na segunda pergunta (Questão 22 (Q22) da Tabela 4.2), é um complemento a pergunta anterior, em que é questionado aos participantes se a elaboração de um termo, e este sendo assinado pelos responsáveis ao preenchimento o deixariam mais seguros, mas em sua maioria foi considerado desnecessário, pois o fato de se ter a LGPD já cumpriria a função de um termo, tendo que responder legalmente em caso do não cumprimento de suas normas.

Em vista do ponto financeiro, a terceira pergunta (Questão 23 (Q23) da Tabela 4.2) contesta se a empresa em que o colaborador trabalharia implementaria uma questão mesmo que tivesse um alto custo, e os resultados foram bastante positivos, pois as respostas reforçaram de que a segurança dos dados é de extrema importância mesmo tendo um alto custo, e ainda foi citado várias vezes a questão de que o custo de implementação é de pequena importância se comparado as multas e consequências do não cumprimento das normas da LGPD.

A quarta e a sexta pergunta (Questão 24 (Q24) e 26 (Q26) respectivamente da Tabela

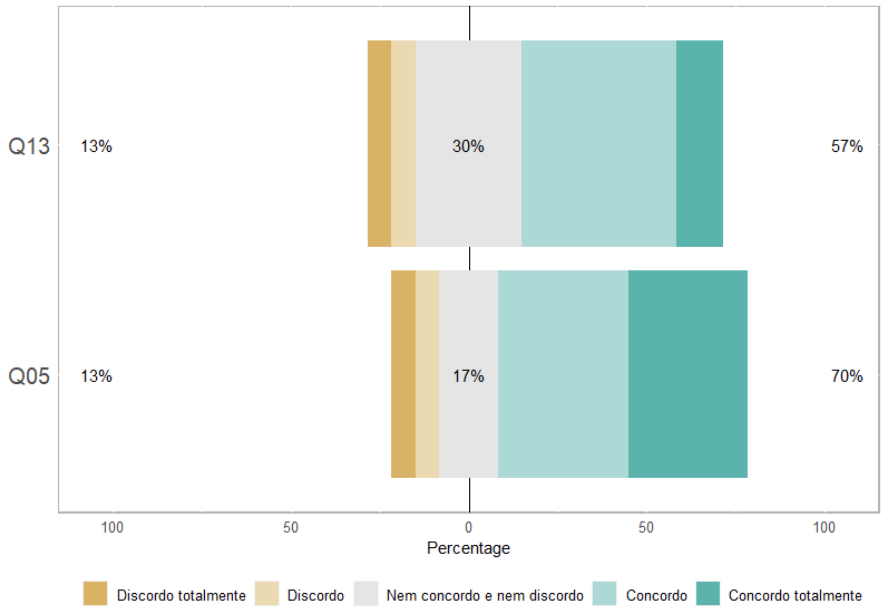

Figura 4.4: Questões relacionadas aos processos de comunicação

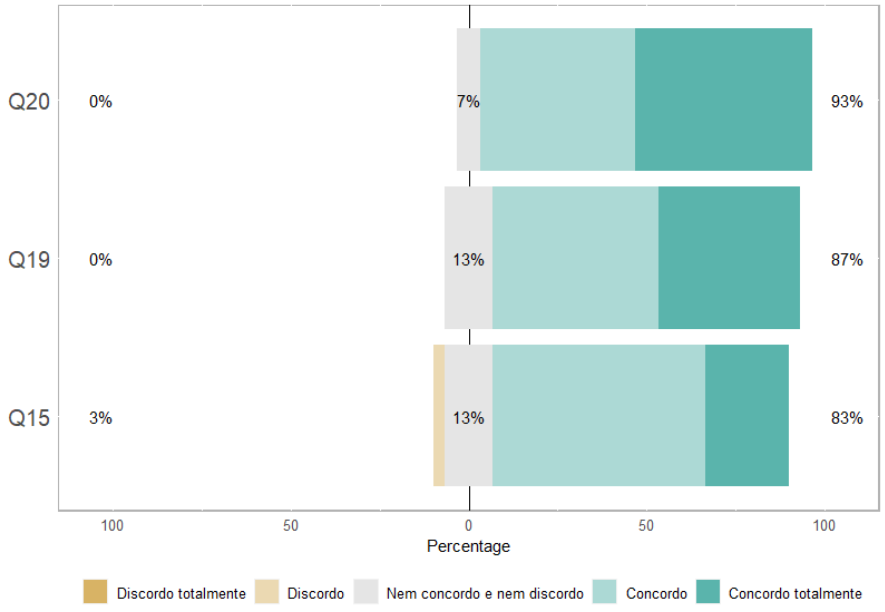

Figura 4.5: Questões relacionadas aos processos de reparação

4.2) são diretamente associadas as colunas “Foi implementado?”, “Justificativa” e “Responsável” do checklist (Tabela 4.1), em que é questionado a importância do usuário ter acesso a informações tais como se uma etapa foi implementada, como foi implementada, quem foi o responsável pela implementação, e a necessidade de realizar registros dos tratamentos executados, que são os objetivos destas colunas presentes no checklist, e por quase unanimidade foi confirmado que é fundamental para as organizações.

Por fim, foi levantado também, se como usuário, o que lhe deixaria mais seguro em relação as respostas (Questão 25 (Q25) da Tabela 4.2), observou-se questões relevantes, como a criação de um selo de qualidade reconhecido pela comunidade, ou se o checklist fosse implementado de modo terceirizado, pois assim reduziria o risco da empresa omitir algo para atrair clientes.

Q01	É importante definir um responsável pela empresa para a implementação da LGPD e responder por possíveis consequências?
Q02	É importante definir a finalidade do sistema?
Q03	É importante definir quem o sistema pode afetar e como?
Q04	É importante pedir o consentimento do usuário para o uso de seus dados?
Q05	É importante que cada decisão do seu sistema possa ser entendida pelo usuário final?
Q06	É importante avaliar quais os dados que serão mantidos, sua origem e a base legal para seu processamento?
Q07	É importante que as decisões tomadas no processamento dos dados seja registrada para uma possível auditoria no futuro?
Q08	É importante que as decisões tomadas pelo sistema sejam registradas afim de uma possível auditoria no futuro?
