

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ADMINISTRAÇÃO
DEPARTAMENTO DE CIÊNCIAS ADMINISTRATIVAS**

FERNANDA BORGES DE OLIVEIRA

**A DIMENSÃO ÉTICA NO USO DA INTELIGÊNCIA ARTIFICIAL:
UMA ANÁLISE APLICADA AO SETOR FINANCEIRO**

Porto Alegre

2023

FERNANDA BORGES DE OLIVEIRA

**A DIMENSÃO ÉTICA NO USO DA INTELIGÊNCIA ARTIFICIAL: UMA
ANÁLISE APLICADA AO SETOR FINANCEIRO**

Trabalho de conclusão de curso de graduação apresentado ao Departamento de *Ciências Administrativas* da Universidade Federal do Rio Grande do Sul como requisito parcial para a obtenção do título de Bacharel em Administração.

Orientadora: Profa. Dra. Raquel Janissek-Muniz

**Porto Alegre
2023**

FERNANDA BORGES DE OLIVEIRA

**A DIMENSÃO ÉTICA NO USO DA INTELIGÊNCIA ARTIFICIAL:
UMA ANÁLISE APLICADA AO SETOR FINANCEIRO**

Trabalho de conclusão de curso de graduação apresentado ao Departamento de Ciências Administrativas da Universidade Federal do Rio Grande do Sul como requisito parcial para a obtenção do título de Bacharel em Administração.

Orientadora: Profa. Dra. Raquel Janissek-Muniz

Aprovado em: ____ de _____ de _____.

BANCA EXAMINADORA

Professora Dra. Daniela Francisco Brauner – Escola de Administração UFRGS

Professora Dra. Natalia Marroni Borges – Escola de Administração UFRGS

Professora Dra. Raquel Janissek-Muniz – Escola de Administração UFRGS

RESUMO

O poder transformador das soluções de Inteligência Artificial (IA) é indiscutível. Seja a nível inovativo, de automatização de processos, de aumento da eficiência ou de redução de custos, os impactos positivos que a implantação da IA oferece são diversos e se aplicam a inúmeros setores. Todavia, as possíveis consequências negativas destes sistemas também merecem atenção. Em setores como o setor financeiro, de alta importância no desenvolvimento socioeconômico global e na vida das pessoas, tanto físicas como jurídicas, estas discussões se tornam ainda mais importantes e urgentes, sendo necessário que a ética seja capaz de acompanhar as mudanças e transformações tecnológicas. Neste contexto, o presente trabalho realiza uma Revisão Sistemática da Literatura a respeito da dimensão ética do uso da Inteligência Artificial no setor financeiro. A busca na base de dados Scopus resultou em 16 publicações que, em diferentes esferas, abordam o tema. Como resultado, a análise dos textos apontou que os principais dilemas éticos residem na falta de transparência no processo de tomada de decisão pelas máquinas e o que isto pode englobar, ou seja, a utilização de dados privados e a propagação de vieses, preconceitos e discriminação.

Palavras-chave: Inteligência Artificial; Ética; Data ethics; Dilemas éticos; Tomada de decisão.

LISTA DE FIGURAS

Figura 1 – Áreas relacionadas com a Inteligência Artificial	24
Figura 2 – Fluxograma da RSL – busca por artigos	42
Figura 3 – Distribuição dos textos selecionados por ano de publicação	44
Figura 4 – Nuvem de palavras das palavras-chave dos artigos selecionados	45
Figura 5 – Desafios Éticos do <i>machine learning</i> em <i>fintechs</i>	53

LISTA DE QUADROS

Quadro 1 – Desafios Éticos no Uso da Inteligência Artificial	31
Quadro 2 – Consolidação dos principais conceitos da revisão teórica	37
Quadro 3 – Filtros da RSL – seleção de artigos	43
Quadro 4 – Publicações selecionadas e analisadas neste trabalho	45
Quadro 5 – Dilemas éticos no uso da inteligência artificial no setor financeiro	53

SUMÁRIO

1 INTRODUÇÃO	15
1.1 OBJETIVOS	19
1.1.1 Objetivo Geral	19
1.1.2 Objetivos Específicos	19
1.2 JUSTIFICATIVA	19
1.3 ESTRUTURA DO TRABALHO	21
2 FUNDAMENTAÇÃO TEÓRICA	21
2.1 INTELIGÊNCIA ARTIFICIAL	22
2.1.1 Inteligência Artificial (IA)	22
2.1.2 Oportunidades e Desafios da IA	26
2.2 DIMENSÃO ÉTICA DA INTELIGÊNCIA ARTIFICIAL	28
2.3 UTILIZAÇÃO DA INTELIGÊNCIA ARTIFICIAL NO SETOR FINANCEIRO	32
3 PROCEDIMENTOS METODOLÓGICOS	38
3.1 ENQUADRAMENTO DA PESQUISA	38
3.2 CONTEXTUALIZAÇÃO DA PESQUISA	39
3.3 OPERACIONALIZAÇÃO DE PESQUISA	40
3.3.1 Coleta de Dados	40
3.3.2. Procedimentos de Análise	42
4 ANÁLISE DOS RESULTADOS E DISCUSSÃO	42
5 CONSIDERAÇÕES FINAIS	55
REFERÊNCIAS	58

1 INTRODUÇÃO

Segundo Schawab (2016), os dados estão transformando a sociedade, sublinhando o que muitos chamam de Quarta Revolução Industrial. A tecnologia está evoluindo de maneira cada vez mais rápida, impactando diretamente a forma como se dão todas as relações sociais, sejam elas pessoais ou profissionais. Valderramas (2020) pontua que a tecnologia alterou de tal forma a vida humana que não se discute atualmente se as mudanças virão, mas sim como adaptar-se a elas.

Neste contexto, o advento da Inteligência Artificial (IA) assumiu um papel de destaque e pode ser considerado um dos grandes vetores das transformações tecnológicas recentes. Mikalef e Gupta (2021, p. 2), definem-a como “a habilidade de um sistema identificar, interpretar, fazer inferências e aprender através de dados, para alcançar determinados objetivos organizacionais e sociais”. Na mesma linha, Gokul (2018) a define como “o uso de ferramentas computacionais na realização de tarefas que, tradicionalmente, requeriam um raciocínio humano sofisticado”. Já John McCarthy (2007) conceitua como a ciência e a engenharia de criar máquinas inteligentes, especialmente programas de computador inteligentes. Para Charniak e McDermott (1985) a IA é um processo que busca humanizar máquinas, e contemplar cada vez mais a sociedade, a partir do estudo da mente humana, com modelos computacionais eficientes e que facilitam as tarefas diárias dos seres humanos.

Os impactos que a Inteligência Artificial vem gerando à sociedade são inegáveis. Esta tecnologia tem sido objeto de expectativa tanto da cultura popular quanto da científica, com potencial para transformar negócios, bem como a relação entre pessoas e tecnologia em geral (ACCENTURE, 2023). Segundo Doneda *et al.* (2018), o desenvolvimento e a implementação de tecnologias de IA proporcionaram efeitos que, muitas vezes, não podem mais ser compreendidos em termos meramente quantitativos, e que implicam uma mudança na subjetividade das relações entre as pessoas e a tecnologia. Contudo, o mesmo autor ainda pontua que assim como qualquer outra tecnologia, mas em maior escala, com maior velocidade e âmbito mais amplo, essas tecnologias apresentam vários desafios e do indivíduo, em especial a sua autonomia, personalidade e igualdade. Garcia (2020), sugere que se a máquina receber dados e informações carregados de vieses e preconceitos de

raça, de gênero, de escolha sexual, de forma física ou de qualquer outro traço, ela irá não só aprender com eles como perpetuá-los, durante o seu processo de aprendizado e logo, ao seu processo de decisão automatizada.

Urge, portanto, a necessidade de voltar a atenção às discussões que envolvem o uso indevido, a generalização, os vieses, o preconceito e a discriminação que as máquinas podem produzir. Vieira (2019) aponta que a inteligência artificial pode até ser inteligente, porém não possui a *expertise* de um humano, isto é, não é sábia. Suas limitações estão baseadas no fato de que quem instrui as máquinas são os humanos e, desta forma, ao “ensinar” as máquinas, os humanos também introduzem os preconceitos, os quais podem também ser exponenciados por meio do *machine learning*, o que amplia a capacidade dos sistemas de IA reproduzir os vieses humanos (VIEIRA, 2019). Ainda segundo Vieira (2019), as pessoas devem tomar iniciativa para corrigir os vieses existentes em algoritmos, antes que seja tarde demais. Britz (2007) cunhou a expressão “injustiça pela generalização” que faz referência justamente aos resultados discriminatórios gerados pela generalização, prática muito utilizada nas decisões automatizadas.

Nestas circunstâncias, diversos autores discutem sobre as questões éticas por trás do uso dos sistemas de IA. Segundo Doneda *et al.* (2018), a ética, enquanto ramo do conhecimento que lida com a moralidade e, no sentido mais amplo, com a preocupação que os humanos possuem em descobrir a melhor maneira de viver, e de distinguir entre o bem e o mal, o certo e o errado, está evoluindo para tentar acompanhar as transformações digitais. Neste sentido, a *data ethics* foi definida como um novo ramo da ética, que estuda e avalia problemas morais relacionados aos dados (FLORIDI). Este ramo específico da ética marca uma mudança de foco sobre o conteúdo (informação) que pode ser criado, registrado, processado e compartilhado através de meios tecnológicos, para um foco nos dados (FLORIDI, 2016).

Ademais, recentemente um artigo publicado pela Gartner (2022), que explora quais os papéis essenciais para análise de dados, pontuou que uma das posições de trabalho extremamente importantes para o futuro das organizações é o *data ethicist* (eticista de dados).

O eticista de dados pensa nas consequências não intencionais do uso dos dados e determina risco e oportunidades. Quais valores podem ser gerados a partir das novas maneiras de usar os dados, e isto está alinhado com os valores da organização? Dado que nem todas as consequências não intencionais do uso de dados podem ser previstas, o eticista de dados tem o papel de monitorar consequências imprevistas que podem levar a percepções desproporcionais sobre a vida de pessoas. Por fim, o eticista de dados é responsável por tornar todos os stakeholders (partes interessadas) conscientes eticamente (HEIZENBERG; JUDAH; DUNCAN, 2022).

Dado este cenário, uma ramificação dentro do campo de estudo da ética que merece destaque são os dilemas éticos (paradoxo ético ou dilema moral), que se configuram por um problema no processo de tomada de decisão entre duas opções possíveis, onde nenhuma das quais é absolutamente aceitável do ponto de vista ético (BRITES, 2020). Preocupações com dilemas éticos no contexto da Inteligência Artificial constituem um aspecto amplamente debatido atualmente e as principais estão centradas nos riscos políticos e sociais decorrentes da sua utilização maliciosa, ou seja, quando utilizados para a promoção da manipulação e distorção de processos sociais (MOURELLE, 2019; CARREÑO, 2019; SHAW, 2019). Neste contexto, deve ser conferida importante atenção à capacidade desses sistemas de tomarem tais decisões onde há vários caminhos possíveis, e cujas consequências afetam diretamente a vida e a dignidade dos seres humanos, isto é, a capacidade destes sistemas enfrentarem os dilemas éticos.

Apesar das questões éticas envolvidas, inúmeros setores já implantaram e consolidaram a utilização de soluções de IA em suas operações. Roger Spitz, presidente da Techistential e Chairman da Disruptive Futures Institute, disserta que o ritmo acelerado da IA está transformando radicalmente todos os setores, incluindo varejo, finanças, manufatura, saúde e software - e desafiando a sobrevivência dos modelos tradicionais nesses setores (DISTRITO REPORT, 2021). Ainda para ele, os objetivos comuns para a implementação de IA incluem melhorar os processos de negócios, derivar melhores *insights*, construir produtos mais inteligentes e combater ameaças competitivas (DISTRITO REPORT, 2021).

