



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
SISTEMAS DE INFORMAÇÃO

ITÉRBIO ARANTES ANDRADE

**GOVERNANÇA DE ALGORITMOS: IMPACTOS SOCIAIS, RISCOS,
NECESSIDADE E PROCESSOS**

Recife

2023

ITÉRBIO ARANTES ANDRADE

**GOVERNANÇA DE ALGORITMOS: IMPACTOS SOCIAIS, RISCOS,
NECESSIDADE E PROCESSOS**

Trabalho apresentado ao Curso de Graduação em Sistemas de Informação do Centro de Informática da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Bacharel em Sistemas de Informação.

Orientador: Prof. Dr. Hermano Perrelli de Moura

Recife

2023

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Andrade, Itérbio Arantes.

Governança de algoritmos: impactos sociais, riscos, necessidade e processos /
Itérbio Arantes Andrade. - Recife, 2023.
52 : il., tab.

Orientador(a): Hermano Perrelli de Moura

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de
Pernambuco, Centro de Informática, Sistemas de Informação - Bacharelado,
2023.

Inclui referências, apêndices.

1. governança de algoritmos. 2. riscos. 3. inteligência artificial. 4. processos
de governança. 5. sociedade. I. Moura, Hermano Perrelli de. (Orientação). II.
Título.

600 CDD (22.ed.)

ITÉRBIO ARANTES ANDRADE

**GOVERNANÇA DE ALGORITMOS: IMPACTOS SOCIAIS, RISCOS,
NECESSIDADE E PROCESSOS**

Trabalho apresentado ao Curso de Graduação em Sistemas de Informação do Centro de Informática da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Bacharel em Sistemas de Informação.

Data da Defesa do TCC:

Recife, 17 de abril de 2023

BANCA EXAMINADORA

Prof. Hermano Perrelli de Moura (Orientador)
UNIVERSIDADE FEDERAL DE PERNAMBUCO

Prof. Ruy José Guerra Barreto de Queiroz (2º membro da banca)
UNIVERSIDADE FEDERAL DE PERNAMBUCO

AGRADECIMENTOS

Agradeço primordialmente a minha mãe Norma Maria por ter me educado e ensinado com carinho e disciplina. Agradeço a todos que me ajudam ou ajudaram na minha caminhada.

Aos professores Hermano Perrelli e Ruy José Guerra por serem verdadeiros guias. A Valdi Junior pela ajuda e auxílio.

A minha companheira Jéssica Gondim por ter me apoiado e me acompanhado neste processo.

RESUMO

Com o início da era digital, ao passar dos anos, os algoritmos estão cada vez mais presentes na humanidade, como reflexo, os mesmos estão sendo utilizados para realizarem diversas tarefas que seriam ineficazes ou impossíveis para um ser humano. Com o aumento do uso dos algoritmos, alguns riscos são iminentes no que diz respeito a sua aplicação na sociedade, como: preconceito, discriminação social, manipulação, violação de privacidade, entre outros vieses que podem ser “mascarados” através da tecnologia. Para lidar com esses riscos, um processo de governança precisa ser considerado, para que os benefícios da inteligência artificial (IA) sejam maximizados e as ameaças mitigadas. Este trabalho faz uma revisão sistemática da literatura sobre estudos na área de governança de algoritmos, com o objetivo de elaborar uma análise da definição, importância e processos da governança de algoritmos.

Palavras-chave: governança de algoritmos, riscos, inteligência artificial, processos de governança, sociedade.

ABSTRACT

With the beginning of the digital age, over the years, algorithms are increasingly present in humanity, as a result, they are being used to perform various tasks that would be ineffective or impossible for a human being. With the increased use of algorithms, some risks are imminent with regard to their application in society, such as: prejudice, social discrimination, manipulation, violation of privacy, among other biases that can be “masked” through technology. To address these risks, a governance process needs to be considered, so that the benefits of artificial intelligence (AI) are maximized and threats mitigated. This work makes a systematic review of the literature on studies in the area of algorithm governance, with the objective of elaborating an analysis on the definition, importance and processes of algorithm governance.

Keywords: algorithm governance, risks, artificial intelligence, governance processes, society.

LISTA DE ABREVIATURAS E SIGLAS

| | |
|------|--|
| UFPE | Universidade Federal de Pernambuco |
| UFMG | Universidade Federal de Minas Gerais |
| CE | Critério de Exclusão |
| RSL | Revisão Sistemática na Literatura |
| MSL | Mapeamento Sistemático da Literatura |
| QP | Questões de Pesquisa |
| IA | Inteligência Artificial |
| APIs | Interface de Programação de Aplicação |
| AIA | Avaliação de Impacto Algorítmico |
| GDPR | Regulamento Geral de Proteção de Dados |
| ADS | Sistema de Decisão Algorítmica |

LISTA DE FIGURAS

| | |
|-----------------------------------|----|
| Figura 1: Triagem Etapa 1 | 19 |
| Figura 2: Triagem Etapa 2 | 21 |
| Figura 3: Resumo Triagem | 22 |
| Figura 4: Ano de Publicação | 26 |
| Figura 5: Artigos por país | 27 |

LISTA DE TABELAS

| | |
|--|----|
| Tabela 1: Resultado da Triagem 1 | 19 |
| Tabela 2: Resultado da Triagem 2 | 21 |
| Tabela 3: Resultado final da Triagem | 25 |
| Tabela 4: Principais riscos dos algoritmos | 33 |

SUMÁRIO

| | | |
|-------|--|----|
| 1 | INTRODUÇÃO | 11 |
| 1.1 | MOTIVAÇÃO..... | 12 |
| 1.2 | OBJETIVO..... | 12 |
| 1.3 | METODOLOGIA..... | 13 |
| 1.4 | ORGANIZAÇÃO DO DOCUMENTO | 13 |
| 2 | REFERENCIAL TEÓRICO | 14 |
| 2.1 | ALGORITMOS | 14 |
| 2.2 | GOVERNANÇA DE ALGORITMOS | 15 |
| 3 | REVISÃO SISTEMÁTICA DA LITERATURA | 17 |
| 3.1 | DEFINIÇÃO DA STRING DE BUSCA | 17 |
| 3.2 | SELEÇÃO DOS ARTIGOS..... | 18 |
| 3.2.1 | ETAPA 1 | 18 |
| 3.2.2 | ETAPA 2 | 19 |
| 3.2.3 | RESULTADO..... | 21 |
| 4 | RESULTADOS DA PESQUISA..... | 28 |
| 4.1 | IMPACTOS SOCIAIS DOS ALGORITMOS | 28 |
| 4.1.1 | SÍNTESE: IMPACTOS SOCIAIS DOS ALGORIMOS | 30 |
| 4.2 | RISCOS DOS ALGORITMOS | 30 |
| 4.2.1 | SÍNTESE: RISCOS DOS ALGORITMOS | 33 |
| 4.3 | IMPORTÂNCIA DA GOVERNANÇA DE ALGORITMOS | 34 |
| 4.3.1 | SÍNTESE: IMPORTÂNCIA DA GOVERNANÇA DE ALGORITMOS | 36 |
| 4.4 | PROCESSOS DA GOVERNANÇA DE ALGORITMOS | 36 |
| 5 | CONCLUSÃO E TRABALHOS FUTUROS | 43 |
| | REFERÊNCIAS | 45 |
| | APÊNDICE A — PROTOCOLO UTILIZADO NA REVISÃO SISTEMÁTICA DA LITERATURA..... | 49 |

1 INTRODUÇÃO

Com o início da era digital, ao passar dos anos, a tecnologia está cada vez mais presente na humanidade, o que implica no crescimento do uso da mesma para realizar muitas tarefas. Os algoritmos são basicamente instrumentos criados para auxiliar a população de alguma forma, seja em uma tarefa ou decisão, dado um input (entrada) os mesmos irão produzir um output (saída). A disponibilidade de poder de processamento cada vez maior e conjuntos de dados possibilitam que os algoritmos executem tarefas com magnitudes e complexidade insuportáveis para os padrões de uma pessoa, o que os torna imprescindíveis para a raça humana.

Embora os algoritmos forneçam resultados valiosos, os mesmos podem tirar o processo decisório do ser humano, o que pode ser arriscado. Alguns riscos provaram-se iminentes no que diz respeito a sua utilização, como: preconceito, discriminação social, manipulação, violação de privacidade, entre outros vieses que podem ser “mascarados” através da tecnologia. Assim, para promover a integração em diversos processos sociais e econômicos, talvez fosse necessário a criação de algum tipo de governança, a qual instituisse leis para regular o desenvolvimento e implantação dos algoritmos (DANILO & V. ALMEIDA, 2016).

A complexidade dos algoritmos é aumentada pelo uso crescente de técnicas de aprendizado de máquina, com essas técnicas, um algoritmo pode reorganizar e transformar o seu funcionamento interno com base nos dados que analisa. Como Pedro Domingos descreveu uma vez: “Algoritmos de aprendizado são algoritmos que fazem outros algoritmos” (PEDRO DOMINGOS, 2015). Esta afirmação confirma o motivo pelo qual os cientistas de dados, muitas vezes, possuem dificuldade para analisar e descrever as etapas de um algoritmo. É notório, cada vez mais, a dificuldade dos humanos de entenderem, explicarem ou preverem o funcionamento interno dos algoritmos, suas bases e eventuais problemas. Isso é uma situação preocupante, uma vez que os algoritmos “são” confiáveis para tomar decisões importantes, se não fundamentais, que afetam nossas vidas.

Gradativamente trabalhos acadêmicos e públicos trazem à tona a necessidade por mais responsabilidade e transparência dos algoritmos (EPIC.ORG, 2015). É um fato que a tecnologia enriquece nossas vidas e possibilita novas vertentes para o futuro. Contudo, esses sistemas também podem ameaçar os valores fundamentais da sociedade. Leis devem ser criadas e regulamentadas para garantir uma proteção aos usuários contra abuso, exploração, discriminação e vigilância invasiva (EPIC.ORG, 2015). É crucial que exista uma governança para os algoritmos.

1.1 MOTIVAÇÃO

Responsabilidade é um dever social que todo cidadão ou organização deve exercer. Garantir que os direitos humanos não serão violados é um dever do Poder Legislativo e Executivo. Com base nessas afirmações, para implantarmos ou criarmos algo que fará parte da sociedade, a legalidade e os cuidados são fundamentais.

Os algoritmos existem para auxiliar ou realizar uma determinada tarefa, a qual seria ineficaz se realizada parcialmente ou totalmente por um humano. A ideia de implantar sistemas e automatizar processos vem ganhando força conforme a tecnologia avança, os ganhos e benefícios que os algoritmos trazem para a sociedade são imensuráveis. Máquinas treinam outras máquinas, sistemas preveem eventos, robôs conseguem pensar e agir, a evolução tecnológica é iminente.

Apesar dos benefícios evidentes da tecnologia, existe a possibilidade da mesma apresentar algum tipo de risco para a sociedade dependendo do seu desenvolvimento, vale lembrar que um algoritmo em sua fase primária é construído por um humano, o qual pode ser falho. Assim como existem leis para manter um bom convívio social, para garantir que um algoritmo não viole os direitos humanos, os mesmos devem seguir algumas normas.

A governança de algoritmos tem como objetivo regulamentar e fiscalizar os sistemas algorítmicos, garantindo que os mesmos não violem os direitos humanos. Os vieses na implementação de uma tecnologia são os principais problemas. Imaginemos a seguinte situação: Um aeroporto precisa implantar um sistema baseado em I.A para identificar possível trâfico; para implantar um sistema como esse, a organização responsável precisa garantir que os dados fornecidos para treinar a I.A sejam íntegros, sem nenhuma manipulação ou viés. Para assegurar a imparcialidade e ética de um algoritmo, uma governança precisa ser estabelecida.