Q09	É importante que o sistema seja testado em diferentes cenários e seus resultados sejam documentados?
Q10	É importante que a qualidade/integridade dos dados seja garantida pelo sistema?
Q11	É importante que seja definido quem, como, e em quais circunstâncias os dados podem ser acessados?
Q12	É importante que seja registrado/documentado quem acessou esses dados?
Q13	É importante que comunicado ao usuário que uma decisão, conteúdo ou consequência é o resultado de uma decisão algorítmica?
Q14	É importante que seja realizado uma análise para detectar possíveis brechas no sistema?

Q15	É importante que seja avaliado o impacto social do uso do sistema além dos usuários finais?
Q16	É importante que os dados sejam protegidos por meio de medidas processuais e técnicas, tais como criptografia, anonimização e pseudonimização?
Q17	É importante que os funcionários da empresa recebam treinamento sobre os principios básicos da LGPD?
Q18	É importante que sejam definidos datas regulares de auditoria do processamento de dados e controles de segurança?
Q19	É importante que se um imprevisto negativo acontecer os afetados possam buscar reparação?
Q20	É importante que o usuário seja informado de que pode pedir uma revisão de como, quando e por quem seus dados foram usados a qualquer momento?
Q21	Com esse checklist para implementar as normas em seu sistema, você confiaria de que a pessoa responsável preencheria todos os campos de forma verdadeira e justa? Justifique.
Q22	Você se sentiria mais seguro caso o responsável precisasse assinar um termo de responsabilidade e compromisso para o preenchimento do checklist? Justifique.
Q23	Caso uma questão do checklist seja de difícil implementação no sentido financeiro, a sua empresa ainda empregaria? Justifique.
Q24	Você, como usuário, se sentiria mais seguro caso tivesse acesso a informações tais como: processo de implementação da LGPD? Quando foi implementado? Quem foi o responsável por implementar? Caso não tenha sido implementado, o porquê dessa decisão? Justifique.
Q25	Pensando ainda no lado do usuário, você confiaria de que as respostas do checklist condiz com a realidade de implementação da empresa? Caso não, o que lhe deixaria mais seguro em relação às respostas?
Q26	Você acha necessário a organização ter o costume de realizar registros (logs) dos tratamentos executados nos dados pessoais? Justifique.

Tabela 4.2: Questões apresentadas no formulário para a validação do checklist

4.3 Ameaças e Limitações para Validação deste Trabalho

Uma ameaça possível a este trabalho são futuras mudanças na LGPD, ementas que podem mudar o conceito, objetivo ou intuito de suas normas, ou até, a sua não obrigatoriedade. Por ser uma Lei recente, criada em 2018 e entrando em vigor somente em 2021 [6], então provavelmente sofrerá algumas mudanças, podendo até ser acrescentado novas normas.