Um estudo realizado pela Deloitte AI Institute (2021) aponta como vêm sendo usadas as soluções de Inteligência Artificial nos mais diversos setores, como varejo, energia e indústria, governo e serviços públicos, saúde e serviços financeiros. No que tange à indústria do varejo, o estudo aponta que os sistemas de inteligência

artificial estão sendo utilizados para personalizar a experiência do consumidor, auxiliar diretamente na decisão a respeito do mix de produtos e serviços a serem oferecidos, na previsão de demanda e, logo, no planejamento de marketing. Já em relação ao setor de energia e indústrias, soluções de AI são utilizadas visando a melhora no desempenho e performance das máquinas, bem como na previsão de falhas. A respeito do setor governamental, a inteligência artificial está sendo empregada cada vez mais como meio de automatização de processos anteriormente morosos e lentos. Neste contexto, merece o devido destaque a utilização de sistemas automatizados para auxílio em processos jurídicos em inúmeros países, incluindo Brasil e Estados Unidos. No que concerne aos setores da saúde, durante a Pandemia de Covid 19 houve um aumento exponencial da utilização da IA neste setor (GARCIA, 2020). Na época, sistemas foram utilizados para mapeamento de casos e monitoramento do fluxo de pessoas nas cidades, e hoje seguem sendo usados com muita força no setor para auxílio em diagnósticos, gestão de hospitais e medicamentos e na produção de remédios (DELOITTE, 2021; GARCIA, 2020).

Por fim, o estudo da Deloitte (2021) aborda a utilização de sistemas de AI no setor financeiro. Para Donepudi (2017), é comprovado que a Inteligência Artificial e o *Machine Learning* (ML) são muito úteis para o setor bancário, seja pela automatização de processos ou pela rápida resposta aos consumidores. De forma detalhada, a IA e o ML são impactantes no setor financeiro no que tange a prevenção à fraude, o contato dinâmico com clientes através de *chatbots*, a personalização de atendimento aos clientes, a automatização de processos, como a análise de crédito e a análise de ativos financeiros (DELOITTE, 2021).

Sendo este um setor que pode gerar impactos significativos na vida das pessoas e na sociedade como um todo, urge a necessidade de avaliar os dilemas éticos por trás da tomada de decisão das máquinas de IA no âmbito financeiro. Neste sentido, o presente trabalho buscará responder a seguinte questão: quais são os dilemas éticos por trás da utilização de soluções de Inteligência Artificial no setor financeiro?

1.1 OBJETIVOS

1.1.1 Objetivo Geral

Analisar possíveis dilemas éticos presentes na utilização de soluções de Inteligência Artificial (IA) em empresas do Setor Financeiro.

1.1.2 Objetivos Específicos

- Listar as principais soluções de Inteligência Artificial (IA) utilizadas no setor financeiro;
- Mapear os dilemas éticos presentes na utilização de tais soluções;
- Explorar os principais pontos de atenção nos casos de utilização.

1.2 JUSTIFICATIVA

As tecnologias de IA estão sendo rapidamente globalmente adotadas em todos os setores para inovar os modelos operacionais e de negócios - de serviços, de saúde a imóveis, de finanças ao varejo (JANISSEK-MUNIZ; BORGES, 2022). No entanto, segundo Janissek-Muniz e Vassali (2022), enquanto a maior parte da atenção tem sido focada na perspectiva positiva do uso da Inteligência Artificial, a preocupação com potenciais aspectos negativos é crescente. Para os autores, essa visão, de potenciais aspectos negativos ou não intencionais da IA, tem provocado atenção de entidades públicas, privadas e da comunidade científica, tendo em vista seu significativo impacto potencial. Janissek-Muniz e Borges (2022) também apontam que alguns dos enfoques atualmente discutidos de forma mais ampla dentro deste tema são as questões éticas, que envolvem os vieses e possíveis efeitos de preconceito e discriminação (RUSSEL *et al.*, 2015). Cheng *et al.* (2022) reiteram que, considerando a importância, universalidade e as potenciais consequências trazidas pela IA investigar esse fenômeno e possíveis impactos também negativos merece maior exploração. Os mesmos autores ainda sublinham a

importância de que a devida atenção ao assunto seja dada principalmente na área acadêmica.

Contudo, para além do âmbito acadêmico, outros indícios evidenciam a importância do estudo da dimensão ética da Inteligência Artificial. No dia 28 de fevereiro de 2020, em Roma, foi assinado pela Pontifícia Academia da Vida (*Pontifical Academy for Life*), pela Microsoft, IBM, FAO e pelo Ministério Italiano da Inovação, um documento chamado “*Call for an AI Ethics*”, isto é, “*Chamada por uma IA ética*”. Tal documento tem como objetivo principal gerar mobilização internacional e promover um senso de responsabilização coletivo entre governos, instituições e o setor privado em prol da criação de um futuro onde toda inovação digital será acompanhada de forma a respeitar a dignidade humana. Em outubro de 2022, a Universidade de Notre Dame (Indiana, USA) sediou o primeiro “Rome Call for AI Ethics”, encontro global voltado para tratar sobre a temática (THE CALL, 2023).

A assinatura do documento e o encontro mundial sediado na Universidade de Notre Dame reforçam a tamanha importância que é tratar da dimensão ética na utilização da Inteligência Artificial (IA). Santos (2018) ressalta que a IA merece discussões cada vez mais abrangentes em termos de sociedade e que envolvam, sobretudo, as redes sociais virtuais, dadas as fragilidades legislativas nesse campo que abrem margem à violação dos aspectos não apenas éticos e legais, bem como humanos. Outro ponto que reforça a importância da tratativa foi o surgimento de um novo ramo específico da ética, o chamado *data ethics*, que permite uma melhor análise de temas éticos, como privacidade, anonimato, transparência, confiança e responsabilidade, concentrando-se nos seus aspectos procedimentais (como coleta, curadoria, filtragem, criação de dados) e algorítmicos (aqueles que alimentam IA e ML), que exercem um impacto significativo na ação e comportamento humanos (FLORIDI, 2016).

No Brasil, em 06 de dezembro de 2022, foi apresentado ao presidente do Senado um relatório que se constitui como uma proposta de regulamentação, e que tem como objetivo estabelecer princípios, regras, diretrizes e fundamentos para o desenvolvimento e a aplicação da inteligência artificial no Brasil (AGÊNCIA SENADO, 2022). Na exposição, a importância da proposta de lei é justificada dada a complexidade técnica, moral e jurídica da ampliação da tecnologia de IA como

instrumentos auxiliares na tomada de decisões pelos poderes públicos e agentes privados.

No que se refere à escolha do setor financeiro como campo de estudo, a decisão se justifica pela importância do setor na vida das pessoas e no desenvolvimento socioeconômico. Segundo El Qadi *et al.* (2023), as instituições financeiras permitem a economia global crescer. Logo, as decisões automatizadas neste setor possuem impacto potencial significativo tanto de pessoas físicas como jurídicas e, por isto, devem ser tomadas de maneira ética, justa e transparente. Dados tais fatos, a emergência do estudo deste tema fica evidenciada. Em um cenário de “Quarta Revolução Industrial” urge a necessidade de garantir que a ética irá caminhar ao lado das transformações digitais.

1.3 ESTRUTURA DO TRABALHO

O presente trabalho está dividido em cinco seções. Esta primeira é composta pela introdução, englobando o problema de pesquisa, objetivos geral e específico, e a justificativa. A segunda seção consiste no referencial teórico, onde serão apresentados os fundamentos para a posterior análise. Serão abordados os conceitos de inteligência artificial, tomada de decisão automatizada, o conceito de ética e *data ethics*, e os principais casos de utilização da IA no mercado financeiro. Na terceira seção serão apresentados os procedimentos metodológicos e na quarta os resultados encontrados. Por fim, a quinta seção contará com as considerações finais e sugestões para futuros trabalhos.

2 FUNDAMENTAÇÃO TEÓRICA

A fundamentação teórica deste trabalho tem como objetivo discorrer sobre a Inteligência Artificial e sua relação com a ética. Para isto, no item 2.1 será discutido o conceito de Inteligência Artificial, suas capacidades, oportunidades e desafios e, posteriormente, no item 2.2 será analisada a dimensão ética no uso desta tecnologia. Por fim, no item 2.3 os conceitos serão aplicados ao campo de estudo específico do trabalho, isto é, o setor financeiro.

2.1 INTELIGÊNCIA ARTIFICIAL

2.1.1 Inteligência Artificial (IA)

Segundo Sichman (2021), sempre que ocorre um entusiasmo com os resultados de uma tecnologia, existe uma tendência da mídia em fornecer definições e explicações, por vezes não muito precisas, dos seus principais aspectos, como é o caso da IA na atualidade. Ao fazer uma retomada da origem histórica da Inteligência Artificial, Sichman (2021) pontua que o marco inicial da IA foi o verão de 1956. Mais precisamente, no *Dartmouth College Conference*, que contou com a participação de pesquisadores reconhecidos como os pais da área, tal qual John MacCarthy, Marvin Minsky, Alan Newell e Herbert Simon. Posteriormente, estes pesquisadores tiveram trajetórias científicas marcantes neste domínio da computação (IA).

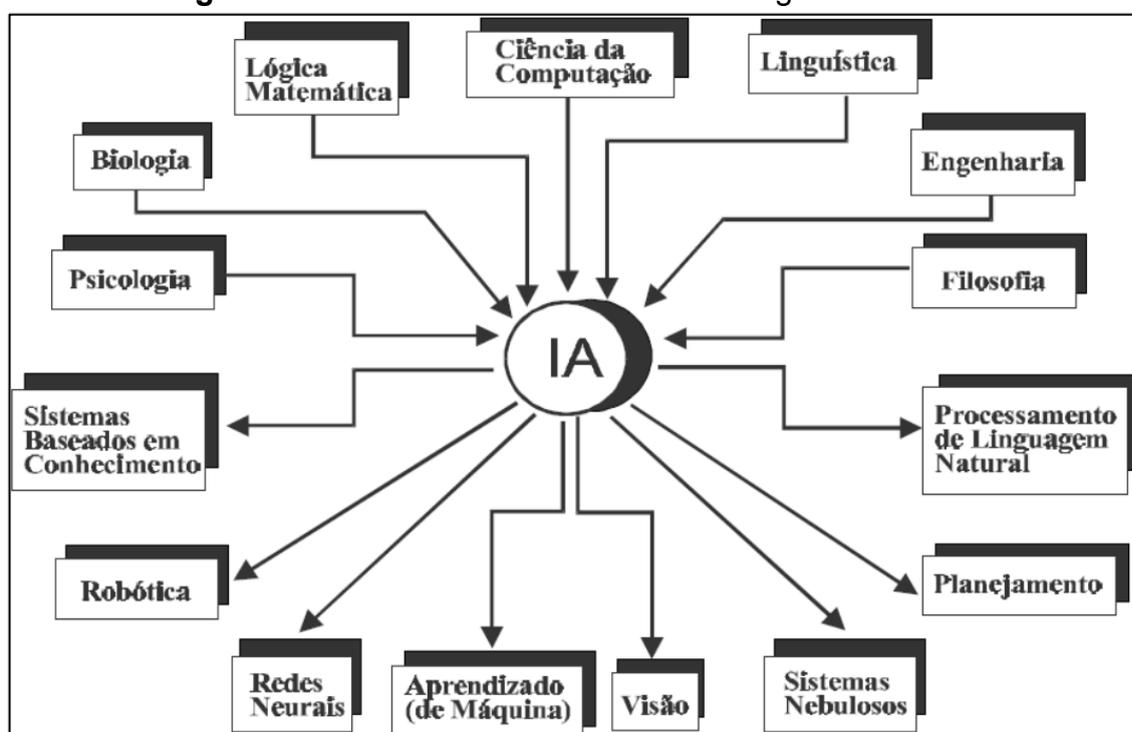
Para John McCarthy (2007), a IA mesma pode ser definida como a ciência e a engenharia de desenvolver máquinas inteligentes, especificamente programas computacionais inteligentes. Segundo ele, está relacionada à tarefa de usar computadores para entender e mimetizar a inteligência humana, porém não precisa se limitar a métodos que são biologicamente observáveis. Sichman (2021) aponta que se trata de um ramo da ciência/engenharia da computação que visa desenvolver sistemas computacionais que solucionam problemas.

Mikalef e Grupta (2021, p. 2) apresentam uma abordagem integrativa para o conceito de IA, definindo-a como “a habilidade de um sistema identificar, interpretar, fazer inferências e aprender através de dados, para alcançar determinados objetivos organizacionais e sociais”. Na mesma linha, Gokul (2018, p. 6) a define como “o uso de ferramentas computacionais na realização de tarefas que, tradicionalmente, requeriam um raciocínio humano sofisticado”. Allas *et al.* (2018) também contribui para a definição de IA, descrevendo-a como a habilidade de máquinas exibirem inteligência humana, resolvendo problemas sem a necessidade de um código que descreva todas as etapas da resolução.