1.2 OBJETIVO

O objetivo deste trabalho compreende elaborar uma análise da definição, importância e processos da governança de algoritmos, com base em estudos e artigos publicados sobre esses pontos. Para atingir tal objetivo, foi realizado uma revisão sistemática na literatura no que diz respeito à governança de algoritmos, para que fosse possível elaborar uma avaliação dos resultados obtidos a partir dos estudos dessa revisão.

Com o intuito de extrair os principais conceitos dos estudos, quatro questões de pesquisa foram consideradas para serem respondidas:

QP1: Quais os impactos sociais dos algoritmos?

QP2: Quais os riscos que os algoritmos podem apresentar?

QP3: Qual a necessidade de uma governança para os algoritmos?

QP4: Quais processos fazem parte da governança de algoritmos?

Ao responder essas perguntas, este estudo servirá de apoio a quem se interessar em entender e avaliar o conceito, causas e processos da governança dos algoritmos. Além disso, as respostas deste trabalho servem de guia para justificar o motivo pelo qual os algoritmos devem passar por uma “vistoria” antes de serem implantados na sociedade.

1.3 METODOLOGIA

Com o objetivo de reunir e avaliar um conjunto de dados de diferentes estudos, e consequentemente responder às questões definidas na seção 1.2 deste trabalho, foi realizado uma revisão sistemática, a qual pode ser definida como um método que propõe responder uma ou mais perguntas de forma objetiva e imparcial (MAICON FALAVIGNA, 2018). O método foi conduzido conforme o processo descrito em (KITCHENHAM, 2004). De acordo com a autora, uma revisão sistemática pode ser dividida em cinco etapas: (i) definição das questões de pesquisa; (ii) realização de buscas por artigos primários relevantes nas fontes de pesquisa escolhidas; (iii) triagem dos artigos (qualificação); (iv) extração de dados baseado nas perguntas de pesquisa; e (v) síntese dos dados.

1.4 ORGANIZAÇÃO DO DOCUMENTO

Este trabalho é resumido em 5 capítulos. O primeiro capítulo conta com a parte introdutória a respeito do tema governança de algoritmos, assim como o objetivo do estudo, a motivação e a metodologia aplicada na realização da pesquisa. O segundo capítulo tem o intuito de discorrer de forma teórica e isolada os conceitos de algoritmos e governança de algoritmos. O terceiro capítulo apresenta as fases da revisão sistemática que foram realizadas. O quarto capítulo evidencia os resultados obtidos a partir da pesquisa realizada. Por fim, o quinto capítulo demonstra uma breve conclusão do trabalho como um todo.

2 REFERENCIAL TEÓRICO

Este capítulo introduz a definição do que são os algoritmos, como eles surgiram e qual sua importância para a sociedade. Além disso, uma breve definição do que se trata a governança de algoritmos e alguns motivos pelo qual este processo deve ser considerado.

2.1 ALGORITMOS

O termo algoritmo deriva do nome de um matemático persa do século IX, conhecido como Abdullah Muhammad ibn Musa Al-Khawarizmi. Seu nome em latim era: *Algoritmi*, o que significava “o sistema de numeração decimal” e foi usado com esse significado por séculos. A partir da década de 1950, com o surgimento dos primeiros computadores, o significado de algoritmos mudou para o que conhecemos hoje (MAREK KOWALKIEWICZ, 2019).

O significado de algoritmos que conhecemos atualmente surgiu com o início da engenharia mecânica e seus processos. No princípio, os algoritmos auxiliaram a matemática, dando base a álgebra da lógica, usando variáveis nos cálculos. Os primeiros indícios dos algoritmos incluem a aproximação de Pi de Arquimedes, a função do máximo divisor comum em números de Euclides e o cálculo de números primos de Erastóstenes (SOUVIK DAS, 2016).

Em 1800 ocorreram várias conquistas que auxiliaram na evolução dos algoritmos, a primeira, marcada como: a Álgebra Booleana, foi estabelecida pelo matemático inglês George Boole. Em 1847, Boole unificou a lógica com cálculos, formando a álgebra binária, a qual deu origem à base da lógica computacional. Por fim, mediante as descobertas citadas, nascem os algoritmos computacionais (SOUVIK DAS, 2016).

Um algoritmo pode ser descrito como um conjunto de instruções bem definidas com o objetivo de solucionar um determinado problema. O conceito de algoritmo é frequentemente ilustrado na literatura pelo exemplo de uma receita culinária, pois, assim como os algoritmos, uma receita consiste em passos específicos e sucessivos para realizar uma determinada tarefa. Na ciência da computação um algoritmo pode ser resumido como um conjunto de regras finitas e rigorosas para realizar um determinado procedimento computacional.

Os algoritmos podem ser classificados em diferentes tipos, de acordo com suas características (propriedades e domínios) - algoritmos combinatórios lidam com contagem e enumeração, algoritmos numéricos produzem respostas numéricas para problemas equacionais, enquanto algoritmos probabilísticos produzem resultados dentro de determinados limites. (PAUL DOURISH, 2016).

Conforme definido por Dourish: “um algoritmo é uma descrição abstrata e formalizada de um procedimento computacional entrelaçado com outras formas computacionais, como dados, estrutura de dados, programas e processos” (PAUL DOURISH, 2016). Um algoritmo também pode ser definido como um sistema de várias partes, o que pode incluir algoritmos de aprendizado de máquina, modelos, funções matemáticas, dados de treinamento, aplicação, hardware e redes. Pode ser descrito como um conjunto sociotécnico: uma combinação de técnicas e componentes sociais, onde os atores sociais e os artefatos tecnológicos se cruzam como parte de um sistema unificado (V. ALMEIDA & F. FILGUEIRAS, 2022).

Grande parte dos sistemas de decisão algorítmica que são geridos por um próprio algoritmo, geralmente, dependem da análise de grandes quantidades de dados para advir informações e inferir correlações para a tomada de decisões. YouTube, Spotify, Amazon e Uber são alguns exemplos de sistemas algorítmicos, com serviços conduzidos por algoritmos.

Contudo, independente das características funcionais, todos os algoritmos possuem um objetivo em comum: Auxiliar na execução de uma determinada tarefa que seria ineficaz ou impossível se realizada por um ser humano.

2.2 GOVERNANÇA DE ALGORITMOS

Governança é a forma como as regras, normas e ações são estruturadas, sustentadas, reguladas e responsabilizadas (BEVIR & MARK, 2012). Por sua vez, a governança de algoritmos pode ser resumidamente definida como a forma de regular e responsabilizar os resultados dos algoritmos. A governança de algoritmos pode ser aplicada em diferentes contextos, desde o ponto de vista estritamente legal e regulatório até um ponto de vista puramente técnico que visa analisar a arquitetura e implementação da tecnologia. Em grande parte, o foco da regência é direcionado para a transparência e garantias técnicas dos algoritmos. Governança de algoritmos é um componente central para que as normas regulatórias do mundo digital sejam construídas. Conforme proposto por Floridi, a governança de algoritmos pode ser vista como a “prática de estabelecer e implementar políticas, procedimentos e padrões para o desenvolvimento, uso e gerenciamento adequados da Infoesfera (L. FLORIDI, 2018).

Existem dois significados que podem ser confundidos, mas estão associados ao processo de governar algoritmos na sociedade. O primeiro, conhecido puramente como governança de algoritmos, refere-se a práticas e políticas para controlar e regular os algoritmos e seus impactos nos indivíduos e na sociedade. A governança geralmente é baseada em ferramentas que não

funcionam propriamente no algoritmo, e sim em elementos de seu ambiente, como conjuntos de dados e elementos de design. O intuito desse processo é na prestação de contas, transparência, garantias técnicas, natureza do algoritmo, seu contexto e riscos. (V. ALMEIDA & F. FILGUEIRAS, 2022). O segundo processo é a governança por algoritmos, a qual se refere a sistemas e mecanismos que “regem” seus próprios feitos, implementados com objetivos bem especificados em direção a feitos econômicos, políticos e sociais. Um caso para exemplificar a governança por algoritmos é o Uber, que faz uso de um gerenciamento algorítmico, permitindo que seja necessário um número mínimo de gerentes humanos para supervisionar milhares de motoristas de forma otimizada e em grande escala.

Para viabilizar e tornar necessário uma governança, alguns fatores relevantes são considerados, como: a natureza do algoritmo, seu contexto, ou análise de risco. Geralmente, quando uma opção de governança é feita, ela visa reduzir os problemas causados pelo algoritmo. Deve tentar preservar a sua eficácia e reduzir os resultados indesejáveis.

Vale lembrar que, algumas técnicas utilizadas na governança não agem diretamente no algoritmo, mas nos dados que são utilizados para seu funcionamento. Isso acontece em várias ferramentas já presentes na legislação de proteção de dados. Em alguns países, leis de transparência e equidade são aplicadas diretamente aos algoritmos e às plataformas que suportam seu funcionamento (DANILO & V. ALMEIDA, 2016).

3 REVISÃO SISTEMÁTICA DA LITERATURA

Este capítulo apresenta uma revisão sistemática da literatura usando a abordagem definida por Kitchenham (KITCHENHAM, 2004). O processo realizado é descrito em cada seção do capítulo.

3.1 DEFINIÇÃO DA STRING DE BUSCA

Como fontes de artigos publicados, a busca realizada incluiu 3 bases de dados: IEEE, ACM e Science Direct. Todas essas bibliotecas digitais possuem páginas Web onde foi possível realizar as buscas pelas palavras chaves de interesse.

As pesquisas nessas fontes foram feitas inicialmente considerando algumas palavras chaves e formaram as seguintes Strings de busca:

- Versão 1: (“governing algorithms” OR “governance algorithm” OR “algorithmic governance” OR “algorithm regulation”) AND (“society” OR “organization”) AND (“transparency” OR “evidence” OR “accountability” OR “regulation”)
- Versão 2: (“algorithm governance” OR “algorithmic governance”) AND (“social” OR “organization”) AND (“model” OR “method” OR “technique” OR “regulation” OR “privacy”)

Após a análise dos resultados algumas palavras chaves foram desconsideradas, pois os artigos retornados não contemplavam o foco da revisão sistemática. Testes de busca foram realizados com o intuito de refinar a String para a pesquisa, o que resultou na seguinte consulta:

- (“algorithm governance” OR “algorithmic governance”) AND (“society” OR “social” OR “organization”) AND (“model” OR “process” OR “regulation” OR “privacy”)

Mostrou-se necessário o acréscimo do termo *society* para garantir que seriam retornados resultados que envolvessem aspectos da sociedade. Além disso, o termo *process* foi utilizado

como substituto/sinônimo dos termos *method* e *technique* com o intuito de gerar uma String mais enxuta e precisa.

3.2 SELEÇÃO DOS ARTIGOS

A primeira busca retornou um total de 232 artigos, levando em consideração as três fontes. Para selecionar os artigos mais relevantes para o estudo foram realizadas algumas etapas de filtragem, detalhadas nas seções a seguir.

3.2.1 ETAPA 1

A primeira etapa consistiu em definir alguns critérios de exclusão, considerando que a primeira busca não foi realizada a partir de algum filtro em específico, foi necessário realizar uma triagem por idioma, data, acesso e tipo de artigo. Seguem os critérios definidos:

- CE01 - Artigos que não estejam no período de 2016 a 2022
- CE02 - Artigos secundários (outra RSL ou MS)
- CE03 - Artigos com conteúdo pago
- CE04 - Artigos duplicados ou similares
- CE05 - Artigos que não estejam relacionados com as perguntas de pesquisa
- CE06 - Artigos que não estejam em inglês

A Figura 1 mostra a relação entre os artigos selecionados inicialmente e os artigos selecionados após a aplicação dos critérios de exclusão.