O checklist apresentado possui algumas limitações, como o usuário não atender a todas questões, seja por questões pessoais, financeiras ou estratégicas da empresa. Outra limitação é o não preenchimento de algumas colunas, ou o preenchimento incorreto. Além disso, uma outra possível implicação é a omissão de fatos por parte do responsável de cada questão, seja por motivos pessoais ou jurídicos.

Capítulo 5

Conclusão

O trabalho permitiu, depois de investigar na literatura, perceber que poucas metodologias auxiliariam as empresas de tecnologia e seus colaboradores a implementar sistemas em conformidade com a LGPD, e em sua maioria percebeu-se a falta de comprometimento com a ética e privacidade dos usuários, por isso, foi elaborado o checklist para ajudar os interessados a desenvolver estes sistemas respeitando a LGPD de um modo direto e prático.

O checklist final então, constitui de seções que devem conter em seu sistema, onde cada questão levantada irá acatar uma parte das exigências da LGPD, assim, resultando em um sistema ético, transparente e confiável. Além das questões, foram criadas colunas para garantir a rastreabilidade destas decisões tomadas, com campos para registrar a justificativa e o responsável por cada implementação.

As metodologias encontradas na literatura contribuíram bastante para o andamento deste trabalho, seja o checklist feito por Marinho et al. [20] que de forma sucinta abrangeu grande parte da LGPD, o método ECCOLA [68] que proporcionou perceber pontos específicos relacionados a algoritmos automatizados e processamento de dados em relação a ética, e Lima et al. [67] que abordou a questão empresarial, seja as dificuldades de implementação, ou os riscos financeiros do não cumprimento da Lei.

Por fim, como trabalho futuro, seria a aplicação prática desta metodologia em sistemas Machine Learning ou de Inteligência Artificial, ou em empresas de tecnologia, para avaliar o seu tempo de implementação, o impacto causado na empresa, e o seu custo beneficio, para que assim tenha mais dados concretos de seu funcionamento, proporcionando aos seus possíveis usuários as dificuldades e/ou pontos que devem ser alterados ou aprimorados.

Referências

[1] Mendes, Laura Schertel, Danilo Doneda, Ingo Wolfgang Sarlet, Otavio Luiz Rodrigues Jr. e Bruno Bioni: *Tratado de Proteção de Dados Pessoais*, volume 1. GEN - Grupo Editorial Nacional, 2020, ISBN 978-85-309-9219-4. https://integrada.minhabiblioteca.com.br/#/books/9788530992200/. ix, 1, 4, 8, 9, 13, 14, 15, 16, 17

[2] Silva, Fabrício Machado da, Maikon Lucian Lenz, Pedro Henrique Chagas Freitas e Sidney Cerqueira Bispo dos Santos: *Inteligência Artificial*, volume 1. Grupo A, 2019. https://integrada.minhabiblioteca.com.br/#/books/9788595029392/, Acessado em 12/05/2021. ix, 24, 28

[3] Lenz, Maikon Lucian, Fabiano Berlink Neumann, Rodrigo Santarelli e Douglas Salvador: *Fundamentos de Aprendizagem de Máquina*, volume 1. Grupo GEN, 2020. https://integrada.minhabiblioteca.com.br/#/books/9786556900902/, Acessado em 14/05/2021. ix, 25, 27, 28, 29, 30

[4] Goldsteen, Abigail, Gilad Ezov, Ron Shmelkin, Micha Moffie e Ariel Farkash: *Data minimization for gdpr compliance in machine learning models*. Cornell University, 1, 2020. https://arxiv.org/abs/2008.04113. ix, 34, 36, 37, 38, 39, 44, 45, 48

[5] Vakkuri, Ville, Kai Kristian Kemell e Pekka Abrahamsson: *Eccola - a method for implementing ethically aligned ai systems*. IEEE Explorer, 1:195–204, 2020. https://ieeexplore.ieee.org/document/9226275, Acessado em 23/04/2021. ix, 3, 40, 41, 42, 43, 44, 45, 49, 50, 53

[6] Brasil: *Lei geral de proteção de dados pessoais (lgpd)*. Diário Oficial [da] República Federativa do Brasil, 2019. http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709compilado.htm. 1, 2, 3, 4, 5, 6, 8, 9, 10, 12, 13, 14, 15, 16, 22, 46, 53, 57, 59, 64