Cheng *et al.* (2022) apresentam um conceito mais amplo, caracterizando-a como uma inovação extremamente disruptiva que atraiu considerável atenção de

profissionais e acadêmicos e que oferece oportunidades sem precedentes para mudanças fundamentais e atualizações abrangentes em muitos setores. Para Santos Gomes (2010) a Inteligência Artificial é um ramo da Ciência da Computação cujo interesse é fazer com que os computadores pensem ou se comportem de forma inteligente, mimetizando a inteligência humana. Segundo ele, por ser um tópico muito amplo, a IA também está relacionada com psicologia, biologia, lógica matemática, linguística, engenharia, filosofia, entre outras áreas científicas, conforme está ilustrado na Figura 1.

Figura 1 – Áreas relacionadas com a Inteligência Artificial



Fonte: MONARD; BARANAUKAS, 2000, p. 2

A Inteligência Artificial sistematiza e automatiza tarefas intelectuais e, portanto, é potencialmente relevante para qualquer esfera da atividade intelectual humana. Neste sentido, podemos dizer que ela é um campo universal (RUSSELL; NORVIG, 2010).

Ainda que se julgue importante e relevante explorar as definições e conceitos de Inteligência Artificial, Sichman (2021) aponta que ao invés de tentar fornecer uma definição acadêmica de IA, o mais adequado seria tentar caracterizar quais são os objetivos da área. Segundo ele, uma das primeiras tentativas desta abordagem foi

proposta Rich e Knight (199). Para os autores, o objetivo da IA é desenvolver sistemas para realizar tarefas que, no momento: (i) são mais bem realizadas por seres humanos que por máquinas, ou (ii) não possuem solução algorítmica viável pela computação convencional (SICHMAN, 2021). Em suma, o domínio de IA se caracteriza por ser uma coleção de modelos, técnicas e tecnologias (busca, raciocínio e representação de conhecimento, mecanismos de decisão, percepção, planejamento, processamento de linguagem natural, tratamento de incertezas, aprendizado de máquina) que, isoladamente ou agrupadas, resolvem problemas anteriormente processados apenas pela inteligência humana (SICHMAN, 2021).

Em relação à dimensão técnica da IA, pode-se dividi-la em quatro capacidades distintas, sendo elas: *Machine Learning*, *Deep Learning*, Processamento Natural de Linguagem e Visão Computacional. O *Machine Learning* - traduzido para o português como o aprendizado de máquinas - são técnicas orientadas que buscam permitir que os sistemas aprendam automaticamente a partir de um grande volume de dados (LUDERMIR, 2021). A mesma autora (2021) esclarece a respeito dos três principais tipos de Aprendizado de Máquina, sendo eles: Supervisionado, Não Supervisionado e por Reforço.

Ludermir disserta que o Aprendizado Supervisionado trata de um modelo onde para cada exemplo apresentado ao algoritmo de aprendizado é necessário apresentar a resposta desejada (ou seja, um rótulo informando a que classe o exemplo pertence, no caso de um problema de classificação de imagens, por exemplo, como distinguir imagens de gatos e de cachorros). Neste caso, o objetivo do algoritmo é construir um classificador que possa determinar corretamente a classe de novos exemplos ainda não rotulados. Já no Aprendizado Não Supervisionado o algoritmo agrupa os exemplos pelas similaridades dos seus atributos, ou seja, algoritmo analisa os exemplos fornecidos e tenta determinar se alguns deles podem ser agrupados de alguma maneira, formando agrupamentos ou *clusters*. Por fim, Ludermir (2021) explica que no Aprendizado por Reforço o algoritmo não recebe a resposta correta, mas recebe um sinal de reforço, de recompensa ou punição. O algoritmo faz uma hipótese baseado nos exemplos e determina se essa hipótese foi boa ou ruim. Ainda em relação ao Aprendizado de Máquina, vale conferir uma atenção especial a técnica chamada de Redes Neurais

Artificiais (RNA) e que tem tido sucesso em resolver muitos problemas (LUDERMIR, 2021). As RNA são modelos matemáticos que se inspiram nas estruturas neurais biológicas e que têm a capacidade computacional adquirida por meio de aprendizado. O processamento da informação em RNA é feito nos neurônios artificiais, conhecidos como neurônio McCulloch e Pitts (MCCULLOCH; PITTS, 1943).

A segunda capacidade da IA é o *Deep Learning* - traduzido para o português como aprendizado profundo . Trata-se de um método específico de aprendizado de máquina o qual é baseado em um grupo de algoritmos que buscam moldar abstrações de alto nível de dados utilizando-se de um grafo profundo com diversas camadas de processamento, compostas por diversas alterações lineares e não lineares (CHAGAS, 2019). O mesmo autor pontua que ao invés de organizar as informações para atuarem através de equações predeterminadas, esse aprendizado determina padrões básicos dessas informações e ensina os computadores a desenvolver-se através da identificação dos padrões em camadas de processamento. Já em relação ao Processamento Natural de Linguagem, trata-se do campo da Inteligência Artificial que lida com a tradução e manipulação de textos, tal qual os tradutores automáticos (LUDERMIR, 2021). A Visão Computacional é a área da IA que se dedica a se dedica a “ensinar” a máquina a compreender e interpretar imagens JANISSEK-MUNIZ; BORGES, 2022).

Dado o exponencial desenvolvimento das capacidades acima, são inúmeras as oportunidades, bem como os desafios, que os sistemas de Inteligência Artificial irão encarar. Segundo Janissek-Muniz e Vassali (2022), é inegável o potencial transformador da Inteligência Artificial, que impacta desde a capacidade inovativa e tecnológica das organizações até seu desempenho em termos de processos e economia. Ludermir (2021) pontua o uso da IA vem trazendo muitos benefícios para a sociedade, tais como: melhorias nos serviços de saúde, voz para texto, melhorias na educação, energia limpa e barata, detecção de fraudes, meios de transportes mais seguros (aplicativos de transporte), rápidos (rota otimizadas) e limpos. Contudo, Cheng *et al.* (2022) reiteram que, considerando a importância, universalidade e os potenciais impactos trazidos pela IA, investigar esse fenômeno e possíveis impactos também negativos merece maior exploração. Estes impactos

negativos são vistos como os desafios a serem enfrentados pela Inteligência Artificial. Juntamente com as oportunidades da IA, tais desafios serão abordados com maior detalhamento na seção a seguir.

2.1.2 Oportunidades e Desafios da IA

Segundo Sichman (2021) estamos atravessando um período onde os futuros benefícios que a IA pode prover são vistos com muito otimismo e euforia. Segundo ele, tal otimismo se justifica por uma conjunção de três fatores fundamentais: (i) o custo de processamento e de memória nunca foi tão barato; (ii) o surgimento de novos paradigmas, como as redes neurais profundas, possibilitados pelo primeiro fator e produzindo inegáveis avanços científicos; e (iii) uma quantidade de dados gigantesca disponível na internet em razão do grande uso de recursos tais como redes e mídias sociais - o que se configura como uma grande oportunidade dado que, quais mais dados processar, mais soluções a IA poderá oferecer à sociedade. Tal entusiasmo, entretanto, tem sido acompanhado por uma série de temores, alguns dos quais fundados (SICHMAN, 2021).

Segundo Cheng *et al.* (2022), considerando a importância, universalidade e as potenciais consequências trazidas pela IA, a investigação dos possíveis impactos negativos merece maior exploração. Na mesma linha, Janissek-Muniz e Vassali (2022) pontuam que enquanto maior parte da atenção tem focado na perspectiva positiva do desenvolvimento e uso dessa tecnologia, existe uma preocupação crescente com potenciais aspectos negativos. Segundo os autores, a Inteligência Artificial também pode apresentar vieses menos positivos devido ao seu envolvimento - direto ou indireto - no aumento exponencial da complexidade subjacente, gerando níveis mais elevados de incerteza, desigualdade, instrumentalização, reificação, alienação, psicopatologias, etc. (CALVO, 2020B, 2021; GONZÁLEZ-ESTEBAN; CALVO, 2022; PRUNKL *et al.*, 2021). Ainda para eles, esta visão de potenciais aspectos negativos ou não intencionais da IA recentemente têm provocado atenção de entidades públicas, privadas, e da comunidade científica, dado seu significativo impacto potencial.

Para Ludermir (2021), um exemplo um destes potenciais aspectos negativos são os desafios relacionados às questões sociais. Para ela, a perda de vagas de trabalho é um dos impactos danosos trazidos pela IA na medida que pode representar um intensificador de desigualdades sociais, dado que as vagas de emprego que exigem menor qualificação são as mais afetadas. No entanto, para autores como Valderramas (2020), Doneda *et al.* (2018), Andrade Martins (2021), Garcia (2020) e Cheg *et al.* (2021) o grande desafio a ser enfrentado pela Inteligência Artificial está relacionado à dimensão ética por trás da utilização dos sistemas. Por este motivo, a próxima seção será dedicada exclusivamente à análise da dimensão ética no uso de sistemas de IA.

2.2 DIMENSÃO ÉTICA DA INTELIGÊNCIA ARTIFICIAL

Aristóteles, considerado o fundador da ética como uma disciplina da Filosofia, conceitua a ética como o conjunto de valores morais e princípios que orientam a conduta humana na sociedade. Para Doneda *et al.* (2018), a ética é o ramo do conhecimento que lida com a moralidade e, no sentido mais amplo, com a preocupação que os humanos sempre tiveram em descobrir a melhor maneira de viver, e de distinguir entre o bem e o mal, o certo e o errado. Já segundo Valderramas (2020), a ética estuda os valores morais e princípios do comportamento humano, sendo a ciência que discute como o homem se relaciona com a sociedade.

A ética não está em nenhum campo de atuação, pois trata-se de um exercício inerente de todas as profissões, portanto ela não pode e não deve em hipótese nenhuma ser desprezada, tampouco diante da enormidade de números, palavras, imagens e outros símbolos que remetem à intimidade das pessoas ou instituições e que estão presentes em proporções nunca antes vista na história da humanidade. Partindo-se da ideia de que sendo a ética um dos princípios norteadores do caráter do ser humano, as máquinas também deveriam ter o mesmo norte como base essencial de coleta, tratamento, classificação, distribuição e divulgação da informação com a singular diferença que este princípio estaria codificado por meio de símbolos (VALDERRAMAS, 2020, p. 8)

Valderramas (2020) sublinha a importância das máquinas também serem norteadas por condutas éticas. Neste contexto, Ludermir (2021) contribui pontuando que precisam ser avaliadas e resolvidas ao menos parcialmente em um curto espaço de tempo as diretrizes da dimensão ética no uso da IA. Logo, a relevância

que tal tema vem ganhando culminou no surgimento do *data ethics*. Este foi definido como um novo ramo da ética, que estuda e avalia problemas morais relacionados aos dados e o seu surgimento marcou uma mudança de foco sobre o conteúdo (informação) que pode ser criado, registrado, processado e compartilhado através de meios tecnológicos para um foco nos dados (FLORIDI, 2016).

Analisando cada uma das suas vertentes, a ética dos dados se concentra em problemas éticos colocados pela recolha e análise de grandes conjuntos de dados, abordando questões relacionadas à criação de perfis, publicidade, re-identificação de indivíduos, privacidade em grupo, discriminação e transparência, entre outras; já a ética dos algoritmos enfoca questões levantadas pela crescente complexidade e autonomia de algoritmos (incluindo inteligência artificial e agentes artificiais, por exemplo, como “bots” de internet e outras aplicações de *machine learning*), ou seja, temas relativos à responsabilidade moral e à responsabilidade de projetistas e cientistas de dados; à análise ética e à auditoria de algoritmos; à avaliação de possíveis resultados indesejáveis resultantes do desenvolvimento de produtos e pesquisas conduzidas por essas tecnologias. Por sua vez, a “ética das práticas” (incluindo ética profissional e deontologia) aborda as questões urgentes relativas às responsabilidades das pessoas e das organizações encarregadas de processos, estratégias e políticas de dados, nomeadamente em relação a questões de consentimento, privacidade dos utilizadores e uso secundário (FLORIDI, 2016)

Sinteticamente, o *data ethics* permite uma melhor análise de temas éticos, como privacidade, anonimato, transparência, confiança e responsabilidade, concentrando-se nos seus aspetos procedimentais (como coleta, curadoria, filtragem, criação de dados) e algorítmicos (aqueles que alimentam IA e ML), que não se traduzem, direta ou necessariamente, em informação, mas ainda assim exercem um impacto significativo na ação e comportamento humanos (FLORIDI, 2016).