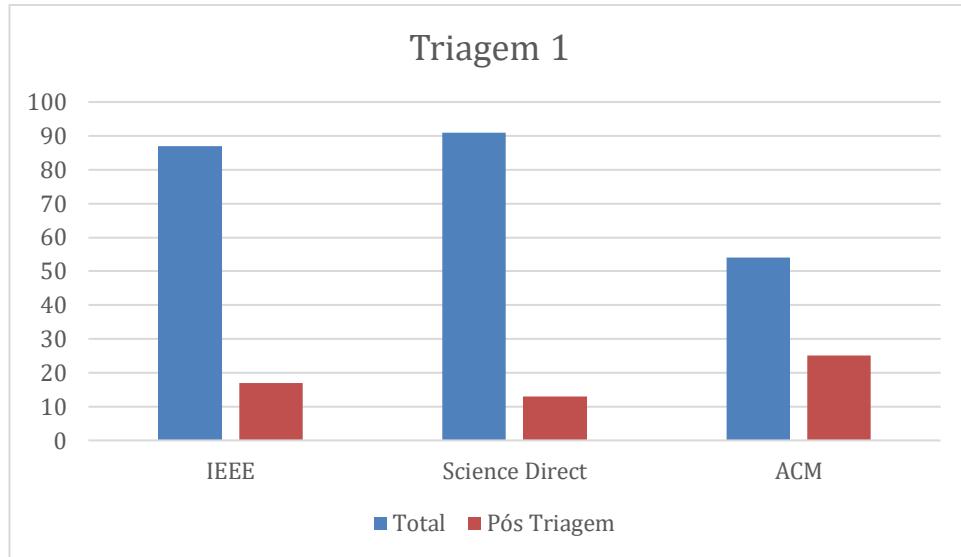


Figura 1: Triagem Etapa 1

A Tabela 1 mostra os resultados dos artigos selecionados após a aplicação dos critérios de exclusão.

| | |
|----------------|----|
| IEEE | 17 |
| Science Direct | 13 |
| ACM | 25 |

Tabela 1: Resultado da Triagem 1

Para análise do critério C05 os títulos e resumos dos artigos foram considerados, de forma que fosse possível identificar se o artigo analisado estava coerente com o tema proposto. Nesta etapa foram marcados como C05 todos os artigos que não continham dados relacionados a governança associada a algoritmos e vice e versa. Não foram identificados artigos repetidos ou similares entre as fontes de pesquisa, logo, o critério C04 não foi aplicado de forma direta.

3.2.2 ETAPA 2

A etapa 1 não levava em consideração a qualidade dos artigos selecionados, por isso foi necessário realizar uma leitura mais aprofundada dos artigos, considerando a introdução e conclusão dos textos, além de analisar partes de cada artigo que tinham relação com as

perguntas de pesquisa definidas, de forma que fosse verificado se as mesmas seriam respondidas.

Para realizar esta etapa foram utilizados alguns critérios para qualificação dos estudos:

- Contexto claro
- Metodologia bem definida
- Aplicação Prática
- Discussão relevantes e consistentes

O objetivo final era que cada estudo possuísse uma nota de qualidade, a qual seria formada pela média das notas de cada critério definido acima. Para valorar os critérios foram utilizadas a seguinte escala:

- **Nota 0:** Não atende
- **Nota 0,5:** Atende parcialmente
- **Nota 1:** Atende totalmente

Para seleção de qualificação dos artigos, foram considerados os estudos com notas iguais ou superiores a 0,5.

A Figura 2 mostra a relação entre os artigos selecionados na Etapa 1 e os artigos selecionados após a aplicação dos critérios de qualificação.

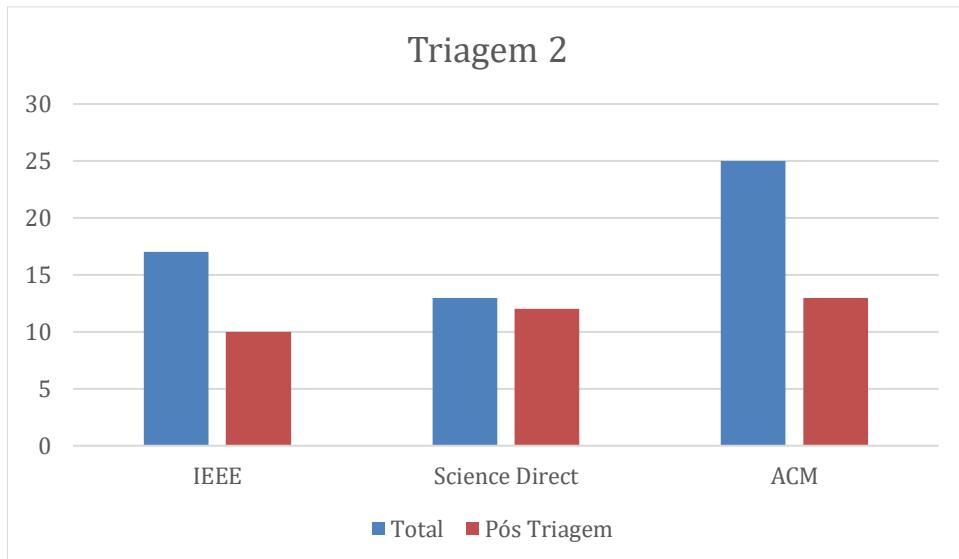


Figura 2: Triagem Etapa 2

A Tabela 2 mostra os resultados dos artigos selecionados após a aplicação dos critérios de qualificação.

| | |
|----------------|----|
| IEEE | 10 |
| Science Direct | 12 |
| ACM | 13 |

Tabela 2: Resultado da Triagem 2

3.2.3 RESULTADO

O objetivo deste processo foi selecionar os artigos mais relevantes para responder às questões propostas sobre o tema governança de algoritmos. Todos os artigos apresentados nesta seção foram utilizados para responder às perguntas de pesquisa, além de serem utilizados como referências para elaboração do conteúdo deste documento. Os resultados apresentados a seguir são referentes aos artigos selecionados.

A Figura 3 demonstra a quantidade de artigos por fonte após cada etapa que foi evidenciada neste capítulo.

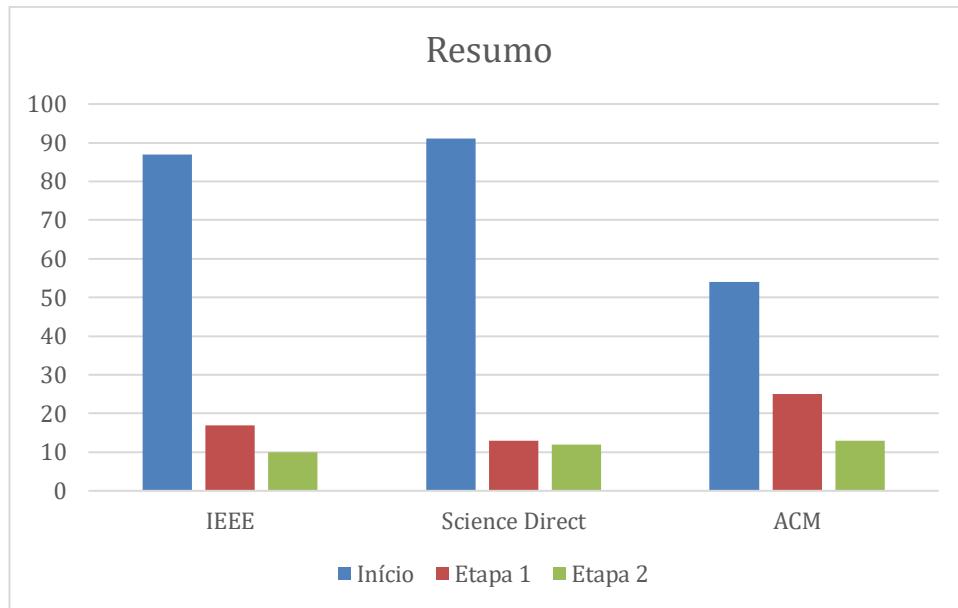


Figura 3: Resumo Triagem

A Tabela 3 apresenta o identificador, título, ano e fonte de cada um dos 35 artigos selecionados.

| ID | Título | Ano | Fonte |
|----|--|------|-------|
| A1 | A Layered Model for AI Governance | 2017 | IEEE |
| A2 | Algorithmic Governance in Smart Cities: The Conundrum and the Potential of Pervasive Computing Solutions | 2018 | IEEE |
| A3 | Algorithms and Institutions: How Social Sciences Can Contribute to Governance of Algorithms | 2022 | IEEE |
| A4 | Decision Provenance: Harnessing Data Flow for Accountable Systems | 2019 | IEEE |
| A5 | Framework for Trustworthy Software Development | 2019 | IEEE |

| | | | |
|-----|---|------|---------------|
| A6 | Governance and Communication of Algorithmic Decision Making: A Case Study on Public Sector | 2021 | IEEE |
| A7 | Procurement as AI Governance | 2021 | IEEE |
| A8 | Trusted Artificial Intelligence: Technique Requirements and Best Practices | 2021 | IEEE |
| A9 | U.S. Artificial Intelligence Governance in the Obama–Trump Years | 2021 | IEEE |
| A10 | What Is Algorithm Governance? | 2016 | IEEE |
| A11 | The flaws of policies requiring human oversight of government algorithms | 2022 | ScienceDirect |
| A12 | Data governance: Organizing data for trustworthy Artificial Intelligence | 2020 | ScienceDirect |
| A13 | Information privacy, impact assessment, and the place of ethics | 2020 | ScienceDirect |
| A14 | Deep learning and principal–agent problems of algorithmic governance: The new materialism perspective | 2020 | ScienceDirect |
| A15 | Governance of artificial intelligence: A risk and guideline-based integrative framework | 2022 | ScienceDirect |
| A16 | An artificial intelligence algorithmic approach to ethical decision-making in human resource management processes | 2022 | ScienceDirect |

| | | | |
|-----|--|------|---------------|
| A17 | Unpacking algorithms as technologies of power: Syrian refugees and data experts on algorithmic governance | 2021 | ScienceDirect |
| A18 | Hard choices in artificial intelligence | 2021 | ScienceDirect |
| A19 | Opportunity for renewal or disruptive force? How artificial intelligence alters democratic politics | 2020 | ScienceDirect |
| A20 | Legal and human rights issues of AI: Gaps, challenges and vulnerabilities | 2020 | ScienceDirect |
| A21 | Assessing behavioral data science privacy issues in government artificial intelligence deployment | 2022 | ScienceDirect |
| A22 | A shallow defence of a technocracy of artificial intelligence: Examining the political harms of algorithmic governance in the domain of government | 2020 | ScienceDirect |
| A23 | Algorithmic Impact Assessments and Accountability: The Co-Construction of Impacts | 2021 | ACM |
| A24 | Algorithms Off-Limits? If Digital Trade Law Restricts Access to Source Code of Software Then Accountability Will Suffer | 2022 | ACM |
| A25 | An Action-Oriented AI Policy Toolkit for Technology Audits by Community Advocates and Activists | 2021 | ACM |

| | | | |
|-----|--|------|-----|
| A26 | Fighting Back Algocracy: The Need for New Participatory Approaches to Technology Assessment | 2020 | ACM |
| A27 | Governing Algorithmic Systems with Impact Assessments: Six Observations | 2021 | ACM |
| A28 | Human Autonomy in Algorithmic Management | 2022 | ACM |
| A29 | Is a Decentralized Autonomous Organization a Panopticon? Algorithmic Governance as Creating and Mitigating Vulnerabilities in DAOs | 2021 | ACM |
| A30 | Is the Data Fair? An Assessment of the Data Quality of Algorithmic Policing Systems | 2020 | ACM |
| A31 | Mind the Gap: Towards an Understanding of Government Decision-Making Based on Artificial Intelligence | 2022 | ACM |
| A32 | Multi-Layered Explanations from Algorithmic Impact Assessments in the GDPR | 2020 | ACM |
| A33 | Power Dynamics and Value Conflicts in Designing and Maintaining Socio-Technical Algorithmic Processes | 2022 | ACM |
| A34 | Robots Enact Malignant Stereotypes | 2022 | ACM |
| A35 | The Algorithmic Imprint | 2022 | ACM |

Tabela 3: Resultado final da Triagem

Após a análise dos artigos selecionados foi realizado um mapeamento dos estudos por ano de publicação. Ao agrupar os estudos por período, foi possível confirmar a relevância do tema governança de algoritmos nos últimos anos. Uma vez que os algoritmos cada vez mais ganham espaço na sociedade, faz-se suscetível que a quantidade de publicações de trabalhos relacionados ao tema aumente gradativamente.