[7] Parliament, European e Council of the European Union: *General data protection regulation*. European Commission, 1:99, 2018. https://gdpr-info.eu/. 1, 2, 3, 9, 10, 11, 14, 17, 18, 19, 20, 21, 22

[8] Communications Networks, Directorate General for, Content, Technology (European Commission) e Grupa ekspertów wysokiego szczebla ds. sztucznej inteligencji: *Orientações éticas para uma IA de confiança*, volume 1. Comissão Europeia, 2019, ISBN 978-92-76-11998-2. https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1, Acessado em 21/04/2021. 1

[9] Ryan, Mark e Bernd Carsten Stahl: *Artificial intelligence ethics guidelines for developers and users: clarifying their content and normative implications*, volume 19. Journal of Information, Communication and Ethics in Society, 2020. https://www.emerald.com/insight/content/doi/10.1108/JICES-12-2019-0138/full/html, Acessado em 28/04/2021. 2

[10] Mitrou, Lilian: *Data protection, artificial intelligence and cognitive services: Is the general data protection regulation (gdpr) 'artificial intelligence-proof'?* SSRN Electronic Journal, 1:90, 2018. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3386914, Acessado em 01/05/2021. 2

[11] Sartor, Giovanni, European Parliament, European Parliamentary Research Service e Scientific Foresight Unit: *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence: study.* European Union, 2020, ISBN 9789284667710. http://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU(2020)641530_EN.pdf, acesso em 2021-05-02, Acessado em 01/05/2021. 2, 3, 18, 19, 20, 33

[12] Lima, Taisa Maria Macena de e Maria de Fátima Freire de Sá: *Inteligência artificial e Lei Geral de Proteção de Dados Pessoais: o direito à explicação nas decisões automatizadas.* Revista Brasileira de Direito Civil, 26(04):20, 2020, ISSN 25944932, 23586974. https://rbdcivil.ibdcivil.org.br/rbdc/index, acesso em 2021-05-02, Acessado em 01/05/2021. 2

[13] O'Reilly (Firm): *Big data now: current perspectives from O'Reilly Media.* O'Reilly Media, Inc, 2012, ISBN 978-1-4493-5671-2 978-93-5023-970-4. https://www.oreilly.com/data/free/files/big-data-now-2012.pdf, Acessado em 07/05/2021. 2

[14] Dias, Guilherme Ataídes e Américo Augusto Nogueira Vieira: *Big data: questões éticas e legais emergentes.* Ciência da Informação, 42(2), 2015. http://revista.ibict.br/ciinf/article/view/1380, Acessado em 07/05/2021. 2

[15] Kuner, Christopher, Dan Jerker B. Svantesson, Fred H. Cate, Orla Lynskey e Christopher Millard: *Machine learning with personal data: is data protection law smart enough to meet the challenge?* International Data Privacy Law, 7(1):1–2, 2017, ISSN 2044-3994. https://academic.oup.com/idpl/article-pdf/7/1/1/14043502/ipx003.pdf, Acessado em 07/05/2021. 3

[16] Goodfellow, Ian J, Jonathon Shlens e Christian Szegedy: *Explaining and harnessing adversarial examples.* arXiv preprint arXiv:1412.6572, página 11, 2014, ISSN 1412.6572. https://arxiv.org/pdf/1412.6572.pdf, Acessado em 07/05/2021. 3

[17] Szegedy, Christian, Wojciech Zaremba, Ilya Sutskever, Joan Bruna, Dumitru Erhan, Ian Goodfellow e Rob Fergus: *Intriguing properties of neural networks.* arXiv preprint arXiv:1312.6199, página 10, 2014, ISSN 1312.6199. https://arxiv.org/pdf/1312.6199.pdf, Acessado em 07/05/2021. 3

[18] Xue, Mingfu, Chengxiang Yuan, Heyi Wu, Yushu Zhang e Weiqiang Liu: *Machine learning security: Threats, countermeasures, and evaluations*. IEEE Access, 8:23, 2020. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9064510, Acessado em 07/05/2021. 3

[19] Garcia, Lara Rocha, Edson Aguilera-Fernandes, Rafael Augusto Moreno Gonçalves e Marcos Ribeiro Pereira-Barretto: *Lei Geral de Proteção de Dados (LGPD): Guia de implantação*, volume 1. Editora Blucher, 2020, ISBN 978-65-5506-016-4. https://integrada.minhabiblioteca.com.br/#/books/9786555060164/. 3, 8, 12, 13, 53