Uma evidência da importância do tema se dá no fato do *Institute of Electrical and Electronic Engineers* (IEEE), da *Association for Computing Machinery* (ACM), a Comunidade Comum Europeia, entre outros, possuir definidas diretrizes para um uso responsável da IA (HAGENDORFF, 2020). No Brasil temos projetos de lei na Câmara dos Deputados (Projetos de Lei n.21/20 e 240/20), que remetem ao cumprimento da Lei Geral de Proteção de Dados (LGPD) e outros no Senado Federal (Projetos de Lei n.5051/2019 e 5691/2019) que se propõe a abordar o mesmo tema (LUDERMIR, 2021). Tal relevância existe porque as tecnologias em discussão no presente trabalho não só afetam - inibindo ou induzindo - comportamentos individuais e sociais, como também têm o potencial de alterar

intrinsecamente os próprios indivíduos e a sociedade. Em sua fase de criação, esses sistemas são desenvolvidos com valores específicos incorporados e, quando implementados e adotados, carregam esses valores, moldando e mudando indivíduos, comunidades e sociedades.

Vieira (2019) pontua que vivemos em um mundo digital, onde o homem dialoga com máquinas fazendo com que os algoritmos passem a tomar decisões e determinar avaliações e inclusive ações que outrora eram feitas por humanos. Contudo, sabe-se que os algoritmos são falíveis e limitados, pois são opiniões embutidas em um código, por meio do qual o homem ensina a máquina, ou seja, a máquina poderá tomar decisões enviesadas com base nos dados fornecidos. Logo, torna-se necessário nutrir os dados de forma precisa para que a inteligência artificial não cometa erros e não seja discriminatória pois tudo o que as máquinas sabem foi por que o homem o ensinou e em consequência, também ensinou os preconceitos ou aprenderam por meio do *machine learning*, tendo a capacidade de reproduzir os vieses humanos (VIEIRA, 2019).

Neste contexto, Doneda *et al.* (2018) dedicam parte de seu estudo para discorrer de maneira detalhada sobre os desafios éticos a serem enfrentados pela IA, sintetizados a partir do quadro 1:

Quadro 1 – Desafios Éticos no Uso da Inteligência Artificial

Desafio Ético	Descrição
Redução do controle humano	A IA facilita a delegação de tarefas importantes em sistemas autônomos que - em circunstâncias específicas - devem permanecer, pelo menos em parte, sujeitos à supervisão humana, seja no “loop”, para fins de monitorização, seja no “pós-loop”, para corrigir erros ou danos que eventualmente surjam.
Remoção da responsabilidade humana	Os desenvolvimentos no campo da IA também podem aumentar a tendência para os desresponsabilizar sempre que um sistema de inteligência artificial puder ser responsabilizado por uma falha ou má conduta. Dada a complexa “cadeia” de designers, fornecedores e usos automatizados que essas tecnologias envolvem sem intervenção humana, essa tendência poderá tornar mais difícil a responsabilização das pessoas por falhas específicas da IA.
Desvalorização de competências humanas	Os desenvolvimentos tecnológicos da IA também podem desvalorizar competências humanas, nomeadamente em domínios sensíveis e intensivos, como o diagnóstico médico e a aviação. Se, por exemplo, daqui a algumas décadas, houver uma escassez de especialistas humanos capazes de diagnosticar o câncer, a sociedade estaria mal preparada e equipada para o mau funcionamento da IA e para lidar com eventuais ataques malévolos.

Erosão da autodeterminação humana	A IA pode corroer a autodeterminação pessoal, pois pode induzir a mudanças não planejadas e indesejadas nos comportamentos humanos para acomodar rotinas que facilitam a automação e a vida das pessoas. O poder de previsão da IA e o denominado “nudging”, mesmo que não intencional, devem fomentar, e nunca minar a dignidade humana e a autodeterminação (YANG, 2018).
Facilitação de condutas humanas controversas ou mesmo malévolas	Há também o potencial risco para que essa tecnologia caia nas mãos erradas e para usos mal-intencionados da IA, o que representaria grave ameaça à segurança e prosperidade de todos nós. Uma forma de impedir o uso malévolo da IA é adotar a consideração e o entendimento de que devemos tratar as pessoas como fins em si mesmas, e nunca apenas como um simples meio.
Preconceito e (in)justiça	Como referido no início, as tecnologias podem ser desenvolvidas com valores específicos embutidos nelas. Isto é particularmente verdadeiro no caso da IA, na qual os algoritmos são treinados de maneira a informar ou apoiar decisões que refletem (ou deixam de refletir) tantos os valores como os seus preconceitos humanos. Os sistemas de IA - sem salvaguardas adequadas - podem herdar e propagar preconceitos existentes nos dados em que foram treinados. É o caso de modelos treinados em dados em que as minorias são sub-representadas e, portanto, não aparecem bem refletidas, o que levanta questões de potencial preconceito e discriminação, nomeadamente em decisões automatizadas em matéria de emprego, habitação, crédito e justiça criminal.

Fonte: Doneda *et al.* (2018)

Na mesma lógica, mas partindo para uma abordagem de solução destes desafios, Virginia Dignum (2019) apresenta como deve-se desenvolver e utilizar a IA de modo responsável. A autora advoga que uma postura ética deve ser adotada em três instâncias distintas (SICHMAN, 2021), sendo elas:

- no processo de *projeto* de tais sistemas, garantindo que as equipes tenham em mente e antevejam as possíveis consequências do sistema para os indivíduos e sociedades;
- no projeto do *comportamento* de tais sistemas, visando representar de forma adequada capacidades de raciocínio ético nos agentes inteligentes;
- no *código de conduta* dos projetistas e desenvolvedores, mediante uma regulação adequada e processos de certificação que garantam um comportamento adequado dos atores envolvidos, como já existe em outras profissões.

Neste contexto, ainda que os dilemas éticos se manifestem no uso de sistemas de Inteligência Artificial em diferentes setores, o presente trabalho abordará especificamente os casos de uso e os dilemas éticos no setor financeiro. Este, por sua vez, é um setor que possui alto impacto na vida das pessoas e na sociedade como um todo, dado que seu funcionamento satisfatório permite o

desenvolvimento da economia global (EL QADI *et al.*, 2023). Por isso, os impactos das decisões tomadas através de máquinas merecem maior atenção.

2.3 UTILIZAÇÃO DA INTELIGÊNCIA ARTIFICIAL NO SETOR FINANCEIRO

Segundo Garcia (2021, p. 1), “no panorama empresarial atual, existe uma pressão cada vez maior para a reinvenção de modelos de negócio e para a incorporação de dados/informação na tomada de decisão no que diz respeito à estratégia das empresas”. O mesmo autor ainda pontua que algumas indústrias, a exemplo da fotografia, tiveram o maior impacto pela transição digital no final do século XX e início do século XXI e que outros setores estão passando pela transição para a era de informação com maior intensidade agora. Um destes setores é o financeiro, um setor relativamente retardatário e que atualmente está sofrendo mudanças estruturais com a entrada na era da IA (SHAIKH, 2017). Donepudi (2017) é preciso quando pontua que o uso da Inteligência Artificial é crucial para a sobrevivência dos bancos nos tempos modernos.

De acordo com a Accenture (2019), a nível global 41% das interações bancárias ocorrem a partir de dispositivos móveis, o que justifica a necessidade deste setor estar à frente em termos inovativos (SILVA, 2020). O relatório da Accenture (2019) ainda pontua que o digital não é mais uma diferenciação e que para que os bancos tenham destaque os gestores devem olhar para cinco tendências, sendo que a primeira delas diz respeito a quatro aspectos tecnológicos: “*Distributed Ledgers*, Inteligência Artificial, Realidade Aumentada e Computação Quântica”. Ao tratar sobre esta tendência e sobre quais destas tecnologias terão maior impacto no setor bancário nos próximos três anos, os executivos bancários apontaram como líder a Inteligência Artificial. 90% dos participantes ainda alegaram que já estão utilizando alguma das quatro tecnologias referenciadas pelo relatório, sendo o destaque a Inteligência Artificial, com 43% de adoção (SILVA, 2020). Outro relatório, apresentado pelo *Deutsche Bank Research* (2019), explana que no setor bancário, os dados são essenciais para quase todas as linhas de negócio e que a gestão autónoma destes sem envolvimento humano promove ganhos a nível de agilidade, precisão e eficiência (SILVA, 2020).

Em relação à usabilidade, a Inteligência Artificial pode ser usada no setor bancário para diferentes propostas (DONEPUDI, 2017), aumentando a eficiência do negócio, com baixo custo, de modo a melhorar a entrega do serviço, através, por exemplo, do serviço de mensagens automáticas 24/7, e a experiência do cliente (PRENTICE; NGUYEN, 2020). Ademais, a IA possui a capacidade de tornar os bancos mais seguros, através das ferramentas de detecção de fraude, por exemplo (SHANTHI; PAVITHRA, 2020). Com o aumento do uso de pagamentos online, o número de criminosos cibernéticos também aumentou para fazer fraudes online, mas os algoritmos de IA são capazes de monitorar as transações e a transação é bloqueada caso o algoritmo detectar riscos (KAYA, 2019).

Em conformidade com o relatório do *Deutsche Bank Research* (2019) as soluções de Inteligência Artificial com foco no cliente e nas operações tendem a passar por explorações mais intensivas do que outras no setor financeiro. No relatório são citadas e destacadas os seguintes usos: identificação e prevenção de fraudes e crimes cibernéticos, processos para verificar a identidade do cliente (KYC – *Know Your Customer*), *chatbots*, análise dos dados de documentos legais e relatórios anuais para extrair informações relevantes e por fim os *robo-advisors* orientados e programados para oferecer melhores soluções para a gestão financeira dos clientes. Segundo Garcia (2021), as instituições financeiras conseguem retirar quatro grandes benefícios da IA, que podem ser também visualizadas em uma ótica de front, middle e back-office, sendo elas: capacidade de previsão, precisão, eficiência e escalabilidade (HALPIN; DANNEMILLER, 2019; PWC FINANCIAL SERVICES, 2018).

A capacidade de previsão diz respeito ao papel que os algoritmos podem ter no auxílio aos gestores de portfolio através da análise de grandes conjuntos de dados, permitindo testar investimentos com maior robustez. Para além disso, os sistemas construídos com base em IA são menos propensos a cometer erros humanos e são mais eficientes que os humanos a analisar dados. Assim, os investidores podem focar-se nas áreas onde são mais produtivos (definir estratégias, atualizar os algoritmos). Por fim, em relação à escalabilidade, a IA pode não funcionar em todo o tipo de negócios ou em todas as funções, mas é uma fonte de vantagens de custo no processamento de dados. Empresas que utilizem algoritmos de IA podem, muitas vezes, atingir a escalabilidade dos mesmos (GARCIA, 2021, p. 14)

No que tange às ferramentas de IA usadas no setor financeiro, Garcia (2021), propõe uma divisão baseada em suas aplicações, sendo elas: aplicações centradas no cliente, gestão de investimentos e do *trading*, aplicações *back-office*, e risco e *compliance*. Em relação à primeira divisão, voltada para o atendimento ao cliente, o autor pontua que a área é responsável pelas tecnologias com vista a melhorar a experiência dos clientes, compreendendo programas que facilitem a entrada de clientes ou que ajudam na submissão de documentos legais, nomeadamente os *chatbots*, e aplicação de IA ao *marketing*, permitindo a análise de dados relativos aos clientes (PWC FINANCIAL SERVICES, 2018). *Chatbots* são uma tecnologia que dão apoio aos clientes através de algoritmos de processamento de linguagem natural e mecanismos de *machine learning*, tendo sido adotados por diversos bancos para melhorar as suas interfaces de *self-service* e *customer support* (BUCHANAN, 2019; FINANCIAL STABILITY BOARD, 2017). Em relação da IA ao *marketing*, ainda na esfera dos clientes, Jarek e Mazurek (2019), concluíram que a IA tem o potencial de eliminar tarefas demoradas e trabalhosas, como a recolha e análise de dados, permitindo que seja dedicado maior tempo a atividades estratégicas e criativas que conduzam a vantagens competitivas, à inovações no design e na forma como a empresa confere valor aos consumidores e ao desenvolvimento de novas competências para lidar com as novas tecnologias usadas (GARCIA, 2020). Não obstante, Cao (2020) pontua que a IA pode ser usada para analisar o comportamento do consumidor, através da sua localização e utilização de aplicações mobile e, posteriormente, realizar campanhas de marketing personalizadas, bem como prever a evolução do perfil dos consumidores e antecipar novas oportunidades no mercado.