A Figura 4 mostra a quantidade de artigos publicados por ano:

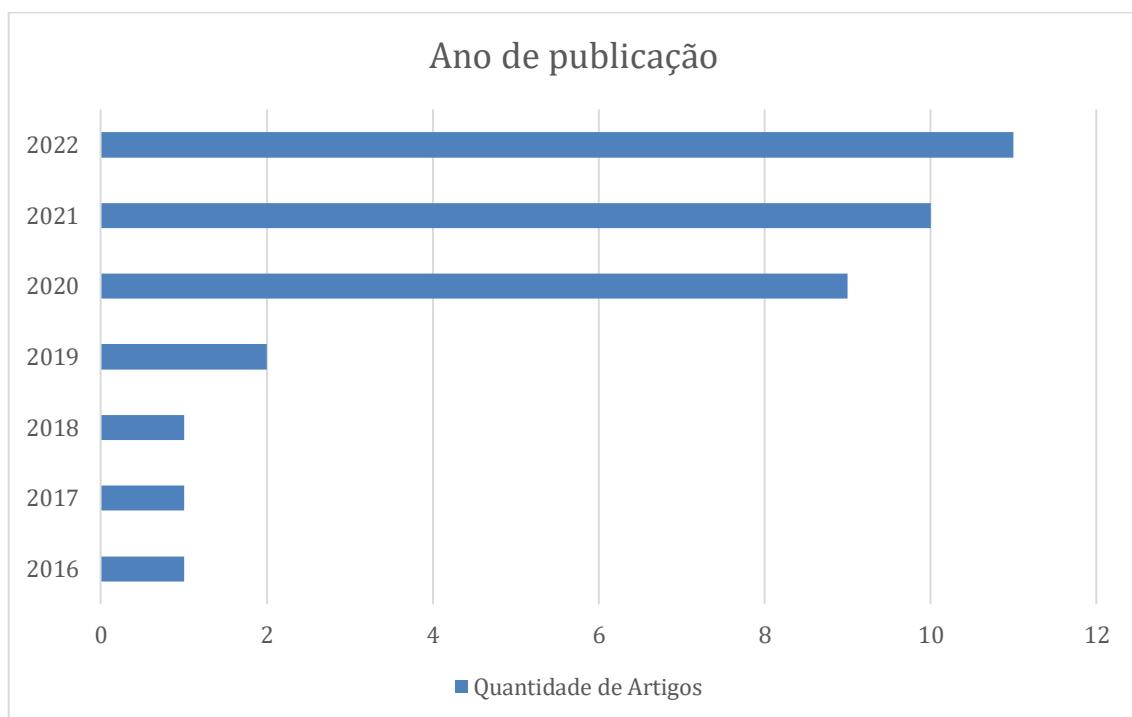


Figura 4: Ano de Publicação

Em vários países do mundo a IA já contribui de forma integral na realização de processos, tomadas de decisão, monitoramento e segurança, aspectos importantes que fazem parte da sociedade como um todo. Com a intenção de identificar em quais países os estudos foram divulgados, foi realizado um mapeamento das publicações dos estudos por país, no intuito de descobrir em quais locais o tema é recorrente e validar a distribuição geográfica sobre a temática.

A Figura 5 mostra a quantidade de artigos publicados por país:

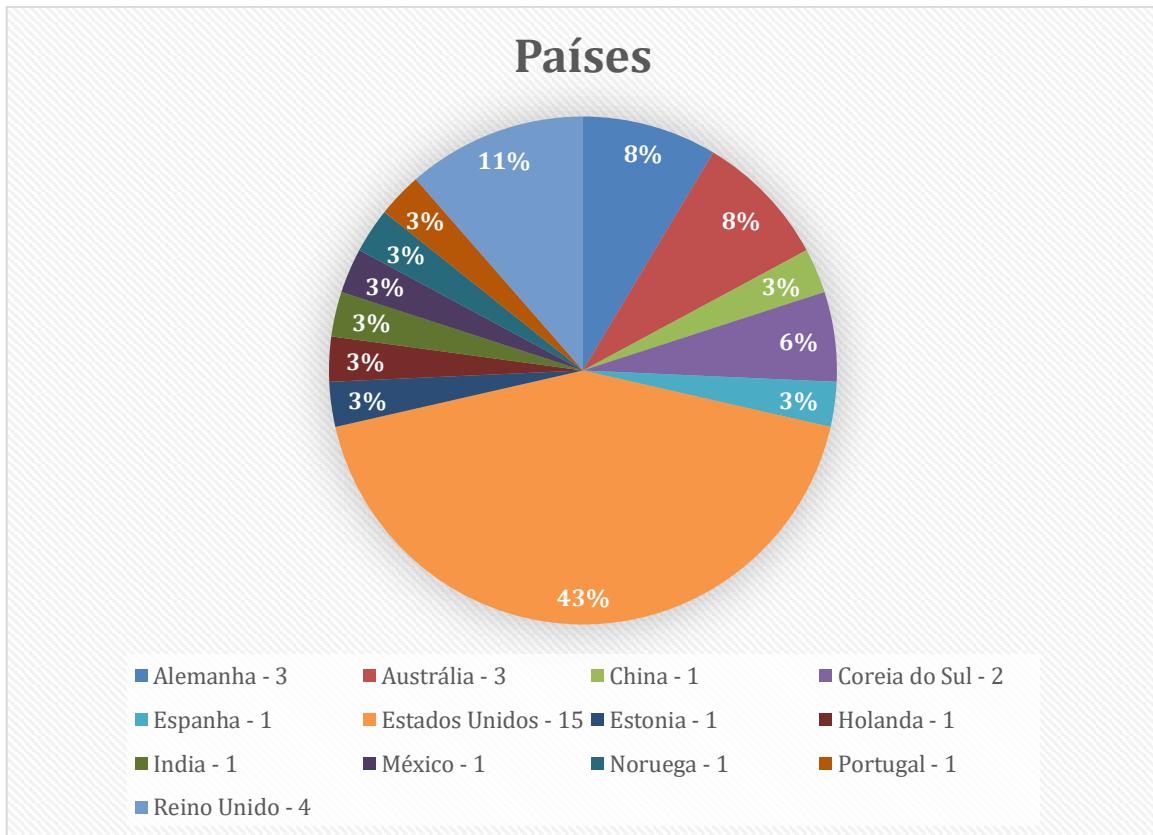


Figura 5: Artigos por país

Conforme apresentado na Figura 5, o país com a maior quantidade de artigos publicados foi os Estados Unidos, seguido do Reino Unido, Alemanha e Austrália.

Outra métrica que foi analisada tinha como objetivo identificar se um autor publicou mais de um estudo dentre os 35 selecionados. O autor Virgílio A.F. Almeida, professor da Universidade Federal de Minas Gerais (UFMG) de Ciência da computação, fez parte da publicação de 3 artigos. Vale ressaltar que os artigos publicados pelo professor tiveram nota de qualificação 1.0 (nota máxima) após a Etapa 2 descrita na Seção 3.2.2 deste trabalho, o que confirma a relevância do autor a respeito do tema governança de algoritmos.

4 RESULTADOS DA PESQUISA

Este capítulo é dividido em quatro seções e apresenta os resultados das perguntas de pesquisas definidas nos objetivos deste trabalho. A primeira seção visa apresentar os resultados obtidos que respondem à pergunta de pesquisa QP1, a segunda seção a pergunta de pesquisa QP2 e assim sucessivamente.

4.1 IMPACTOS SOCIAIS DOS ALGORITMOS

A primeira questão deste estudo visa entender os impactos que os algoritmos podem causar na sociedade.

Os estudos mostraram que os algoritmos gradativamente estão ganhando espaço na humanidade. Atualmente, sistemas e dispositivos eletrônicos estão sendo utilizados cada vez mais para tomar decisões, realizar avaliações e analisar causas que impactam nossas vidas. Muitos setores da sociedade estão adotando rapidamente tecnologias digitais e Big Data, resultando em sistemas autônomos e tomada de decisões algorítmicas em bilhões de vidas humanas. A tendência é que a tecnologia auxilie ou substitua (em alguns casos) algumas atividades realizadas por pessoas, devido ao fato do poder da computação e o conjunto de dados possibilitarem que os algoritmos executem tarefas com uma magnitude e complexidade impossíveis para os padrões humanos (DANILO & V. ALMEIDA, 2016).

Atualmente os algoritmos desempenham um papel fundamental na sociedade. Os algoritmos são o que muitas vezes decidem se alguém será contratado ou demitido, se será aprovado ou reprovado em uma prova, se terá acesso a um determinado benefício social, quais notícias ele verá na mídia e nas redes sociais, qual o melhor caminho para chegar em um local. Em muitos campos, as tecnologias digitais estão permitindo a construção ou reengenharia da sociedade, tendo os algoritmos como elemento central deste processo (F. FILGUEIRAS & R. F. MENDONÇA, 2022).

Os estudos em Ciências Sociais gradativamente identificam problemas, processos e resultados do poder dos algoritmos na sociedade. Conforme descrito por F. Filgueiras: “O poder é um dos conceitos centrais nas Ciências Sociais e é amplamente definido como a capacidade de uma entidade individual ou coletiva de influenciar o comportamento, as ações e as crenças dos outros, independentemente da vontade desses outros. Concebido como uma influência, essa capacidade é inherentemente relacional” (F. FILGUEIRAS & R. F. MENDONÇA, 2022). Devido ao poder que os algoritmos estão exercendo na sociedade, é comum que o

comportamento dos indivíduos que fazem parte da mesma seja afetado. Essa mudança, pouco a pouco, institucionaliza-se e passa a ser vista como natural. Os algoritmos, como as instituições, são dispositivos que auxiliam na resolução de algum problema ou tomada de decisão. Conforme compreendemos algoritmos como instituições, entendemos que eles possuem uma “habilidade” inerente, que é estruturar o comportamento humano, afetando diretamente o desígnio das nossas escolhas (F. FILGUEIRAS & R. F. MENDONÇA, 2022).

O rápido desenvolvimento da Inteligência Artificial (IA) nos últimos anos impactou diretamente aspectos da saúde, transporte, mídia, segurança e outros campos (T. ZHANG & Y. QIN, 2021). A evolução dos algoritmos permitiu que algumas tecnologias surgissem, como: sistemas de previsão de crimes, testes para vários tipos de doenças, carros autônomos, entre várias outras aplicações. Vários desses sistemas afetam diretamente a sociedade em virtude de estarem ligados a atividades cotidianas. A IA já é utilizada como ferramenta de apoio, a mesma pode auxiliar médicos a encontrar um determinado diagnóstico corretamente, ou mesmo proporcionar maior racionalidade e imparcialidade em decisões judiciais (DANILO & V. ALMEIDA, 2016).