[20] Marinho, Fernando: *Os 10 Mandamentos da LGPD - Como Implementar a Lei Geral de Proteção de Dados em 14 Passos*, volume 1. GEN - Grupo Editorial Nacional, 2020, ISBN 978-85-97-02599-6. https://integrada.minhabiblioteca.com.br/#/books/9788597026009/. 4, 21, 22, 47, 48, 49, 50, 51, 53, 58, 65

[21] Prodanov, Cleber Cristiano e Ernani Cesar de Freitas.: *Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico*, volume 2. Universidade Freevale, 2013, ISBN 978-85-7717-158-3. https://aedmoodle.ufpa.br/pluginfile.php/291348/mod_resource/content/3/2.1-E-book-Metodologia-do-Trabalho-Cientifico-2.pdf. 5, 6, 7

[22] Pinheiro, Patricia Peck: *Segurança Digital - Proteção de Dados nas Empresas*, volume 2. GEN - Grupo Editorial Nacional, 2020, ISBN 978-85-97-02639-9. https://integrada.minhabiblioteca.com.br/#/books/9788597026405/. 10

[23] Pinheiro, Patricia Peck: *Proteção de Dados Pessoais: Comentários a Lei 13.709/2018 (LGPD)*, volume 2. Editora Saraiva, 2020, ISBN 978-85-536-1362-5. https://integrada.minhabiblioteca.com.br/#/books/9788553613625/cfi/83!/4/4@0.00:47.4, Acessado em 21/04/2021. 11, 13, 14, 16, 17, 53, 59

[24] Brasil: *Dispõe sobre a arbitragem*. Diário Oficial [da] República Federativa do Brasil, 1996. http://www.planalto.gov.br/ccivil_03/LEIS/L9307.htm. 12

[25] Brasil: *Guia de boas práticas – lei geral de proteção de dados (lgpd)*. Comitê Central de Governançã de Dados. Secretaria de Governo Digital, 1–65:69, 2020. https://www.gov.br/governodigital/pt-br/governanca-de-dados/guias-operacionais-para-adequacao-a-lgpd, Acessado em 21/04/2021. 15, 16, 17, 48

[26] Team, I.T.G.P.: *EU General Data Protection Regulation (GDPR) – An implementation and compliance guide, fourth edition*. IT Governance Publishing, 2020, ISBN 9781787782501. https://books.google.com.br/books?id=LicDEAAAQBAJ. 18, 19, 20

[27] De Hert, Paul, Vagelis Papakonstantinou, Gianclaudio Malgieri, Laurent Beslay e Ignacio Sanchez: *The right to data portability in the gdpr: Towards user-centric interoperability of digital services*. Computer Law & Security Review, 34(2):193–203, 2018, ISSN 0267-3649. https://www.sciencedirect.com/science/article/pii/S0267364917303333. 18, 20

[28] Truong, Nguyen, Kai Sun, Siyao Wang, Florian Guitton e YiKe Guo: *Privacy preservation in federated learning: An insightful survey from the gdpr perspective*. Computers & Security, 110:102402, 2021, ISSN 0167-4048. https://www.sciencedirect.com/science/article/pii/S0167404821002261. 18, 20, 53

[29] Li, He, Lu Yu e Wu He: *The impact of gdpr on global technology development*. Journal of Global Information Technology Management, 22(1):1–6, 2019. https://doi.org/10.1080/1097198X.2019.1569186. 19, 20

[30] Gal, Michal S e Oshrit Aviv: *The Competitive Effects of the GDPR*. Journal of Competition Law & Economics, 16(3):349–391, maio 2020, ISSN 1744-6414. https://doi.org/10.1093/joclec/nhaa012. 19, 20

[31] Chatzipoulidis, Aristeidis, Theodosios Tsiakis e Theodoros Kargidis: *A readiness assessment tool for gdpr compliance certification*. Computer Fraud & Security, 2019(8):14–19, 2019, ISSN 1361-3723. https://www.sciencedirect.com/science/article/pii/S1361372319300867. 19, 21, 53