Dando sequência à divisão das ferramentas da IA utilizadas no setor financeiro, Garcia (2021) aborda as tecnologias voltadas para a gestão de investimento e o *trading*.

Ao nível da gestão de investimentos e do *trading*, existem ferramentas de IA que permitem a execução de posições ou a gestão de portfólios. Através da análise do histórico de comportamentos do cliente, o algoritmo pode antecipar a próxima posição que este vai comprar e avaliar o perfil de risco do cliente em questão. Adicionalmente, as ferramentas de IA podem utilizar a grande quantidade de dados disponíveis para detetar movimentos nos preços de ativos no mercado, de forma a gerar retornos mais elevados (FINANCIAL STABILITY BOARD, 2017).

Segundo Halpin e Dannemiller (2019), na área de gestão de investimentos, a utilização de IA amplia a capacidade de reconhecer padrões, antecipar e prever eventos futuros e tomar boas decisões de investimento (GARCIA, 2020). Relativamente ao terceiro elemento da divisão das ferramentas, que trata das aplicações de *back-office*, estas podem centrar-se na otimização de capital e na análise de impacto no mercado (FINANCIAL STABILITY BOARD, 2017). Na análise de impacto no mercado, as instituições podem utilizar IA para obter informação e modelos históricos ou criar robôs de trading que aprendem a reagir a alterações no mercado, assim como ferramentas de IA que minimizem o impacto da negociação de grandes posições, ao nível dos preços e liquidez no mercado (GARCIA, 2020).

Por fim, a última classe da divisão trata das ferramentas voltadas à gestão do risco e *compliance*.

Segundo o Financial Stability Board (2017), a IA e a machine learning podem ser utilizadas no âmbito da RegTech, assim como algoritmos de processamento de linguagem natural, para monitorar o comportamento de traders, através de, por exemplo, análises a e-mails, documentos ou mensagens, promovendo a transparência nos mercados financeiros. Além disso, também é possível utilizar algoritmos para identificar transações suspeitas para que possam ser alvo de investigações mais aprofundadas (GARCIA, 2021).

Além dos pontos citados, cabe conceder atenção especial à utilização da IA na análise de concessão de crédito. Segundo Johnson *et al.* (2019), as decisões de crédito tomadas a partir de meios digitais reduzem os custos e a morosidade dos processos. Além disso, os autores pontuam que estudos empíricos estabelecem que modelos estatísticos avançados e análises preditivas aumentam a capacidade dos credores de calcular os riscos de inadimplências. Defensores do uso da Inteligência Artificial neste contexto pontuam que a introdução de plataformas de tomada de decisão automatizada (ADM - *automated decision-making*) mitigam os riscos de discriminação contra grupos específicos, uma vez que um sistema de computador não é, por si só, de executar processos mentais (seja cálculo consciente ou quase inconsciente emoções ou viés inconsciente) associados à ação discriminatória dos humanos (JOHNSON *et al.*, 2019). Contudo, pontua Johnson *et al.* (2019), a ADM pode apenas mudar o local de discriminação da mesa do gerente do banco para a tela do computador ou para o cientista de dados, porque dados jamais são brutos. Os dados utilizados pela ADM são sempre coletados, analisados e usados por

peças que podem ter os mesmos cálculos conscientes, emoções pouco conscientes ou vieses inconscientes em jogo em suas próprias observações (Johnson *et al.*, 2019). Neste contexto, tal qual pontuam Cheng *et. al* (2022) e Janissek-Muniz e Vassali (2022) urge a necessidade de nos atentarmos para as possíveis consequências negativas do uso da IA, principalmente em casos com impacto direto da vida das pessoas, tal qual na análise de concessão de crédito.

O Quadro 2 sintetiza os principais elementos identificados no referencial teórico, os quais servirão de base para a coleta de dados:

Quadro 2 – Consolidação dos principais conceitos da revisão teórica

Conceito	Autor	Ano	Descrição
Inteligência Artificial	Mikalef e Gupta	2021	Habilidade de um sistema identificar, interpretar, fazer inferências e aprender através de dados, para alcançar determinados objetivos organizacionais e sociais
	Gokul	2018	Uso de ferramentas computacionais na realização de tarefas que, tradicionalmente, requeriam um raciocínio humano sofisticado
	Charniak e McDermott	1985	Processo que busca humanizar máquinas, e contemplar cada vez mais a sociedade, através do estudo da mente humana, com modelos computacionais eficientes e que facilitam as tarefas diárias dos seres humanos
	Sichman	2021	Ramo da ciência/engenharia da computação que visa desenvolver sistemas computacionais que solucionam problemas
	Santos Gomes	2010	Ramo da Ciência da Computação cujo interesse é fazer com que os computadores pensem ou se comportem de forma inteligente, mimetizando a inteligência humana
Machine Learning	Ludermir	2021	Traduzido para o português como o aprendizado de máquinas - constitui-se por técnicas orientadas que buscam permitir que os sistemas aprendam automaticamente a partir de um grande volume de dados
Deep Learning	Chagas	2019	Método específico de aprendizado de máquina permite que modelos computacionais compostos de inúmeras camadas de processamento possam “aprender” representações de dados com diversos níveis de

			abstração. Esses métodos melhoraram o reconhecimento de fala, de objetos visuais, detecção de objetos, dentre outras possibilidades
Processamento Natural de Linguagem	Ludermir	2021	Campo da Inteligência Artificial que lida com a tradução e manipulação de textos, tal qual os tradutores automáticos
Visão Computacional	Janissek-Muniz e Borges	2022	Área da IA que se dedica a se dedica a “ensinar” a máquina a compreender e interpretar imagens
Ética	Doneda <i>et al.</i>	2018	Ramo do conhecimento que lida com a moralidade e, no sentido mais amplo, com a preocupação que os humanos sempre tiveram em descobrir a melhor maneira de viver, e de distinguir entre o bem e o mal, o certo e o errado
	Aristóteles	-	Conjunto de valores morais e princípios que orientam a conduta humana na sociedade.
<i>Data Ethics</i>	Floridi	2016	Definido como um novo ramo da ética, que estuda e avalia problemas morais relacionados aos dados e o seu surgimento marcou uma mudança de foco sobre o conteúdo (informação) que pode ser criado, registrado, processado e compartilhado através de meios tecnológicos para um foco nos dados
	Floridi	2016	A ética dos dados se concentra em problemas éticos colocados pela recolha e análise de grandes conjuntos de dados, abordando questões relacionadas à criação de perfis, publicidade, re-identificação de indivíduos, privacidade em grupo, discriminação e transparência, entre outras
Setor Financeiro	Mishkin	2019	Composto por instituições e mercados financeiros que ajudam a combinar a poupança de indivíduos e empresas com as necessidades de investimento de outras empresas e indivíduos. Inclui bancos, mercados financeiros, empresas de seguros e outras instituições que desempenham um papel fundamental na alocação eficiente de recursos financeiros.

Fonte: autora

Posta a explanação técnica e conceitual dos elementos, somada à consolidação destes conceitos, parte-se para a apresentação dos procedimentos metodológicos utilizados visando identificar os casos de utilização de sistemas de Inteligência Artificial no setor financeiro e discutir seus dilemas éticos.

3 PROCEDIMENTOS METODOLÓGICOS

3.1 ENQUADRAMENTO DA PESQUISA

No que tange aos procedimentos metodológicos utilizados, esta pesquisa possui uma abordagem qualitativa, com objetivo de analisar com profundidade situações particulares ou complexas (OLIVEIRA, 2011). Para tanto, optou-se pelo desenvolvimento de uma Revisão Sistemática da Literatura (RSL), compreendida como uma revisão de fontes secundárias em que considera resultados e contribuições das pesquisas para um tópico previamente determinado mediante a aplicação de métodos explícitos de busca (LAMB, 2013; LINDE; WILLICH, 2003).

Segundo Watson e Webster (2002) a revisão sistemática de literatura tem como objetivo gerar conhecimentos estruturados sobre um tema de pesquisa e pode ser usada para fazer predições razoáveis sobre o tema pesquisado, o que harmoniza com o objetivo do presente estudo. Para além disto, o modelo de revisão se mostra interessante para o caso por representar um estudo científico passível de reprodução, visto que a busca acadêmica ocorre conforme critérios pré-estabelecidos (MOURA, 2018). Não obstante, Sampaio e Mancini (2007, p. 84) descrevem, “as revisões sistemáticas nos permitem incorporar um espectro maior de resultados relevantes, ao invés de limitar as nossas conclusões à leitura de somente alguns artigos”. Neste sentido, a seguir será aprofundado como foi operacionalizada a pesquisa e, mais especificamente, quais foram os critérios utilizados para definir os “resultados relevantes”.

3.2 CONTEXTUALIZAÇÃO DA PESQUISA

O presente trabalho visa responder a seguinte questão: “Quais são os dilemas éticos por trás da utilização de soluções de Inteligência Artificial no setor financeiro?”. Serão investigadas quais são as soluções de Inteligência Artificial utilizadas no setor financeiro e os dilemas éticos envolvidos nestes casos.

A escolha do setor financeiro como campo de pesquisa deu-se devido à sua importância na vida das pessoas e das organizações, sejam elas públicas ou privadas. Segundo El Qadi *et al.* (2023), as instituições financeiras são responsáveis pelo crescimento da economia global. Além disso, de acordo com o relatório da Insider Intelligence sobre o uso da IA no setor bancário, 80% dos bancos estão cientes dos benefícios potenciais apresentados pela IA, e muitos já implementaram as tecnologias em seus processos (EL. QADI *et al.*, 2023). Não obstante, Kapoor e Bisht (2020) pontuam que se espera que os gastos com IA no setor financeiro na Ásia-Pacífico atinjam US\$ 4,29 bilhões em 2024, o que comprova o ritmo acelerado da consolidação dos uso destas tecnologias no setor.

A pesquisa se dará através da análise de bases secundárias, isto é, estudos científicos já realizados, através de uma Revisão Sistemática da Literatura, descrita a seguir.

3.3 OPERACIONALIZAÇÃO DE PESQUISA

Sendo o método de análise escolhido a Revisão Sistemática da Literatura (RSL), a presente seção abordará com maior detalhamento os passos realizados para operacionalização da análise.