Alguns artigos mostraram o quanto a IA pode beneficiar a sociedade no quesito evolutivo. Segundo F. Zambonelli: “as tecnologias da computação podem tornar as cidades e vários ambientes ‘inteligentes’, ou seja, capazes de atingir a consciência dos processos físicos e sociais e de afetá-los dinamicamente de maneira proposital”. Viver em um ambiente inteligente consequentemente nos torna mais inteligentes, ao aumentar nosso nível de consciência sobre as atividades urbanas em andamento. Além disso, ao apoiar e facilitar nossas atividades habituais, como: dirigir ou encontrar um restaurante mais próximo, a vida acaba tornando-se mais agradável e menos estressante, outrossim, cidades inteligentes podem tornar o ambiente mais sustentável. (F. ZAMBONELLI & F. SALIM, 2018).

Grande parte dos trabalhos analisados mostraram os benefícios da IA em órgãos governamentais, com o uso de ferramentas algorítmicas os mesmos encontraram maneiras de lidar com uma série de funções públicas de forma ágil. As agências estão cada vez mais confiando em algoritmos de aprendizado de máquina para auxiliar ou lidar com processos de tomada de decisão em vários cenários (L. M. B. DOR & C. COGLIANESE, 2021). O interesse em fornecer soluções baseadas em IA já é constatado na maioria dos governos em todo o mundo, devido a geração de valor público e melhoria nos serviços (qualidade). Além disso, as técnicas de IA podem melhorar a interoperabilidade dos dados e análise de Big Data. Com a exploração da Big Data, pode ser possível gerar cenários e modelos úteis para a tomada de decisões estratégicas. Como resultado comum, a IA pode executar tarefas rotineiras das

organizações, liberando os trabalhadores para realizar atividades que podem agregar um maior valor e precisam de uma intervenção humana (VALLE-CRUZ D & GARCÍA-CONTRERAS R, 2022).

4.1.1 SÍNTESE: IMPACTOS SOCIAIS DOS ALGORITMOS

Contudo, embora os algoritmos afetem a sociedade positivamente em uma ampla gama de contextos, o potencial dessas ferramentas causarem danos pode vir à tona quando o governo depende do aprendizado de máquina para tomar decisões. Não é à toa que alguns membros da organização não apoiam o uso da IA no governo. Vale lembrar que existem impactos altamente consequentes no bem-estar e na liberdade das pessoas, visto o poder do governo de tomar decisões.

Em contraste com a maioria das ideias apresentadas nesta seção, os algoritmos também podem trazer riscos para a sociedade, grupo de usuários, governo e organizações. Com base nessa afirmação, a segunda pergunta da pesquisa se provou pertinente, uma vez que tende a levantar estudos que indiquem quais os riscos que os algoritmos podem apresentar.

4.2 RISCOS DOS ALGORITMOS

A segunda questão deste trabalho evidencia os possíveis riscos que os algoritmos podem implicar ao serem empregados em uma organização ou ambiente.

Embora possamos reconhecer os benefícios que os algoritmos trazem para a sociedade, no que concerne ao potencial de tornar nossas vidas melhores e desempenhar um papel importante no apoio à eficiência e inovação, os mesmos podem trazer riscos e perigos. As ameaças dos sistemas baseados em IA com tomada de decisão automatizada têm sido amplamente difundidas nos noticiários e artigos acadêmicos. Discriminação social, nodesinformação e discurso de ódio, são alguns de vários incidentes nos últimos anos causados por sistemas algorítmicos. Conforme apontado por Noble: “Sistemas algorítmicos de tomada de decisão rotineiramente beneficiam pessoas e grupos que já são privilegiados, enquanto discriminam os demais” (S. U. NOBLE, 2018).

O fato dos algoritmos fazerem parte da nossa vida, traz à tona uma lista de preocupações sobre o impacto do mesmo no meio social, dado os riscos que podem estar associados com sua utilização, como: formas de manipulação através da tecnologia, preconceitos, censura, discriminação social, violações de privacidade, direito de propriedade e muito mais (DANILO

& V. ALMEIDA, 2016). A IA trouxe algumas questões sociais e éticas, um estudo mostrou que em outubro de 2019, uma ferramenta desenvolvida na Amazon com o objetivo de recrutar pessoas para um determinado trabalho, provou-se tendenciosa contra candidatas do sexo feminino. Além disso, em 2021, o Facebook foi culpado por recomendações através de publicidades discriminadas. Muitos artigos resumem essas questões sociais e éticas relacionadas a IA como casos de informação, viés algorítmico, segurança e privacidade (T. ZHANG & Y. QIN, 2021). A adoção dos algoritmos trouxe vários desafios e ameaças, questões como privacidade, resultados tendenciosos e vieses algorítmicos são grandes barreiras para aplicações em larga escala de IA.

A automatização de processos nos setores governamentais é acompanhada de um pessimismo cibernético. Atualmente, com os riscos existentes da IA, várias armadilhas são motivo de preocupação para indivíduos (exemplo: projetistas) que precisam tomar uma decisão e utilizam a tecnologia como ferramenta de auxílio em suas atividades. A qualidade dos dados e os vieses em modelos baseados em IA são um perigo para a tomada de decisão, devido a tornarem a resolução tendenciosa. Manipulação e discriminação através da tecnologia são outras ameaças encontradas ao implantar a IA em órgãos governamentais (VALLE-CRUZ D & GARCÍA-CONTRERAS R, 2022). Problemas relacionados à “opacidade” algorítmica e a falta de explicabilidade de algumas técnicas são um dos maiores problemas enfrentados para a aceitação da IA nas organizações públicas. Esses problemas acabam resultando em perda do controle humano sobre a tomada de decisões e compreensão pública de como os algoritmos de caixa-preta geram seus resultados. As pessoas que são afetadas por decisões governamentais determinadas por uma IA exigem saber como as conclusões foram alcançadas, o que é um grande desafio, visto a natureza frequentemente opaca e não intuitiva dos algoritmos. O uso da IA pelo governo tem sido uma fonte de inúmeras injustiças, visto que os algoritmos tendem a apresentar erros e vieses, levando a decisões baseadas em informações incorretas e que agravam as desigualdades. Além disso, tomar decisões por meio de lógica rígida e baseadas em regras de algoritmo, viola o princípio de que as decisões do governo devem responder às circunstâncias de cada pessoa (GREEN B, 2022).

Por meio de uma série de trabalhos, pesquisas e apresentações públicas, a organização Gender Shades demonstrou como o reconhecimento facial disponível atualmente e vendidos por grandes empresas de tecnologias são substancialmente mais imprecisos com rostos escuros em geral, especialmente para rostos de mulheres de pele escura. A decorrência desse viés algoritmo resultou em possíveis tratamentos injustos pelas entidades que usaram essas APIs, incluindo agências governamentais (METCALF J & MOSS E, 2021).

Destacado por vários artigos, a noção de impacto (riscos) tem sido fundamental nas discussões sobre justiça algorítmica em várias conferências e workshops. Os debates mais pertinentes têm se concentrado sobre o “impacto díspar” de sistema algorítmicos, no que diz respeito ao que os sistemas replicam, estendem ou consolidam em relação a discriminação social. Impacto díspar é definido como resultados desiguais de um sistema algorítmico, que pode ou não ser intencional por parte de qualquer pessoa envolvida no desenvolvimento ou funcionamento desse sistema. Essa definição contrapõe o resultado discriminatório de uma IA que surge de decisões intencionais tomadas por aqueles envolvidos na construção ou operação de um sistema algorítmico (WATKINS EA & SINGH R, 2021).

Um dos estudos analisados (T. ZHANG & Y. QIN, 2021) segmentou os problemas e ameaças que o desenvolvimento e adoção da IA trouxeram em quatro grupos:

I. Casulos de Informação

Com base em dados, a IA consegue indicar informações personalizadas conforme o gosto de cada usuário, dinâmica amplamente utilizada no que diz respeito a recomendação de filmes, notícias, vídeos e outros conteúdos. Todavia, apesar de parecer algo integralmente benéfico, informações recomendadas por sistemas algorítmicos podem se restringir em uma gama específica e acabam causando casulos de informação, o que pode ser uma grave ameaça à democracia do povo e do país (CÉDRIC GOSSART, 2014).

II. Conluio algorítmico

Sistemas baseados em IA podem estabelecer o valor de um item com base no preço de seus concorrentes e dados sobre clientes, logo, os algoritmos podem ser instruídos e utilizados para conluio, uma vez que os mesmos podem atingir níveis significativos de colusão tácita, “ferindo” gravemente os direitos e interesses dos consumidores (EZRACHI, ARIEL & MAURICE E. STUCKE, 2017).

III. Viés algorítmico

O viés algorítmico é identificado quando um sistema produz uma resposta injusta e discriminatória em relação a alguém ou grupo social. O caso a seguir foi apresentado pelo autor T. Zhang no intuito de exemplificar um viés algorítmico: “COMPAS é um sistema de

aprendizado de máquina amplamente utilizado nos EUA, que pode prever potenciais futuros criminosos com base em regras confidenciais e dados históricos. Em 2016, Jeremy Goldhaber Fiebert mostrou que o COMPAS é tendencioso contra negros réus, já que o sistema é mais propenso a rotular negros réus como maior risco de cometer crimes violentos no futuro” (MARISA VASCONCELOS, 2018).

IV. Desafios de segurança e privacidade

Como qualquer software, sistemas baseados em inteligência artificial não estão imunes a ataques maliciosos. Vale destacar que sistemas de IA são vulneráveis a ataques emergentes específicos, como aprendizado contraditório e envenenamento de dados (DANILO VASCONCELLOS VARGAS & KOUICHI SAKURAI, 2019).

4.2.1 SÍNTESE: RISCOS DOS ALGORITMOS

A tabela a seguir mostra os principais riscos, identificados a partir dos estudos, que os algoritmos podem apresentar.

| | |
|----|-----------------------------------|
| R1 | Viés Algorítmico |
| R2 | Manipulação através da tecnologia |
| R3 | Discriminação Social |
| R4 | Falta de Segurança e Privacidade |
| R5 | Casulos de Informação |
| R6 | Impacto díspar |

Tabela 4: Principais riscos dos algoritmos

Com o aumento do uso de algoritmos para cumprir tarefas complexas, vários riscos e problemas podem surgir conforme demonstrado nesta seção. Para lidar com esses riscos, o processo de governança de algoritmos deve ser considerado.

4.3 IMPORTÂNCIA DA GOVERNANÇA DE ALGORITMOS

Esta seção tem como objetivo demonstrar a necessidade de uma governança para os algoritmos, dado o impacto e riscos que os mesmos podem causar conforme demonstrado nas seções anteriores.

Em razão do impacto social a partir das decisões algorítmicas, surge um importante questionamento: os algoritmos precisam de métodos para sua governança, ou as regras e normas de governança da internet são o suficiente? Técnicas de governança devem ser adotadas com o intuito de proteger grupos vulneráveis da sociedade e aumentar o bem-estar público. Para alcançar tal governança faz-se necessário uma congregação entre a sociedade, governos e empresas privadas (F. FILGUEIRAS & R. F. MENDONÇA, 2022). A integração dos algoritmos em vários processos sociais e econômicos pode ser bastante valiosa, desde que instrumentos que permitam algum tipo de governança fossem projetados. De acordo com Virgílio Almeida: “Isso poderia impedir que os sistemas algorítmicos afetassem negativamente o equilíbrio de poder, em favor daqueles que podem exercer o poder real sobre como são usados, além de maximizar seus benefícios e reduzir seus riscos”. Geralmente, quando surge uma ideia de criar uma governança para algo, ela tende a reduzir os problemas causados por tal. No caso da governança para IA, deve-se tentar preservar a eficácia dos algoritmos, reduzindo seus resultados indesejáveis (DANILO & V. ALMEIDA, 2016).