[32] Iramina, Aline: *Rgpd v. lgpd: AdoÇÃo estratÉgica da abordagem responsiva na elaboraÇÃo da lei geral de proteÇÃo de dados do brasil e do regulamento geral de proteÇÃo de dados da uniÃo europeia*. Adoção Estratégica da Abordagem Responsiva na Elaboração da Lei Geral de Proteção de Dados do Brasil e do Regulamento Geral de Proteção de Dados da União Europeia. Revista de Direito, Estado e Telecomunicações, páginas 91–117, 2020. https://core.ac.uk/reader/337598120, Acessado em 18/05/2021. 22

[33] Coppin, Ben: *Inteligência Artificial*, volume 1. Grupo GEN, 2010. https://integrada.minhabiblioteca.com.br/#/books/9788595156104/, Acessado em 12/05/2021. 23

[34] Russel, Stuart e Peter Norvig: *Inteligência Artificial*, volume 3. Grupo GEN, 2013. https://integrada.minhabiblioteca.com.br/#/books/9788595156104/, Acessado em 12/05/2021. 23, 24, 25

[35] Siau, Keng e Weiyu Wang: *Artificial Intelligence (AI) Ethics: Ethics of AI and Ethical AI*. Journal of Database Management, 31(2):74–87, 2020, ISSN 1063-8016, 1533-8010. http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/JDM.2020040105, acesso em 2021-05-18. 23

[36] Morais, Isabelly Soares de, Priscila de Fátima Gonçalves, Cleverson Lopes Ledur, Ramiro Sebastião Córdova Junior, Maurício de Oliveira Saraiva e Sandra Rovena Frigeri: *Introdução a Big Data e Internet das Coisas (IoT)*. Grupo A, 20181, ISBN 978-85-9502-764-0. https://integrada.minhabiblioteca.com.br/#/books/9788595027640/. 23, 25, 29, 30, 32

[37] Faceli, Katti, Ana Carolina Lorena, João Gama, Tiago Agostinho de Almeida e André C. P. L. F. de Carvalho: *Inteligência Artificial - Uma Abordagem de Aprendizado de Máquina*, volume 2. Grupo GEN, 2021. https://integrada.minhabiblioteca.com.br/#/books/9788521637509/, Acessado em 12/05/2021. 25, 27, 28, 30

[38] Ertel, Wolfgang: *Introduction to Artificial Intelligence*. Undergraduate Topics in Computer Science. Springer International Publishing : Imprint: Springer, Cham, 2nd ed. 2017 edição, 2017, ISBN 9783319584874. 25

[39] Floridi, Luciano, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge, Robert Madelin, Ugo Pagallo, Francesca Rossi, Burkhard Schafer, Peggy Valcke e Effy Vayena: *AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations*. Minds and Machines, 28(4):689–707, 2018, ISSN 1572-8641. https://doi.org/10.1007/s11023-018-9482-5. 26

[40] Jobin, Anna, Marcello Ienca e Effy Vayena: *The global landscape of AI ethics guidelines*. Nature Machine Intelligence, 1(9):389–399, setembro 2019, ISSN 2522-5839. http://www.nature.com/articles/s42256-019-0088-2, acesso em 2021-05-18. 26

[41] Felzmann, Heike, Eduard Fosch Villaronga, Christoph Lutz e Aurelia Tamò-Larrieux: *Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns*. Big Data & Society, 6(1):2053951719860542, 2019. https://doi.org/10.1177/2053951719860542. 27

[42] Coelho, André L.V., Everlândio Fernandes e Katti Faceli: *Multi-objective design of hierarchical consensus functions for clustering ensembles via genetic programming*. Decision Support Systems, 51(4):794–809, 2011, ISSN 01679236. https://linkinghub.elsevier.com/retrieve/pii/S0167923611000431, acesso em 2021-05-14. 27, 30

[43] Matarić, Maja J.: *Introdução á robótica*. Editora Blucher, 2nd ed. 2017 edição, 2014, ISBN 9788521208549. https://integrada.minhabiblioteca.com.br/#/books/9788521208549/, acesso em 2021-05-18. 28, 29

[44] Xie, Ming, Hui Chen e Zhencheng Hu: *New Foundation of Artificial Intelligence*. WorldScientific, 2021, ISBN 9789814271639. https://doi.org/10.1142/7264. 29

[45] Gomes, Elisabeth e Fabiane Braga: *Inteligencia Competitiva Tempos Big Data*, volume 1. Editora Alta Books, 2017, ISBN 9788550804101. https://integrada.minhabiblioteca.com.br/#/books/9788550804101/, acesso em 2021-05-19. 31, 32