3.3.1 Coleta de Dados

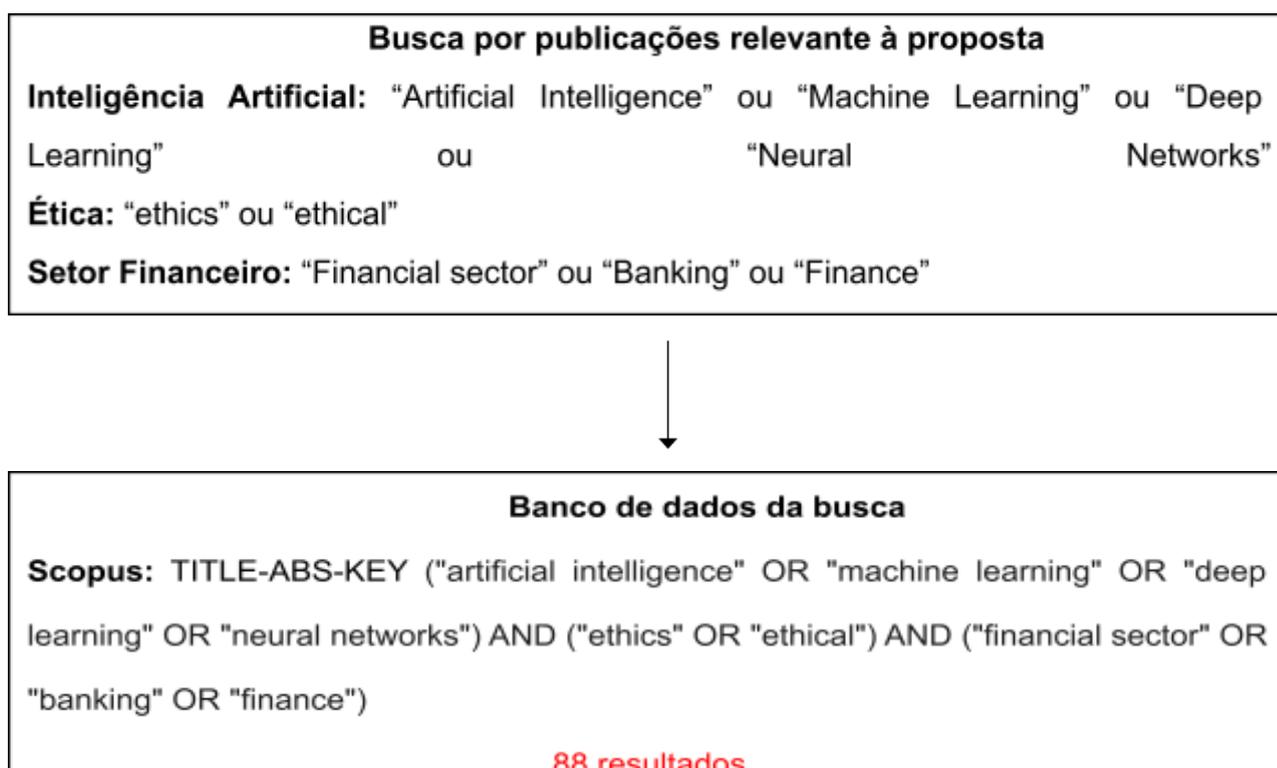
Segundo Gomes e Caminha (2014), as etapas mais empregadas para a realização da RSL, que foram elaboradas originalmente pelo Instituto Cochrane, são as seguintes: formulação da pergunta, localização e seleção dos estudos, avaliação crítica dos estudos, coleta de dados, análise e apresentação dos dados, interpretação dos dados e aprimoramento e atualização da revisão (MOURA, 2018). Isto posto, as etapas seguidas para a coleta de dados foram as seguintes:

1. Definição das palavras-chave de busca em sites de periódicos científicos;
2. definição dos sites de periódicos científicos a serem utilizados;
3. busca pelas palavras-chave nas fontes previamente definidas;
4. estabelecimento de filtros para distinguir os trabalhos de fato relevantes a pesquisa;

5. leitura dos artigos selecionados, sintetizando os principais conceitos propostos;
6. interpretação dos dados e análise crítica.

Cabe ressaltar que, tendo em vista a maior quantidade de artigos disponíveis, optou-se por operacionalizar a busca em inglês. Assim, as palavras-chave definidas na primeira etapa foram: “*Artificial Intelligence*”, “*Machine Learning*”, “*Deep Learning*”, “*Neural Networks*”, “*Ethics*”, “*Ethical*”, “*Finance Sector*”, “*Finance*” e “*Banking*”. Em relação a bases de periódicos científicos escolhidos, a escolhida para utilização da pesquisa foi a *Scopus*, devido a qualidade e a quantidade de artigos disponíveis. A figura 2 sistematiza os procedimentos de pesquisa e os resultados obtidos:

Figura 2 – Fluxograma da RSL – busca por artigos



Fonte: autora

Feitas as primeiras pesquisas e visando operacionalizar a etapa (4) do método proposto pelo Instituto Cochrane, sendo ela “estabelecimento de filtros para

distinguir os trabalhos de fato relevantes a pesquisa”, foram estabelecidos os seguintes critérios de exclusão:

Quadro 3 – Filtros da RSL – seleção de artigos

Ordem	Filtro	Justificativa
1°	Data da publicação	Tendo em vista o crescente desenvolvimento dos sistemas de Inteligência Artificial nos últimos anos, a pesquisa valeu-se da análise de artigos publicados nos últimos 5 anos, visando assim focar os esforços em análise mais atuais.
2°	Idioma	Artigos em inglês devido à maior disponibilidade de trabalhos publicados
3°	Seleção de textos	A pesquisa abordou apenas artigos científicos, excluindo livros e capítulos de livros, assim como artigos repetidos
4°	Análise dos Resumos	Nesta etapa foram lidos os resumos dos artigos visando manter apenas aqueles que estivessem verdadeiramente relacionados ao tema
5°	Leitura dos Artigos	Ao final do processo, os artigos selecionados foram lidos por inteiro e, em cima deles, foi realizada a análise dos resultados

Fonte: autora

Após a aplicação dos critérios de exclusão, 80 trabalhos foram excluídos em função da data de publicação, nenhum pelo idioma e 59 por não serem artigos científicos. Depois da leitura dos resumos, mais 13 foram excluídos por não abordarem de maneira central nenhum dos temas pesquisados. Logo, no processo de exclusão, 16 textos foram selecionados para a realização do trabalho.

3.3.2. Procedimentos de Análise

A análise dos artigos selecionados sistematicamente foi realizada de maneira qualitativa. Logo, o conteúdo de todos os textos escolhidos foi lido e analisado de maneira integral. A partir disto, extraiu-se informações e teses importantes que pudessem, em diferentes esferas, contribuir para o alcance dos objetivos do presente trabalho, sendo eles listar as principais soluções de Inteligência Artificial (IA) utilizadas no setor financeiro, mapear os dilemas éticos presentes na utilização de tais soluções e explorar os principais pontos de atenção nos casos de utilização.

4 ANÁLISE DOS RESULTADOS E DISCUSSÃO

Tendo em vista que o presente trabalho se propõe a identificar os dilemas éticos presentes na utilização de soluções de Inteligência Artificial (IA) em empresas do setor financeiro, através da Revisão Sistemática da Literatura (RSL) buscou-se analisar os dois pontos principais da temática escolhida, sendo eles: utilização da IA no setor financeiro e ética no uso da tecnologia em questão. Para isto, foram selecionados para análise artigos que tratassem sobre estes temas, ainda que separadamente.

Em relação à natureza dos artigos analisados, é possível perceber que o número de publicações aumentou ano a ano. Embora ainda estejamos no início de 2023, já estão listados alguns trabalhos publicados neste ano. Em 2022, foram sete artigos publicados, enquanto que em 2021 três. Sobre 2020 e 2019, cada ano apresentou apenas dois trabalhos publicados. A partir de tal análise pode-se perceber que o tema vem se tornando cada vez mais relevante a nível acadêmico. A figura a seguir ilustra a dimensão a distribuição dos trabalhos pelo ano de publicação:

Figura 3 – Distribuição dos textos selecionados por ano de publicação



Fonte: autora

transformative symbiosis in favour of financial technology				<i>learning</i> no setor financeiro. Discutem também o lado ético e regulamentar da utilização dos sistemas inteligentes.
Managers' understanding of artificial intelligence in relation to marketing financial services: insights from a cross-country study	Mogaji, E., Nguyen, N.P.	2022	International Journal of Bank Marketing 40(6), pp. 1272-1298	Estuda o conhecimento dos gerentes do setor financeiro a respeito das soluções de AI que estão sendo implementadas no setor de marketing das suas instituições e coleta suas percepções a respeito, considerando que eles são importantes <i>stakeholders</i> para o processo de implantação.
Artificial intelligence in customer-facing financial services: a systematic literature review and agenda for future research	Hentzen, J.K., Hoffmann, A., Dolan, R., Pala, E.	2022	International Journal of Bank Marketing 40(6), pp. 1299-1336	Revisão sistemática da literatura sobre sistemas de Inteligência Artificial voltados para os clientes utilizados no setor financeiro.
Consumer perceived ethicality of banks in the era of digitalisation: The case of Poland	Piotrowski, D.	2022	Economics and Business Review 8(1), pp. 90-114	Análise sobre a percepção dos consumidores em relação a o que faz um banco ser considerado ético.
The Different Faces of AI Ethics Across the World: A Principle-To-Practice Gap Analysis	Tidjon, L.N., Khomh, F.	2022	IEEE Transactions on Artificial Intelligence pp. 1-20	Avaliação contextual dos princípios éticos atuais em relação ao uso da IA em diferentes continentes e análise dos gaps entre a teoria e a prática das diretrizes éticas.
Ethically Responsible Machine Learning in Fintech	Rizinski, M., Peshov, H., Mishev, K., (...), Vodenska, I., Trajanov, D.	2022	IEEE Access 10, pp. 97531-97554	Estudo da dimensão ética e do uso responsável de sistemas de aprendizado de máquinas em <i>fintechs</i> .
An Overview of Artificial Intelligence Ethics	Huang, C., Zhang, Z., Mao, B., Yao, X.	2022	IEEE Transactions on Artificial Intelligence pp. 1-21	Análise abrangente sobre os riscos e questões éticas no uso da Inteligência Artificial.
Cognitive architectures for artificial intelligence ethics	Bickley, S.J., Torgler, B.	2022	AI and Society	Análise ampla da ética no uso da IA. Dá ênfase para a necessidade de se poder compreender como a inteligência chega nas soluções que chega através de maior transparência e, isto é, ênfase a IA explicável.
In pursuit of socially-minded data-intensive innovation in banking:	Aitken, M., Ng, M., Horsfall, D., (...), van	2021	Technology in Society 66,101666	Estudo da percepção e expectativas dos consumidores sobre a dimensão ética nas inovações

A focus group study of public expectations of digital innovation in banking	Moorsel, A., Elliott, K.			digitais propostas pelos bancos.
Ethical and political-economic dimensions and potential reforms of the hybrid leveraged, high frequency, artificial intelligence trading model	Nielsen, R.P.	2021	Business and Professional Ethics Journal 40(2), pp. 189-222	Questões éticas relacionadas ao uso de sistemas de inteligência artificial em operações no mercado financeiro.
AI human impact: toward a model for ethical investing in AI-intensive companies	Brusseu, J.	2021	Journal of Sustainable Finance and Investment	Propõe um modelo de análise da crítica para investidores a respeito do quão éticas são as empresas que usam a inteligência artificial intensamente.
Access to Finance for Artificial Intelligence Regulation in the Financial Services Industry	Lee, J.	2020	European Business Organization Law Review 21(4), pp. 731-757	Discussão sobre as dimensões regulamentares e éticas para melhorar e estimular o acesso ao setor financeiro. Argumentam que a IA deve ser desenvolvida visando garantir acesso, proteção ao consumidor e integridade do mercado.
It's All About Data: How to Make Good Decisions in a World Awash with Information	Mahdavi, M., Kazemi, H.	2020	Journal of Financial Data Science 2(2), pp. 8-16	Análise de como a Inteligência Artificial pode ajudar profissionais do setor financeiro a tomarem boas decisões.
User abnormal behavior recommendation via multilayer network	Song, C., Liu, W., Liu, Z., Liu, X.	2019	PLoS ONE 14(12), e0224684	Proposição de um modelo de recomendação de serviços financeiros sem a utilização de dados privados.
Super intelligent financial services	Fourie, L., Bennett, T.K.	2019	Journal of Payments Strategy and Systems 13(2), pp. 151-164	Análise ampla de como as organizações do setor financeiro estão usando os sistemas de Inteligência Artificial.

Fonte: autora

Dos 16 estudos sistematicamente selecionados, 4 abordam apenas a dimensão ética da IA, sem delimitar setores de aplicação. Outros 4 estudam a utilização dos sistemas de Inteligência Artificial no setor financeiro, sem necessariamente incluir a dimensão ética na análise. Os outros 8 textos citados relacionam, ainda que em diferentes esferas, a ética no uso da IA especificamente

no setor financeiro. Visando atingir o objetivo do presente trabalho, em um primeiro momento se discutirá sobre os textos que tratam apenas sobre a dimensão ética no uso da Inteligência Artificial e após especificamente sobre o uso da IA no setor financeiro. Por fim, será estabelecida a relação entre ambos os temas.

Em seu trabalho, Tidjon e Khomh (2022) pontuam que existe uma lacuna entre os princípios éticos e a prática real dos mesmos no uso da IA ao redor do mundo. Especificamente em relação ao âmbito ético da Inteligência Artificial, os autores colaboram indicando que existem vários problemas nesta questão, principalmente no que tange às questões de segurança da informação, privacidade, preconceito e responsabilidade. Sinalizam que as tecnologias de IA têm o potencial de perpetuar e aumentar preconceitos já existentes em áreas como concessões de crédito, emprego e justiça, dado que as máquinas aprendem através do conhecimento humano (TIGJON; KHOMH, 2022). Para além disto, os autores sublinham a importância de tornar a inteligência artificial transparente para que seja possível entender e garantir que as tomadas de decisões advindas das máquinas sejam justas e éticas.