A falta de confiança nos sistemas e aplicativos de IA geram boa parte dos problemas relacionados aos algoritmos. As pessoas não confiam nas decisões tomadas por sistemas algorítmicos, o que promove um impacto negativo na adoção do mesmo. Para promover ampla adoção da IA independente do ambiente, faz-se necessário garantir que os aplicativos de IA sejam confiáveis, o que tem sido um consenso entre governos e organizações (T. ZHANG & Y. QIN, 2021). Diariamente os algoritmos são utilizados para tomar alguma decisão nas vidas das pessoas, devido a isto, há uma convicção cada vez maior da necessidade de garantir a confiança dessas decisões. Organizações e governos em todo mundo estão tomando medidas para alcançar esse objetivo, pois, alcançar a IA confiável é o que precisamos. Entretanto, não há uma definição uniforme do que seria um sistema algorítmico confiável. Alguns estudos demonstraram que os requisitos para conquistar uma IA confiável são: garantir robustez e segurança, governança de dados e privacidade, transparência, interpretabilidade e responsabilidade (T. ZHANG & Y. QIN, 2021).

A maioria dos sistemas de IA são considerados “caixas pretas”, causando uma grande assimetria de informações entre desenvolvedores e usuário final. Parte da sociedade

expressaram preocupação com as consequências não intencionais e possíveis prejuízos advindos do uso da tecnologia. Enquanto a IA tem o potencial de moldar a vida de bilhões de pessoas, apenas alguns especialistas realmente conseguem entender seu funcionamento. Os sistemas baseados em IA costumam ser vistos como inescrutáveis, o que resulta em grandes assimetrias de informação entre desenvolvedores de tais sistemas e outras partes que fazem uso do mesmo (U. GASSER & V. A. F. ALMEIDA, 2017). Com o objetivo de preencher essa lacuna de informação, vários estudos propuseram uma estrutura conceitual para pensar sobre governança de algoritmos, de forma resumida, garantir transparência, responsabilidade e explicabilidade foram considerados um dos princípios básicos para cogitar um regime de governança para IA.

Em um artigo, os autores Inioluwa Deborah Raji e Buolamwini evidenciaram que APIs que são previamente auditadas tendem a ser significativamente menos tendenciosas do que aquelas que não são auditadas antecipadamente, indicando que auditorias críticas de terceiros podem resultar em pressão significativa para mudanças nas práticas de desenvolvimento de sistemas. Embora essas auditorias possam ser bem-sucedidas em forçar as empresas de tecnologia a mudar suas práticas, Raji argumenta sobre a possibilidade das empresas, organizações ou terceiros realizarem, de forma obrigatória, avaliações dos impactos dos algoritmos (AIAs), o que pode trazer um resultado mais significativo em relação a uma auditoria simples (INIOLUWA DEBORAH RAJI & JOY BUOLAMWINI, 2019).

Diante dos ganhos e ameaças a partir da tomada de decisão algorítmica, os formuladores de políticas exploram abordagens regulatórias que podem permitir que os governos alcancem os benefícios dos algoritmos enquanto evitam os riscos dos mesmos. Como os governos adotam algoritmos para informar ou tomar decisões, a regulamentação faz-se necessária para evitar a produção de injustiças e a violação de princípios jurídicos fundamentais (GREEN B, 2022). Para que a confiança em IA seja construída no setor público, dado que muitos membros da organização suspeitam de sistemas baseados em IA, os governos devem demonstrar às partes interessadas e aos indivíduos afetados pelos algoritmos que o mesmo foi projetado e implementado com transparência, responsabilidade, consistência e sem preconceito. De formuladores de políticas a salas de diretorias, a ideia de regular e implementar uma AIA como uma forma de responsabilidade algorítmica está ganhando força, pois, essas avaliações são vistas eventualmente como úteis para antecipar, evitar e mitigar as consequências negativas dos sistemas de decisão algorítmica (SINGH R & ELISH MC, 2021). Para superar o desafio da transparência e abordar as preocupações a respeito da IA injusta ou tendenciosa, as instituições governamentais precisam estipular padrões que deslindem claramente quais informações

qualquer empresa deve divulgar sobre o design e a operação de seu algoritmo. Uma vez estipulado e administrado de forma criteriosa, a aquisição de produtos e serviços de IA pode tornar o governo mais transparente e confiável, otimizando e padronizando o processo de tomada de decisão para minimizar erros, inconsistências e preconceitos decorrentes de humanos nos sistemas de IA (SINGH R & ELISH MC, 2021).

4.3.1 SÍNTESI: IMPORTÂNCIA DA GOVERNANÇA DE ALGORITMOS

Uma das questões mais abordadas nos artigos analisados foi a governança de IA proporcionando um sentido ético, moral e não discriminatório. Em razão disso, vários países e órgãos internacionais têm se dedicado em gerar diretrizes sobre a utilização de IA. Sendo assim, é extremamente necessário que um processo de governança seja considerado, e o mesmo deve garantir o uso ético da IA, para que os direitos humanos não sejam violados.

4.4 PROCESSOS DA GOVERNANÇA DE ALGORITMOS

A questão de pesquisa QP4 visa demonstrar quais as técnicas fazem parte do processo da governança de algoritmos. Essa questão é primordial para entender as estratégias que são necessárias para implantar uma governança e seus desafios.

Conforme demonstrado na seção anterior, existe uma importância no que diz respeito à governança de algoritmos e por algoritmos, pois ambas concepções envolvem questões sociais e éticas. A forma como essas governanças devem ser estruturadas está em constante debate devido às questões sociais e interesse público de cada ambiente. Assim como instituições, os algoritmos possuem um objetivo no processo de solução de problemas ou tomada de decisões. Entendendo os algoritmos como instituições, podemos afirmar que os mesmos devem possuir papéis relacionados à transparência e responsabilidade. Conhecendo o funcionamento dos algoritmos, pode-se afirmar que o design é um direito essencial que um cidadão pode exigir para que o mesmo possa entender como o sistema algorítmico, que influencia seu meio, tomou tal decisão. Saber como, quando e com que objetivo um algoritmo que faz parte do meio social foi criado, torna a ordem dos algoritmos mais democrática, além de produzir formas de reflexão que conduzem as ações dos desenvolvedores de sistemas (F. FILGUEIRAS & R. F. MENDONÇA, 2022). Pensar em algoritmos como instituições, exige entender os valores sociais que fazem parte do mesmo e a maneira como eles influem na realidade, tal como a maneira que são transformados. Do mesmo modo que instituições políticas devem estar

relacionadas com princípios que compõem a ação coletiva, os algoritmos também precisam integrar tais elementos, levando em conta que esses princípios podem mudar com o tempo e devem ser ajustados conforme necessidade. (F. FILGUEIRAS & R. F. MENDONÇA, 2022).

Para certificar uma governança eficaz e buscar melhorias para a mesma, órgãos concentram-se na contratação governamental como meio de promover o uso ético da IA, ou seja, sua utilização responsável, respeitando os valores-chave, como transparência, responsabilidade, não discriminação, privacidade e segurança. O uso governamental de algoritmos privados seria por meio da elaboração de cláusulas contratuais que obrigam os contratantes a cumprir os padrões éticos, além de divulgar informações suficientes para satisfazer as preocupações do público em geral. Os estudos mostraram que ao incluir encargos de IA ética e transparente em contratos públicos, funcionários do governo poderiam ajudar a promover maior legitimidade pública na utilização da IA, assim como promover um conjunto mais amplo de normas sobre o uso responsável de técnicas de aprendizado de máquina (L. M. B. DOR & C. COGLIANESE, 2021). Várias instituições governamentais já estão desenvolvendo conjuntos semelhantes de diretrizes ou padrões éticos. A União Europeia, por exemplo, estabeleceu um processo para reconhecer que os valores em que se baseiam nossas sociedades precisam ser totalmente integrados na forma como a IA se desenvolve. A UE emitiu um Regulamento Geral de Proteção de Dados (GDPR) que aborda questões de privacidade, transparência e responsabilidade levantado pela IA, além disso, o regime oferece um “direito à explicação” que obriga as empresas a explicar a finalidade de um algoritmo e o tipo de dados que o mesmo utiliza para tomar decisões automatizadas (U. GASSER & V. A. F. ALMEIDA, 2017). Nos Estados Unidos, o Escritório de Política de Ciência e Tecnologia da Casa Branca aconselhou as agências a se concentrarem na justiça, transparência e gestão de riscos em seus usos de IA. Os contratos de produtos e serviços relacionados a sistemas algorítmicos devem, no mínimo, exigir que as empresas se comprometam a fornecer acesso às informações necessárias ao longo de seus contratos, em alguns casos, durante o período em que o governo continuar a usar o produto ou sistema desenvolvido pelos contratantes. Esses contratos também podem exigir que os desenvolvedores sejam submetidos a auditorias governamentais ou públicas, no intuito de expor a programação e dados que foram utilizados na implementação do sistema, com o objetivo de monitorar os impactos não intencionais de preconceitos baseados em raça ou gênero ou resultado errôneos. À medida que mais órgãos governamentais implantam sistemas de IA a partir de um contrato, as cláusulas contratuais acabam “evoluindo” para enquadrar as normas éticas que uma IA deve seguir. Esses princípios se tornarão cada vez mais importantes, incentivando outras empresas a tomarem medidas semelhantes para seguir a

tendência da indústria. A sociologia conhece esse processo como isomorfismo mimético (L. M. B. DOR & C. COGLIANESE, 2021). As empresas já começaram a pensar de forma mais crítica sobre suas aquisições de sistemas algorítmicos, e esperam que seus fornecedores de sistemas baseados em IA cumpram certas normas éticas. Quando a aquisição é utilizada como técnica para a governança de algoritmos, ela causa um efeito amplo, que incentiva a indústria como um todo, de forma que garanta que os sistemas algorítmicos atendam às metas de desempenho, e ao mesmo tempo, conquiste a confiança e aceitação do público.

A governança de algoritmos, muitas vezes, é baseada em ferramentas que não funcionam no algoritmo em si, mas em elementos de seu ambiente. Os conjuntos de dados são utilizados como parte central das tarefas executadas pelos algoritmos, trazendo à tona a necessidade de validar se os mesmos estão sendo utilizados de forma lícita, ética e não demonstram qualquer tipo de viés. Os estudos sobre modelagem algorítmica demonstraram que dados incorretos e tendenciosos produzem resultados tendenciosos, cenário que pode ser comparado com um conceito bastante comum da ciência da computação: “Entra lixo, sai lixo” (UDOH ES, 2020).

A “IA confiável” foi uma concepção comum nos artigos desse estudo, governos de todo o mundo vêm realizando ações para promover tal ideia. Em janeiro de 2019, a Proteção de Dados Pessoais de Cingapura apresenta uma estrutura de governança para algoritmos, que tem por objetivo promover a confiança dos usuários na adoção e implantação da IA. A estrutura propõe alguns princípios éticos e ações como: centrada no ser humano, explicável, transparente e justa. Em abril de 2019, o grupo de especialistas de Alto Nível da União Europeia em IA publicou oficialmente as Diretrizes Éticas para Inteligência Artificial (Comissão Europeia, 2019). O regulamento apresenta o conceito de IA confiável, uma base para os desenvolvedores utilizarem em suas construções de aprendizado de máquina. A IA confiável deve ser legal, ética e robusta e existem sete requisitos principais que os sistemas de IA devem seguir para serem considerados confiáveis segundo a Comissão Europeia. A lista a seguir visa especificar cada requisito:

I. Ação e Supervisão Humana

Os sistemas de IA devem apoiar a autonomia e a tomada de decisões dos seres humanos. Isto exige que os sistemas de IA funcionem como facilitadores de uma sociedade democrática, apoiando o usuário e promovendo os direitos fundamentais, além de permitir a supervisão humana (Comissão Europeia, 2019).

II. Transparência

Os dados utilizados no sistema, assim como os modelos de negócio devem ser transparentes. Métodos de rastreabilidade podem ajudar a atingir esse princípio. Os sistemas algorítmicos devem ser explicados de maneira personalizada às partes interessadas. Os usuários do sistema devem ser informados que estão atuando com um sistema de IA e devem ser informados sobre a capacidade e deficiências do sistema (Comissão Europeia, 2019).