[46] Zolynski, Célia: *Os big data e os dados pessoais entre os princípios da proteção e da inovação*. Revista de Direito, Estado e Telecomunicações. Brasília, 12, 2020. https://doi.org/10.26512/lstr.v12i1.30007. 32

[47] Ribeiro, Ana Lídia Lira: *Discriminação em algoritmos de inteligência artificial: uma análise acerca da LGPD como instrumento normativo mitigador de vieses discriminatórios*. Universidade Federal do Ceará, 1, 2021. http://www.repositorio.ufc.br/handle/riufc/57947, acesso em 2021-05-19. 33

[48] Gruschka, Nils, Vasileios Mavroeidis, Kamer Vishi e Meiko Jensen: *Privacy issues and data protection in big data: A case study analysis under gdpr*. IEEE Explorer, 1:5027–5033, 2018. https://ieeexplore.ieee.org/abstract/document/8622621. 33, 34

[49] Bonatti, Piero A. e Sabrina Kirrane: *Big data and analytics in the age of the gdpr.* IEEE Explorer, 1:7–16, 2019, ISSN 2642-7273. https://ieeexplore.ieee.org/abstract/document/8818223. 34

[50] Yip, Camille, Nian LinReena Han e BanLeong Sng: *Legal and ethical issues in research.* Indian Journal of Anaesthesia, 60(9):684, 2016, ISSN 0019-5049. http://www.ijaweb.org/text.asp?2016/60/9/684/190627, acesso em 2021-09-21. 34

[51] Li, Guiqin, Xuechao Deng, Zhiyuan Gao e Feng Chen: *Analysis on Ethical Problems of Artificial Intelligence Technology.* ACM Press, páginas 101–105, 2019. http://dl.acm.org/citation.cfm?doid=3341042.3341057, acesso em 2021-09-21. 35

[52] Kwan, David, Luiz Marcio Cysneiros e Julio Cesar Sampaio do Prado Leite: *Towards achieving trust through transparency and ethics (pre-print).* CoRR, abs/2107.02959, 2021. https://arxiv.org/abs/2107.02959, acesso em 2021-09-28. 35

[53] Cerqueira, José Antonio Siqueira de: *Exploring ethical requirements elicitation for applications in the context of AI.* Dissertação (Mestrado em Informática) — Universidade de Brasília, página 175, junho 2021. https://repositorio.unb.br/handle/10482/41966, acesso em 2021-09-28. 35

[54] International Conference on Very Large Data Bases e Christoph Koch: *33rd International Conference on Very Large Data Bases: University of Vienna, Austria, September 23-27 2007 : conference proceedings.* Association for Computing Machinery, New York, 2007, ISBN 9781595936493. http://portal.acm.org/toc.cfm?id=1325851, acesso em 2021-08-15, OCLC: 650168232. 36

[55] Brodley, Carla E e Paul E Utgoff: *Multivariate versus univariate decision trees.* Citeseer, 1992. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.48.5989&rep=rep1&type=pdf, acesso em 2021-09-06. 36

[56] Sun, Chen, Ye Tian, Liang Gao, Yishuai Niu, Tianlong Zhang, Hua Li, Yuqing Zhang, Zengqi Yue, Nicole Delepine-Gilon e Jin Yu: *Machine Learning Allows Calibration Models to Predict Trace Element Concentration in Soils with Generalized LIBS Spectra.* Scientific Reports, 9(1):11363, agosto 2019, ISSN 2045-2322. https://www.nature.com/articles/s41598-019-47751-y, acesso em 2021-09-20. 37, 38, 39

[57] Muller, Paul, Shayegan Omidshafiei, Mark Rowland, Karl Tuyls, Julien Pérolat, Siqi Liu, Daniel Hennes, Luke Marris, Marc Lanctot, Edward Hughes, Zhe Wang, Guy Lever, Nicolas Heess, Thore Graepel e Rémi Munos: *A generalized training approach for multiagent learning.* CoRR, abs/1909.12823, 2019. http://arxiv.org/abs/1909.12823, acesso em 2021-09-20. 37, 38

[58] Fung, B.C.M., K. Wang e P.S. Yu: *Top-down specialization for information and privacy preservation.* IEEE Xplore, páginas 205–216, 2005. https://ieeexplore.ieee.org/document/1410123, acesso em 2021-09-07. 37

[59] Djeffal, Christian: *Sustainable AI Development (SAID): On the Road to More Access to Justice*. SSRN Electronic Journal, 2018, ISSN 1556-5068. https://www.ssrn.com/abstract=3298980, acesso em 2021-09-22. 39, 40, 46, 51