Da mesma maneira, Huang *et al.* (2022), em uma análise ampla da ética da Inteligência Artificial, dissertam a respeito dos desafios a serem enfrentados pela IA, bem como possíveis abordagens de tratamento para os mesmos. Os autores conferem destaque aos desafios relacionados à privacidade e segurança de dados, transparência nas tomadas de decisão automatizadas e vieses ou preconceitos algorítmicos. Em relação a como enfrentar tais desafios, o trabalho propõe que sejam desenvolvidos sistemas de monitoramento e auditoria da IA, bem como regulamentos e políticas de governança sobre.

Bickley e Torgler (2022) contribuem dando ampla importância a dimensão ética no uso da IA e alertando para os possíveis riscos que a desatenção neste tema pode levar, tal qual a perpetuação de preconceitos e discriminação. Além disso, relatam que o ponto crucial para uma adoção e integração mais ampla dos sistemas da IA na sociedade é que estes se tornem mais transparentes, explicáveis e responsabilizáveis. Neste sentido, sinalizam que as arquiteturas cognitivas propostas por eles são promissoras para atingir tais objetivos. Na mesma linha, Brosseau (2021) argumenta em seu trabalho que o desenvolvimento acelerado dos sistemas

de IA também tem um potencial negativo e que as empresas que utilizam as tecnologias devem ser responsabilizadas por esses possíveis impactos na sociedade. Ao propor um modelo de avaliação de empresas que utilizam intensivamente tecnologias de Inteligência Artificial visando que os investidores possam avaliar o quão éticas são as empresas antes de investir nelas, o autor cita que deve ser levado em consideração a transparência das organizações em relação ao uso dos sistemas, o comprometimento com a proteção de dados e dos direitos humanos.

Em suma, os quatro trabalhos selecionados sistematicamente que abordam o tema da ética no uso da inteligência artificial, sem delimitar o setor de uso, apresentam grande concordância em relação às principais questões éticas a serem enfrentadas neste contexto, sendo elas a transparência, privacidade, segurança de dados, preconceito e perpetuação de vieses. Todos estes pontos estão centrados na *black box* da IA, isto é, na caixa preta da IA. O funcionamento dos sistemas inteligentes não é claro para todos - os dados utilizados são confidenciais e o processo de tomada de decisão é uma “caixa preta” que eventualmente nem mesmo os desenvolvedores dos sistemas podem explicar (BICKLEY; TORGLER, 2021). Neste sentido, por não se conhecer o caminho para a tomada de decisão das máquinas, torna-se quase impossível auferir a respeito da ética e da justiça destas decisões das máquinas. Em relação a como resolver este problema, Tidjon e Khomh (2022) propõem um esforço coletivo entre governos, indústrias, organizações não governamentais e sociedade no geral para que sejam eliminados os gaps entre os princípios e a prática nas diretrizes éticas do uso da IA. Já para Huang *et al.* (2022) devem ser desenvolvidos sistemas de auditoria para o uso dos sistemas, bem como regulamentos e políticas de governança sobre. Bickley e Torgler (2022) apontam para o uso de arquiteturas cognitivas para compreender a caixa-preta e Brosseau (2021) não propõe claramente uma maneira de enfrentar os dilemas éticos, mas sim de como avaliar a ética das organizações que utilizam a IA.

Dada a análise dos textos que abordam exclusivamente a dimensão ética no uso da inteligência artificial, parte-se para o estudo daqueles que tratam do uso de sistemas de IA no setor financeiro, sem necessariamente abordar aspectos éticos. Em seu trabalho sobre a percepção dos gestores de financeiros sobre o uso dos

sistemas inteligentes, Mogaji e Nguyen (2022) contribuem pontuando inúmeras formas como a inteligência artificial vem sendo usada no setor financeiro, sendo elas carteiras inteligentes, digitalização de pagamentos, serviços assistidos por voz, sistemas de prevenção a fraude e identificação de risco, análise de documentações legais de maneira ágil, sistemas de indicação, identificações de padrões e previsões do mercado de capitais e *chatboots*. Pontuam que a utilização do sistema destas formas torna o sistema financeiro como um todo mais ágil, aprimora a capacidade analítica de gestores através do processamento de grande quantidade de dados, proporciona maior engajamento e satisfação dos clientes, principalmente através dos *chatboots*. Além disso, reiteram que os sistemas de IA possuem ampla capacidade de diminuir os custos operacionais e de aumentar a qualidade dos serviços prestados, diminuindo o risco de erro humano.

Em concordância com Mogaji, Nguyen, (2022) e Hentzen *et al.* (2022), através de uma revisão sistemática da literatura, também citam diversas formas de como a Inteligência Artificial vem sendo utilizada no setor financeiro. Para além dos que haviam sido explorados por Mogaji e Nguyen (2022), dissertam a respeito dos robôs *advisors* e os sistemas automatizados de gestão de portfólio de investimentos. Em relação a estas soluções, relatam que as mesmas reduzem a necessidade de mão de obra humana na medida que existe muita demanda advinda daqueles que temiam fraudes por parte dos seus gestores de portfólio. Além disto, dão importância significativa a utilização de *chatboots* e assistentes virtuais, que promovem a maior proximidade e engajamento dos clientes e reduzem os custos operacionais, tal qual os robôs *advisors* e os sistemas de gestão de portfólio automatizados.

No que tange os outros dois textos que abordam com maior foco a relação da Inteligência Artificial e o setor financeiro, sendo eles os trabalhos de Mahdavi, Kazemi (2020), Fourie e Bennett (2019), ambos os artigos estavam indisponíveis para a leitura. Porém, em relação ao primeiro, além do resumo, também foram disponibilizadas as principais conclusões. Elas apontam que o aumento da quantidade de dados disponíveis em finanças está criando grandes oportunidades em todas as áreas do setor financeiro, incluindo o gerenciamento de riscos e fraudes, construção de portfólio, bancos de investimento e seguros. Ademais, pontuam em relação a importância da presença de profissionais do setor no

acompanhamento destes sistemas, dados os riscos do uso da IA, como a falta de transparência e os possíveis vieses gerados. Tal qual Mahdavi, Kazemi (2020), Mogaji, Nguyen (2022) e Hentzen *et al.* (2022) também fazem apontamentos sobre a dimensão ética no uso da IA, ainda que não seja este o foco central dos trabalhos. Ambos citam que deve haver uma atenção especial voltada a esta temática dados os riscos relativos à privacidade e perpetuação de preconceitos que os sistemas apresentam. Ao final de sua RSL, Hentzen *et al.* (2022) inclusive sugerem para futuras pesquisas trabalhos que abordem os impactos sociais, éticos e regulatórios do uso da IA no Setor Financeiro.

Por fim, parte-se para a exploração de trabalhos que abordam todos os pontos-chave do presente estudo, sendo eles a Inteligência Artificial, ética e o setor financeiro. Vale ressaltar desde já que os trabalhos de Nielsen (2021) e Nguyen *et al.* (2023) estavam indisponíveis para leitura.

El Qadi. *et al* (2023) realizam um estudo da dimensão ética no uso da Inteligência Artificial focado na utilização dos sistemas para análise para concessão de crédito para empresas. Dissertam que o uso de tecnologias de Machine Learning para o credit scoring é muito positivo, no entanto, de maneira conforme aos outros textos, aponta para os riscos éticos da caixa-preta da IA. Indica o fato de não poder se explicar como e por que as decisões automatizadas são tomadas como um grande problema, dado que quando estamos tratando de temas como concessão de crédito estamos tratando de questões altamente importantes para o desenvolvimento socioeconômico. Propõe como solução a *explainable AI*, isto é, Inteligência Artificial Explicável (XAI).

Em uma abordagem voltada à percepção dos consumidores, Piotrowski (2022) conduz um estudo de caso na Polônia buscando entender com maior profundidade como a digitalização influenciou na visão dos clientes sobre a ética. O autor chega a conclusão que os consumidores enxergam como entidades éticas e consideram positivas as transformações digitais quando estas são feitas de maneira responsável socialmente, transparente e conferem privacidade e segurança dos dados. Pode-se relacionar tal conclusão às análises discutidas nos textos anteriores que, em abordagem distinta, também confere importante relevância a temas como privacidade, segurança de dados e transparência.

Com abordagem similar a Piotrowski (2022), Aitken *et al.* (2021) também conduz uma análise focada na percepção de consumidores, mas desta vez na avaliação do público em relação às inovações no setor bancário, abordando os benefícios, os desafios e a responsabilidade social inerente ao tema. Os resultados do estudo mostraram que as inovações digitais melhoram a percepção dos clientes em relação às instituições financeiras ao desenvolver a eficiência e a conveniência dos serviços bancários. No entanto, os resultados também mostram que os participantes tinham a expectativa que tais inovações adotadas fossem conduzidas de maneira ética e socialmente responsáveis. De forma análoga aos outros estudos analisados, os participantes do grupo focal se mostraram preocupados com questões de transparência, privacidade e segurança de dados.

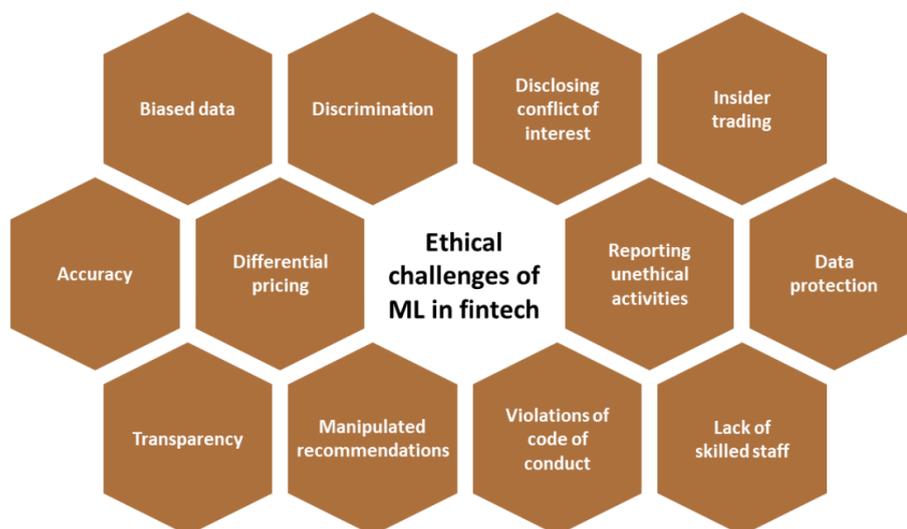
Também em concordância com os outros trabalhos, Lee (2020) diz que a Inteligência Artificial tem o potencial de transformar o setor financeiro ao conferir uma melhora na eficiência e qualidade dos serviços. Por outro lado, pontua que esta relação possui riscos éticos e regulatórios. Lee (2020) apresenta um viés focado em dilemas regulatórios, que não são o foco do presente trabalho. Contudo, contribui citando algumas aplicações da AI voltadas para o mercado de capitais especificamente, tal qual sistemas de recomendação de movimentação financeira, tal qual os robôs *advisors*, e sistemas para regulação de mercado.

Em abordagem diferente das analisadas até então, Song *et al.* (2019) propõe um modelo de recomendação de ativos financeiros ético, partindo do pressuposto que outros modelos utilizam dados privados dos consumidores para fazer as proposições. Estes sistemas requerem um grande volume de dados para treinar as máquinas, e por este aspecto é um desafio projetar estes recursos sem olhar para dados privados (SONG *et al.*, 2020). Neste sentido, os autores propõem um modelo de recomendação ético e com segurança de dados, reiterando, mais uma vez, os dilemas relacionados à privacidade e segurança de dados.

Por fim, Rizinski *et al.* (2022) realizam uma análise ampla e análoga a do presente trabalho discutindo a necessidade de se desenvolver uma Inteligência Artificial ética no setor financeiro. Dão atenção principal aos riscos de discriminação, preconceito, falta de transparência e responsabilidade. A figura 3, proposta pelos

autores, resume todos os dilemas éticos que identificaram na utilização do *machine learning* no setor financeiro, mais especificamente em *fintechs*.