III. Diversidade, não discriminação e equidade

A inclusão e diversidade têm de estar presentes em todo o ciclo de vida do sistema, para que a IA seja acessível a todos, independentemente de qualquer deficiência, raça ou gênero (Comissão Europeia, 2019).

IV. Robustez e Segurança

Os algoritmos devem ser desenvolvidos de forma preventiva, considerando os riscos intencionais e não intencionais, para que a robustez do sistema seja garantida. Eles precisam ser seguros, de forma que haja um plano alternativo caso algo dê errado, além de serem precisos e confiáveis (Comissão Europeia, 2019).

V. Responsabilidade

Este princípio exige que sejam criados mecanismos que assegurem a responsabilidade por parte de quem desenvolveu o sistema e a responsabilização de quem está utilizando o mesmo. Devem haver responsáveis pelos resultados da IA. O conceito de “auditabilidade” prega que os algoritmos devem ser avaliados, assim como seus dados e processos de design, ações que desempenham um papel fundamental para responsabilidade algorítmica (Comissão Europeia, 2019).

VI. Privacidade e governação dos dados

O direito à privacidade deve ser respeitado em todo momento, dados pessoais utilizados pelo sistema não devem ser expostos de forma indevida. Técnicas que garantam uma governança de dados adequada devem ser consideradas, tendo em conta a qualidade e integridade dos dados, assim como o acesso legítimo dos dados (Comissão Europeia, 2019).

VII. Bem-estar social e ambiental

A sustentabilidade e a responsabilidade ecológica dos sistemas algorítmicos devem ser incentivadas, pois a IA deve beneficiar todos os seres humanos, incluindo as gerações futuras. Além disso, deve levar em consideração o meio ambiente, abrangendo outros seres vivos, e seu impacto social (Comissão Europeia, 2019).

Vários gigantes da tecnologia como Microsoft e IBM estão adotando os princípios mencionados anteriormente para atingir a IA confiável. Em 2018, a Microsoft propôs 6 princípios para o uso responsável da IA, como justiça, confiabilidade e segurança, privacidade e segurança, inclusão, transparência e responsabilidade. A IBM também propôs o conceito de IA confiável para que os resultados dos sistemas de IA tenham uma maior credibilidade (T. ZHANG & Y. QIN, 2021).

O termo “avaliação de impacto algorítmico” foi bastante mencionado nos estudos analisados, dado sua importância no contexto governança de algoritmos. A AIA pode ser definida como uma caixa de ferramentas de possíveis processos de responsabilidade, incluindo auditorias algorítmicas. A ideia geral de uma AIA é documentar a construção e o possível impacto de um sistema de decisão algorítmica (ADS), para mitigar possíveis danos a indivíduos e comunidades, principalmente indivíduos e comunidades vulneráveis (SINGH R & ELISH MC, 2021). As avaliações de impacto algorítmico não concebem uma responsabilização direta, a menos que os métodos utilizados para determinar seus impactos sejam submetidos a um fórum que tenha capacidade de analisar e exigir mudanças, caso necessário, em tais avaliações e até no próprio sistema. Para estabelecer a responsabilidade na gestão da AIA, é fundamental que exista uma relação entre o sistema algorítmico e um fórum capaz de reivindicar modificações em tais sistemas. Conforme instituições, empresas e governos começem a implantar as práticas de AIA como uma obrigação, mais perto estaremos de garantir que os sistemas de IA tenham como parte de seu desenvolvimento as ideias de equidade, transparência e justiça. Segundo o autor Metcalf J e Moss E, “A AIA é uma prática de governança que estrutura e sustenta uma relação de prestação de contas, e têm sido amplamente estabelecidas em domínios análogos aos

sistemas algorítmicos” (METCALF J & MOSS E, 2021). O governo Canadense passou a exigir que suas agências que fazem uso de algoritmos adotem as convenções impostas pela AIA (KARLIN M. & CORRIVEAU, 2018). Grandes empresas como Google e Facebook estão promovendo avaliação de impacto algorítmico no que diz respeito aos direitos humanos, com o objetivo de identificar os possíveis danos de suas plataformas e produtos (WAROFKA, 2018). Proposta no Congresso dos EUA em 2019, a Lei de Responsabilidade Algorítmica exigirá que empresas realizem avaliações de impactos em seus ADS que afetam ambientes sociais (H.R. 2231, 2019). A avaliação de impacto algorítmica “remove” a dependência de legislaturas lentas que tentam delinear o que os ADSs podem ou não fazer, uma vez que administradas por uma agência governamental responsável, a AIA pode definir os impactos positivos e negativos antes do desenvolvimento do algoritmo (WATKINS EA & MOSS E, 2021).

Sem dúvida a AIA é uma técnica atraente, mas como qualquer concepção nova está sujeita a riscos e indagações. A governança por meio das avaliações de impactos algorítmicos defronta-se com o risco de informações omissas e falta de fóruns externos que possam exigir responsabilização por danos. Do mesmo modo, as auditorias carecem de conhecimento para identificar os impactos dos sistemas, uma vez que os auditores não possuem acesso formal ao funcionamento interno do mesmo (METCALF J & MOSS E, 2021). O que constitui uma avaliação? Um impacto? Um ADS? Quem decide? Os algoritmos usados por empresas privadas devem estar sujeitos às mesmas formas de responsabilização como aqueles usados por instituições públicas? São perguntas levantadas pelos autores Singh R e Elish MC, para afirmar que não existe uma decisão clara, institucional, regulatória e judicial em torno de qualquer visão do que é uma AIA (SINGH R & ELISH MC, 2021).

Uma peça fundamental e bastante discutida da governança de algoritmos, refere-se a supervisão humana nas tomadas de decisões por algoritmos. As políticas de supervisão humana, por exemplo, permitem que os governos usem algoritmos, mas apenas se um ser humano possuir alguma forma de supervisionar ou controlar a decisão final do mesmo. Isso implica nos algoritmos não poderem “fazer julgamentos” por conta própria, e sim, auxiliar os humanos na tomada de decisão final. Essas políticas visam garantir que os humanos desempenhem um papel de controle de qualidade, no intuito de evitar previsões algorítmicas equivocadas ou tendenciosas. Segundo o autor Green B, essas políticas “visam proteger os direitos humanos e a dignidade, mantendo um humano no processo da tomada de decisão automatizada.” (GREEN B, 2022). A suposição de adotar algoritmos em conjunto da supervisão humana, pode permitir que os governos obtenham o melhor dos dois lados: a consistência, objetividade e precisão dos algoritmos, assim como a análise crítica e contextual do ser humano. Legislações recentes

assumem que a supervisão humana protege o governo contra os danos que os algoritmos podem causar. A Comissão Europeia, em sua proposta de Lei para IA, afirmou que a supervisão humana (acompanhada de outros métodos) faz-se necessária para mitigar os riscos aos direitos fundamentais (COMISSÃO EUROPEIA, 2021). Muitos governos assumem a supervisão humana como um diferenciador para que os algoritmos sejam implantados em seus órgãos e instituições. O Regulamento Geral de Proteção de Dados restringe decisões significativas que considerem apenas a tomada de decisão automatizada (PARLAMENTO EUROPEU E CONSELHO DA UNIÃO EUROPEIA, 2016). O Governo do Canadá exige que agências federais que usam algoritmos de tomada de decisão de alto risco (impactam de alguma forma indivíduos ou comunidades), tenham um processo de intervenção humana, o qual apenas uma pessoa pode tomar a decisão final (GOVERNO DO CANADÁ, 2021). No estado de Washington (EUA) as agências governamentais e locais só podem utilizar tecnologias de reconhecimento facial em determinados casos, uma vez que façam parte de uma decisão de alto impacto, tais casos são sujeitos a uma supervisão humana (LEGISLATIVO DO ESTADO DE WASHINGTON, 2020).

Apesar do destaque que governos e legislações têm dado para a supervisão humana como ferramenta da governança de algoritmos para mitigar riscos, a qualidade funcional dessa ideia não foi completamente indagada. Os legisladores trazem valores como dignidade e direitos humanos como motivadores para a política da supervisão humana de algoritmos, porém não trazem evidências na prática que um ser humano realmente irá conseguir promover esses valores (GREEN B, 2022). A ideia de inserir um humano no processo de decisão algorítmica traz a sensação de satisfazer direitos e princípios sociais, contudo, pesquisas realizadas em sistemas sociotécnicos demonstram frequentemente que tecnologias e pessoas não interagem como esperado (SUCHMAN L & BLOMBERG J, 2019).

5 CONCLUSÃO E TRABALHOS FUTUROS

A geração de valor que os algoritmos trazem para a sociedade é notória. O poder computacional que possuímos hoje proporciona que atividades que seriam ineficazes ou impossíveis de serem realizadas por um ser humano venham a acontecer. Contudo, conforme demonstrado neste trabalho, os algoritmos podem trazer riscos para a sociedade, discriminação social, vieses algorítmicos, violação de dados e resultados tendenciosos são algumas das possíveis ameaças que os sistemas baseados em IA podem apresentar. Tudo que pode impactar, seja de maneira direta ou indireta, um indivíduo ou comunidade precisa ser governado, dado que o processo de governar possui a função de regular e responsabilizar ações que atuam sobre algo ou alguém. O processo de governança de algoritmos precisa ser considerado para que as vantagens de utilizar sistemas tecnológicos sejam maximizadas e as ameaças mitigadas.

É evidente que não existe apenas uma única solução para atingirmos a governança de algoritmos, entretanto, vale ressaltar que basear-se em uma regulamentação intuitivamente atraente, mas ineficaz, pode levar ao pior dos dois mundos: o problema pode permanecer subjacente, mas a presença da regulamentação leva à percepção de que o problema foi resolvido. Apesar dos desafios mencionados em relação a supervisão humana como técnica para a governança de algoritmos, a mesma deve ser considerada como um possível método aplicável, caso os reguladores que introduzirem tal solução consigam garantir de forma prática que seus regulamentos realmente irão fornecer as proteções e benefícios propostos.

As medidas realizadas para regular os algoritmos do governo e do setor privado, devem estar intimamente “ligadas” aos contextos sociais em que estão inseridos, de forma que os padrões que refletem o interesse público sejam definidos e considerados desde a implementação do sistema algorítmico. Estabelecer um órgão fiscalizador que possa garantir a integridade e o cumprimento dos valores de interesse do público na utilização de algoritmos trará grande valia, uma vez que este processo tende a quebrar a “não confiança” da população em relação às tomadas de decisão por parte dos sistemas baseados em IA. Quanto mais abordagens regulatórias voltadas para questões sociotécnicas forem discutidas e aplicadas, mais próximos estaremos de implantar uma governança de algoritmos que promova os direitos humanos sem prejudicar os benefícios que os algoritmos trazem para a sociedade.

5.1 TRABALHOS FUTUROS

Algumas sugestões de possíveis trabalhos futuros são listadas a seguir:

- Trabalho para mapear os desafios da governança de algoritmos, assim como suas possíveis soluções.
- Estudo com avaliações comparativas das práticas utilizadas na governança de algoritmos, com o intuito de buscar os melhores métodos.