[60] Sachs, Jeffrey D., Guido Schmidt-Traub, Mariana Mazzucato, Dirk Messner, Nebojsa Nakicenovic e Johan Rockström: *Six Transformations to achieve the Sustainable Development Goals*. Nature Sustainability, 2(9):805–814, setembro 2019, ISSN 2398-9629. http://www.nature.com/articles/s41893-019-0352-9, acesso em 2021-09-22. 39, 40

[61] Vinuesa, Ricardo, Hossein Azizpour, Iolanda Leite, Madeline Balaam, Virginia Dignum, Sami Domisch, Anna Felländer, Simone Daniela Langhans, Max Tegmark e Francesco Fuso Nerini: *The role of artificial intelligence in achieving the Sustainable Development Goals*. Nature Communications, 11(1):233, 2020, ISSN 2041-1723. http://www.nature.com/articles/s41467-019-14108-y, acesso em 2021-09-22. 40

[62] Jacobson, Ivar, Pan Wei Ng, Paul E. McMahon, Ian Spence e Svante Lidman: *The essence of software engineering: The semat kernel*. Commun. ACM, 55(12):42–49, dezembro 2012, ISSN 0001-0782. https://doi.org/10.1145/2380656.2380670. 40

[63] Jobin, Anna, Marcello Ienca e Effy Vayena: *The global landscape of AI ethics guidelines*. Nature Machine Intelligence, 1(9):389–399, setembro 2019, ISSN 2522-5839. http://www.nature.com/articles/s42256-019-0088-2, acesso em 2021-09-07. 40

[64] Peterson, Martin: *The Ethics of Technology: A Geometric Analysis of Five Moral Principles*. Oxford University Press, junho 2017, ISBN 9780190652289. https://books.google.com.br/books?hl=pt-BR&lr=&id=vTUlDwAAQBAJ&oi=fnd&pg=PT5&dq=the+Ethics+of+Technology:+A+Geometric+Analysis+of+Five+Moral+Principles&ots=fbpqSgsBOR&sig=dq9jdba-4B6nf93-1baa6cHRra8#v=onepage&q=the%20Ethics%20of%20Technology%3A%20A%20Geometric%20Analysis%20of%20Five%20Moral%20Principles&f=false, acesso em 2021-09-28, Google-Books-ID: vTUlDwAAQBAJ. 44, 45, 46

[65] Verma, A. K.: *Sustainable Development and Environmental Ethics*. International Journal on Environmental Sciences, junho 2019, ISSN 09764534. https://papers.ssrn.com/abstract=3689046, acesso em 2021-10-03. 44

[66] Agbese, Mamia: *Implementing artificial intelligence ethics in trustworthy systems development : extending eccola to cover information governance principles*. Information Systems Science, maio 2021, ISSN URN:NBN:FI:JYU-202105283279. https://jyx.jyu.fi/handle/123456789/76034, acesso em 2021-10-04. 45, 50, 51, 57

[67] Lima, Ana Paula Moraes Canto de, Marcelo Crespo e Patricia Peck Pinheiro: *LGPD Aplicada*, volume 1. GEN - Grupo Editorial Nacional, 2020, ISBN 978-85-97-02692-4. https://integrada.minhabiblioteca.com.br/#/books/9788597026931/. 48, 53, 59, 65

[68] Antikainen, Jani, Mamia Agbese, Hanna Kaisa Alanen, Erika Halme, Hannakaisa Isomaki, Marianna Jantunen, Kai Kristian Kemell, Rebekah Rousi, Heidi Vainio-Pekka e Ville Vakkuri: *Governnce of ethical and trustworthy ai systems: Research gaps in the eccola method*. 2021 IEEE 29th International Requirements Engineering Conference Workshops (REW), página 6, 2021, ISSN 9781665418980. 49, 50, 65

[69] Antikainen, Jani, Mamia Agbese, Hanna Kaisa Alanen, Erika Halme, Isomaki Hannakaisa, Marianna Jantunen, Kai Kristian Kemell, Rebekah Rousi, Heidi Vainio-Pekka e Ville Vakkuri: *A Deployment Model to Extend Ethically Aligned AI Implementation Method ECCOLA*. 2021 IEEE 29th International Requirements Engineering Conference Workshops (REW), 1:6, 2021. `https://ieeexplore.ieee.org/Xplore/home.jsp`, acesso em 2021-09-28. 50