REFERÊNCIAS

- Warofka, A. “**An Independent Assessment of the Human Rights Impact of Facebook in Myanmar.**” <https://about.fb.com/news/2018/11/myanmar-hria/>, 2018.
- V. Almeida; F. Filgueiras; R. F. Mendonça. **Algorithms and Institutions: How Social Sciences Can Contribute to Governance of Algorithms.** <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9775553>, 2022.
- Ezrachi, Ariel, and Maurice E. Stucke. "Algorithmic collusion: Problems and counter-measures." [https://one.oecd.org/document/DAF/COMP/WD\(2017\)25/en/pdf](https://one.oecd.org/document/DAF/COMP/WD(2017)25/en/pdf), 2017.
- Green B. **The flaws of policies requiring human oversight of government algorithms.** <https://www.sciencedirect.com/science/article/pii/S0267364922000292>, 2022.
- Lilian Bambirra Assis; Juliana Oliveira Andrade; Antônio Carvalho Neto; Betânia Tanure; Alexandre Carrieri. **O isomorfismo entre executivos nas maiores empresas brasileiras.** Gerais, Rev. Interinst. Psicol. vol.3 no.1 Juiz de fora jul. 2010.
- Joy Buolamwini; Timnit Gebru. **Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification.** <http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>, 2018.
- Gossart, Cédric. "Can digital technologies threaten democracy by creating information cocoons?" Transforming politics and policy in the digital age. IGI Global, 2014.
- Council of the European Union. **General data protection regulation. Off. J. Eur. Union.** <https://eur-lex.europa.eu/eli/reg/2016/679/oj>, 2016.
- Council of the European Union. **General data protection regulation. Off. J. Eur. Union.** <https://eur-lex.europa.eu/eli/reg/2016/679/oj>, 2016.
- Valle-Cruz D,García-Contreras R,Muñoz-Chávez JP. **Mind the Gap: Towards an Understanding of Government Decision-Making Based on Artificial Intelligence.** <https://doi.org/10.1145/3543434.3543445>, 2022.
- Souvik Das. **The Origin and Evolution of Algorithms.** <https://www.digit.in/features/science-and-technology/the-origin-of-algorithms-30045.html>, 2016
- Inioluwa Deborah Raji and Joy Buolamwini. **Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products.** <https://doi.org/10.1145/3306618.3314244>, 2019.

P.Domingos. **The Master Algorithm**, Basic Books, 2015.

D. Doneda; V. A. F. Almeida. **What Is Algorithm Governance?**
<https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7529042>, 2016

Paul Dourish. **Algorithms and their others: Algorithmic culture in context. Big Data & Society**, 2016.

Udoh ES. **Is the Data Fair? An Assessment of the Data Quality of Algorithmic Policing Systems**. <https://doi.org/10.1145/3428502.3428503>, 2020.

Electronic Privacy Information Center, “**Algorithmic Transparency: End Secret Profiling**”. <https://epic.org/algorithmic-transparency>, 2015.

Ethics guidelines for trustworthy AI. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>, 2019.

Ethics guidelines for trustworthy AI. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>, 2019.

European Commission. Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (**Artificial Intelligence Act**). <https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence-artificial-intelligence>, 2021.

L. Floridi. “**Soft ethics, the governance of the digital and the General Data Protection Regulation**”. Philos. Trans. Roy. Soc. A, Math. Phys. Eng. Sci., vol. 376, no. 2133, 2018.

U. Gasser; V. A. F. Almeida. **A Layered Model for AI Governance**.
<https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8114684>, 2017.

Government of Canada. **Directive on Automated Decision-Making**. <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592>, 2021.

H.R. 2231. **Algorithmic Accountability Act**. 116th Congress, 2019.

Metcalf J, Moss E, Watkins EA, Singh R, Elish MC. **Algorithmic Impact Assessments and Accountability: The Co-Construction of Impacts**.
<https://doi.org/10.1145/3442188.3445935>, 2021.

Su, Jiawei, Danilo Vasconcellos Vargas, and Kouichi Sakurai. “**One Pixel Attack for Fooling Deep Neural Networks**.” IEEE Transactions on Evolutionary Computation, 2019.

Irion K. **Algorithms Off-Limits? If Digital Trade Law Restricts Access to Source Code of Software Then Accountability Will Suffer.** <https://doi.org/10.1145/3531146.3533212>, 2021.

Barbara Kitchenham. **Procedures for Performing Systematic Reviews.** <https://www.inf.ufsc.br/~aldo.vw/kitchenham.pdf>, 2004.

Marek Kowalkiewicz. **How did we get here? The story of algorithms.** <https://towardsdatascience.com/how-did-we-get-here-the-story-of-algorithms-9ee186ba2a07>, 2019.

L. M. B. Dor; C. Coglianese. **Procurement as AI Governance.** <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9540751>, 2021.

Suchman L, Blomberg J, Orr JE, Trigg R. **The Data Protection and Privacy Act.** <https://ict.go.ug/wp-content/uploads/2019/03/Data-Protection-and-Privacy-Act-2019.pdf>, 2019.

Karlin, M. and Corriveau. “**The Government of Canada’s Algorithmic Impact Assessment: Take Two.**” <https://medium.com/@supergovernance/the-government-of-canadas-algorithmic-impact-assessment-take-two-8a22a87acf6f>, 2018.

Vasconcelos, Marisa, et al. “**Modeling Epistemological Principles for Bias Mitigation in AI Systems: An Illustration in Hiring Decisions.**” Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society, 2018.

Bevir, Mark. **Governance: A very short introduction.** Oxford, UK: Oxford University, 2012.

S. U. Noble. **Algorithms of Oppression: How Search Engines Reinforce Racism.** New York, 2018.

Watkins EA,Moss E,Metcalf J,Singh R,Elish MC. **Governing Algorithmic Systems with Impact Assessments: Six Observations.** <https://doi.org/10.1145/3461702.3462580>, 2021.

Washington State Legislature. **SB 6280 - 2019-20: Concerning the Use of Facial Recognition Services.** <http://lawfilesext.leg.wa.gov/biennium/2019-20/Pdf/Bills/Senate%20Passed%20Legislature/6280-S.PL.pdf>?q=20210513071229, 2020.

F. Zambonelli; F. Salim; S. W. Loke; W. De Meuter; S. Kanhere. **Algorithmic Governance in Smart Cities: The Conundrum and the Potential of Pervasive Computing Solutions.** <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8371566>, 2018.

T. Zhang; Y. Qin; Q. Li. **Trusted Artificial Intelligence: Technique Requirements and Best Practices.** <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9599411>, 2021.

APÊNDICE A — PROTOCOLO UTILIZADO NA REVISÃO SISTEMÁTICA DA LITERATURA

O protocolo abaixo foi criado e utilizado como guia para a realização da revisão sistemática deste trabalho. Segue abaixo as etapas que foram definidas e aplicadas:

PARTE 1: PLANEJAMENTO

Questões:

- QP1: Quais os impactos sociais dos algoritmos?
- QP2: Quais os riscos que os algoritmos podem apresentar?
- QP3: Qual a necessidade de uma governança para os Algoritmos?
- QP4: Quais processos fazem parte da governança de algoritmos?

PARTE 2: DEFINIÇÃO DE PALAVRAS, STRING E FONTE DE PESQUISA

Palavras-chaves: governance algorithm, society, transparency

Sinônimos por palavras:

- **governance algorithm:** governing algorithms, algorithmic governance, algorithm regulation
- **society:** organization
- **transparency:** evidence, accountability, regulation

String de busca:

String v1:

(“governing algorithms” OR “governance algorithm” OR “algorithmic governance” OR “algorithm regulation”) AND (“society” OR “organization”) AND (“transparency” OR “evidence” OR “accountability” OR “regulation”)

String v2:

(“algorithm governance” OR “governance algorithm” OR “algorithmic governance”) AND (“society” OR “organization”) AND (“transparency” OR “model” OR “accountability” OR “regulation” OR “privacy”)

String v3:

(“algorithm governance” OR “governance algorithm” OR “algorithmic governance”) AND (“social” OR “society” OR “organization”) AND (“transparency” OR “model” OR “accountability” OR “regulation” OR “privacy”)

String v4:

(“algorithm governance” OR “algorithmic governance”) AND (“social” OR “organization”) AND (“model” OR “method” OR “technique” OR “regulation” OR “privacy”)

String Final:

(“algorithm governance” OR “algorithmic governance”) AND (“society” OR “social” OR “organization”) AND (“model” OR “process” OR “regulation” OR “privacy”)

Fontes de Pesquisa:

<http://www.sciencedirect.com> -

<https://www.sciencedirect.com/search?qs=%28%E2%80%9Calgorithm%20governance%E2%80%9D%20OR%20%E2%80%9Calgorithmic%20governance%E2%80%9D%29%20AND%20%28%E2%80%9Csocial%E2%80%9D%20OR%20%E2%80%9Corganization%E2%80%9D%29%20AND%20%28%22model%E2%80%9D%20OR%20%22method%22%20OR%20%E2%80%9Ctechnique%E2%80%9D%20OR%20%22regulation%22%20OR%20%E2%80%9Cprivacy%E2%80%9D%29>

<http://ieeexplore.ieee.org> -

[https://ieeexplore.ieee.org/search/searchresult.jsp?queryText=\(%E2%80%9Calgorithm%20governance%E2%80%9D%20OR%20%E2%80%9Calgorithmic%20governance%E2%80%9D\)%20AND%20\(%E2%80%9Csocial%E2%80%9D%20OR%20%22society%22%20OR%20%E2%80%9Corganization%E2%80%9D\)%20AND%20\(%22model%E2%80%9D%20OR%20%22method%22%20OR%20%E2%80%9Cprocess%E2%80%9D%20OR%20%22regulation%22%20OR%20%E2%80%9Cprivacy%E2%80%9D\)&newsearch=true&type=alt2&highlight=true&returnType=SEARCH&matchPubs=true&pageNumber=1&returnFacets=ALL](https://ieeexplore.ieee.org/search/searchresult.jsp?queryText=(%E2%80%9Calgorithm%20governance%E2%80%9D%20OR%20%E2%80%9Calgorithmic%20governance%E2%80%9D)%20AND%20(%E2%80%9Csocial%E2%80%9D%20OR%20%22society%22%20OR%20%E2%80%9Corganization%E2%80%9D)%20AND%20(%22model%E2%80%9D%20OR%20%22method%22%20OR%20%E2%80%9Cprocess%E2%80%9D%20OR%20%22regulation%22%20OR%20%E2%80%9Cprivacy%E2%80%9D)&newsearch=true&type=alt2&highlight=true&returnType=SEARCH&matchPubs=true&pageNumber=1&returnFacets=ALL)

<http://dl.acm.org> -

<https://dl.acm.org/action/doSearch?fillQuickSearch=false&target=advanced&expand=dl&AllField=AllField%3A%28%28%22algorithm+governance%22+OR+%22algorithmic+governance%22%29+AND+%28%22social%22+OR+%22organization%22%29+AND+%28%22model%22+OR+%22method%22+OR+%22technique%22+OR+%22regulation%22+OR+%22privacy%22%29%29>

PARTE 3: DEFININDO E APLICANDO FILTROS

Os seguintes critérios de exclusão serão aplicados:

- CE01 - Artigos que não estejam no período de 2016 a 2022;
- CE02 - Artigos secundários (outra RSL ou MS);
- CE03 - Artigos com conteúdo pago;
- CE04 - Artigos duplicados ou similares;
- CE05 - Artigos que não estejam relacionados com as perguntas de pesquisa
- CE06 - Artigos que não estejam em inglês

PARTE 4: QUALIFICANDO OS ESTUDOS

- Contexto Claro
- Metodologia bem definida
- Aplicação Prática
- Discussão relevantes e consistentes

Cada estudo deverá ter uma "nota de qualidade", somando as notas de cada critério. Para valorar os critérios foi utilizado a seguinte escala: "0 - não atende; 0.5 - atende parcialmente; 1.0 - atendente totalmente".

PARTE 5: SELEÇÃO DE ARTIGOS

Os estudos com notas iguais ou superior a 50% da maior nota e qualidade deverão ser selecionados para a lista final de artigos, ordenada pelo ano de publicação.

PARTE 6: COLETA DE EVIDÊNCIAS

Criar planilha com a lista dos artigos finais selecionados, incluindo as colunas com as questões de pesquisa. Para cada questão buscar nos estudos trechos do artigo que respondem à questão.