Figura 5 – Desafios Éticos do *machine learning* em *fintechs*



Fonte: Rizinski *et al.* (2022), p. 9

Por fim, Rizinski *et al.* (2022) propõe que a solução para os dilemas éticos citados está na focada na transparência, responsabilidade e justiça, e que para os sistemas serem justos e éticos eles precisam ser desenvolvidos cuidadosamente, os dados utilizados devem ser analisados rigorosamente, os processos devem ser amplamente compreendidos e devem haver testes de validação dos sistemas

Após a análise e estudo aprofundado dos textos, pode-se explorar com amplitude os objetivos do trabalho, sendo eles listar as principais soluções de IA utilizadas no setor financeiro, mapear os dilemas éticos presentes na utilização de tais soluções e a discussão de tais relações. O quadro 5 sintetiza o resultado desta análise.

Quadro 5 - Dilemas éticos no uso da inteligência artificial no setor financeiros

Forma de utilização da IA no setor financeiro	Dilema ético	Discussão
<i>Chatbots</i> e assistentes virtuais	Falta de transparência	Os <i>chatbots</i> ou assistentes virtuais são ferramentas interessantíssimas aos clientes das organizações do setor financeiro porque os mantêm próximos e conferem agilidade ao atendimento dos mesmos. Contudo, devido a caixa-preta da IA, muitas vezes não é possível

		identificar o caminho que o assistente virtual levou para tomar determinada decisão ou dar determinada indicação, o que confere a solução falta de transparência na tomada de decisão e todos os potenciais problemas advindos dela.
Análise de crédito/financiamento	Perpetuação de vieses; preconceito; baixa transparência sobre o processo de tomada de decisão	A análise de crédito ou financiamento é um dos pontos mais críticos no que tange a dimensão ética do uso da Inteligência Artificial no mercado financeiro. Isto porque a decisão final tomada pela máquina, de conceder ou não crédito ou financiamento para pessoas físicas ou jurídicas têm impacto relevante no desenvolvimento socioeconômico. Neste sentido, a falta de transparência de como são tomadas as decisões abrem espaço para auferir que as máquinas estejam propagando preconceito ou vieses que tenham sido ensinadas, intencionalmente ou não, pelos humanos. Tais vieses podem ser em relação ao gênero, raça, escolaridade ou profissão, por exemplo.
Sistemas de recomendações de ativos financeiros/ sistemas automatizados de gestão de portfólio	Segurança e proteção de dados; falta de transparência; substituição da força de trabalho humana	O principal dilema ético enfrentado pelo uso de sistemas de recomendação de ativos financeiros ou de gestão de portfólio automatizados está ligado à segurança de proteção dos dados. Isto porque, como pontuam Song <i>et al.</i> , (2020), estes sistemas requerem um grande volume de dados para treinar as máquinas, e por este aspecto é um desafio projetar estes recursos sem olhar para dados privados. Neste sentido, e aliado à falta de transparência na tomada de decisão das máquinas, torna-se um desafio pontuar qual é o embasamento dos sistemas para recomendar os ativos. Para além disto, Mogaji, E., Nguyen, N.P. (2022) aponta que estes sistemas podem ser uma ameaça para os gestores de portfólio ou assessores de investimento, dado que alguns clientes tendem a confiar mais na recomendação das máquinas pois com estas não possuem o risco de receber sugestões ancoradas em comissões ou fraudulentas.
Análise de documentação/ análise de risco	Perpetuação de vieses; preconceito; falta de transparência	Sistemas de IA implementados para prevenir fraudes e analisar riscos inerentes ao setor financeiro são

		positivos pois tornam tais processos ágeis e rápidos. No entanto, esta análise, por exemplo, pode vir a banir o cliente da abertura de conta em uma instituição financeira dada uma decisão de máquina onde não é possível se avaliar os caminhos que o sistema tomou para chegar a uma conclusão. O problema aqui também está centrado na falta de transparência das soluções.
Marketing Financeiro	Segurança de dados	Na medida que as soluções de Inteligência Artificial são utilizadas para auxiliar o setor de marketing das instituições financeiras e fazer propagandas específicas para cada cliente, surge a questão da utilização de dados privados dos clientes para compreensão do comportamento e preferências dos clientes. O que foi citado por Song <i>et al.</i> (2020) em relação aos sistemas de recomendação de ativos cabe aqui também, ou seja, estes sistemas requerem um grande volume de dados para treinar as máquinas, e por este aspecto é um desafio projetar estes recursos sem olhar para dados privados.

Fonte: autora

Tem-se, a partir da análise sistemática dos artigos, que os principais casos de uso da inteligência artificial no setor financeiro são o uso de *chatboots* ou assistentes virtuais, sistemas de análise de crédito/financiamento automatizados, de recomendações de ativos financeiros e gestão de portfólio automatizada, análise de documentações, análise de risco e sistemas de prevenção a fraude. Pode-se dizer que a IA é utilizada no setor financeiro para automatização de processos, tornando-os mais rápidos e menos custosos. No entanto, este processo de automatização está enfrentando desafios frente a questões éticas, sociais e regulamentares. No que tange a dimensão ética, a questão central está na falta de transparência dos processos de IA, problema também chamado de caixa preta da IA. Este desconhecimento global dos métodos utilizados pela Inteligência Artificial para as tomadas de decisão faz com que não se possa avaliar a existência de preconceitos, vieses e discriminações presentes nestas tecnologias. Mas, dado que as máquinas aprendem através do conhecimento humano, é amplamente possível

que também tenham assimilado tais vieses. Urge, portanto, como citado por vários dos textos analisados, a necessidade do desenvolvimento de sistemas explicáveis e socialmente responsáveis. Em um setor que possui amplo impacto na sociedade e na vida tanto de pessoas físicas como jurídicas, é de extrema importância que tais sistemas de inteligência artificial responsável sejam desenvolvidos de maneira rápida. Na medida em que forem desenvolvidas tais tecnologias, poderão ser expandidas de maneira ainda mais rápida, e os impactos positivos que a Inteligência Artificial traz ao setor financeiro ampliados.

5 CONSIDERAÇÕES FINAIS

Este trabalho buscou identificar os dilemas éticos presentes na utilização da Inteligência Artificial no setor financeiro. A pesquisa foi operacionalizada através da Revisão Sistemática da Literatura, tendo a base de dados Scopus como a fonte de busca. Através da RSL foi possível identificar, escolher e analisar com profundidade pesquisas relevantes ao tema, que permitiram com que fosse alcançado o objetivo do trabalho.

Por meio da Revisão Sistemática, 16 textos que tratavam de aspectos relevantes ao tema foram selecionados, ou seja, artigos que abordassem, ainda que diferentes dimensões e aspectos, a ética no uso da Inteligência Artificial de forma ampla, a utilização da IA no setor financeiro e, por fim, as questões éticas da utilização destas tecnológicas especificamente no setor financeiro. Ao final do estudo dos trabalhos foi possível identificar as principais questões éticas no uso da Inteligência Artificial, bem como as formas de uso da IA predominantes do setor financeiro e, logo, os dilemas éticos presentes na utilização da IA no setor financeiro. Os resultados apontaram que os principais obstáculos éticos de tais sistemas no setor financeiro residem na falta de transparência na tomada de decisão automatizada, o que, potencialmente, pode envolver a utilização de dados privados e a propagação de vieses, preconceito e discriminação. Na medida que as tecnologias de IA são utilizadas no setor financeiro para atender os clientes virtualmente (*chatbots*), para análise de crédito e financiamento, em sistemas de recomendação de investimentos e na análise de risco, por exemplo, a falta de

clareza e o potencial enviesamento das decisões automatizadas se tornam preocupantes, pois possuem impacto direto no desenvolvimento sociedade.

No entanto, ainda que o tema seja de extrema relevância, tendo em vista e relevância ímpar do setor financeiro na vida das pessoas, tanto físicas como jurídicas, até o presente momento são poucos os trabalhos que abordam o tema. A figura 3, no entanto, mostra como o número de publicações vem crescendo a cada ano. Neste cenário, o presente trabalho contribui para a academia na medida que discute um tema que deveria ser mais explorado. Não obstante, colabora ao propor novas agendas de pesquisa.

Na prática, o presente trabalho favorece a sociedade ao trazer para discussão e alertar em relação a dilemas éticos na utilização de sistemas que, tradicionalmente, são vistos apenas de maneira positiva. Como pontuam Janissek-Muniz e Vassali (2022), a maior parte da atenção tem focado na perspectiva positiva do desenvolvimento da Inteligência Artificial. Logo, trabalhos que tratem de possíveis aspectos negativos e alertem em relação aos mesmos se tornam de grande valia para a sociedade, que passa a ficar atenta a como tais dilemas podem impactar sua vida, assumindo postura de maior questionamento e não apenas de aceitação frente às máquinas.

No que tange o setor financeiro, o presente trabalho também coopera ao informar a respeito de potenciais adversidades que as organizações do setor podem vir a enfrentar ao utilizarem tais sistemas. Contudo, vale sinalizar que as contribuições do presente trabalho podem ser facilmente estendidas para outros setores da economia, pois as problemáticas de falta de transparência, segurança de dados e propagação de vieses não se limitam apenas ao âmbito financeiro. As contribuições deste trabalho às organizações, do setor financeiro ou não, se concentram no alerta que o estudo faz dos possíveis impactos negativos dos sistemas de Inteligência Artificial. Sichman (2021) pontua que estamos atravessando um período onde os benefícios que a IA pode prover são vistos com muita euforia. No entanto, é de suma importância que as organizações não assumam postura eufórica e implementem as soluções de Inteligência Artificial sem se atentar aos possíveis impactos negativos das mesmas, dadas as potenciais consequências desfavoráveis que uma empresa pode arcar caso seja de conhecimento global o uso

não ético de soluções, mesmo que de maneira não proposital. Em um contexto onde as informações se propagam de maneira extremamente rápida, a informação de que uma organização está utilizando sistemas de IA de maneira antiética pode ser fatal para a sobrevivência da mesma. Assim sendo, este estudo contribui com as instituições, principalmente as financeiras, ao discutir as dimensões éticas do uso da Inteligência Artificial.

Para além das contribuições, faz-se necessário pontuar a respeito das limitações da pesquisa, que estiveram centradas no número pequeno número de publicações analisadas. Ademais, foi utilizada apenas uma base de publicações. Para futuras pesquisas, recomenda-se a utilização de mais bases e, conseqüentemente, mais trabalhos. Não obstante, o trabalho se limitou a análise de bases secundárias, ficando carente de evidências práticas do uso das soluções no setor financeiro e os dilemas vivenciados no dia a dia. Neste sentido, além da análise de mais bases de publicações, tem-se como sugestões de pesquisas futuras trabalhos sobre a Inteligência Artificial explicável no setor financeiro, solução proposta, mas pouco explorada por alguns autores frente ao problema da falta de transparência dos sistemas. Ademais, trabalhos que abordem estudos de caso da utilização dos sistemas de IA no setor financeiros e seus dilemas éticos vivenciados na prática podem também ser de grande valia.

REFERÊNCIAS

- ACCENTURE. **Banking Technology Vision 2019**. Accenture. [S. l.], 2019. Disponível em: https://bankingblog.accenture.com/wp-content/uploads/2019/08/Accenture-Technology-Vision-Banking-2019_DACH.pdf. Acesso em:
- AGÊNCIA Senado. Comissão conclui texto sobre regulação da inteligência artificial no Brasil. *In: Senado Notícias*. Brasília, 06 dez. 2022. Disponível em: <https://www12.senado.leg.br/noticias/materias/2022/12/06/comissao-conclui-texto-sobre-regulacao-da-inteligencia-artificial-no-brasil>. Acesso em 17 fev. 2023.
- AITKEN, Mhairi *et al.* In pursuit of socially-minded data-intensive innovation in banking: A focus group study of public expectations of digital innovation in banking. **Technology in Society**, [s. l.], v. 66, p. 101666, 2021. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0160791X2100141X>. Acesso em: 13 mar. 2023.
- ALLAS, T. *et al.* (2018). Artificial intelligence is getting ready for business, but are businesses ready for AI? **McKinsey Analytics**: Analytics comes of age, [s. l.], p. 18–34, 2018. Disponível em: <https://www.mckinsey.com/~media/McKinsey/BusinessFunctions/McKinseyAnalytics/OurInsights/Analyticscomesofage/Analyticscomesofage.ashx>. Acesso em:
- ARENCIBIA, Mario González; CARDERO, Dagmaris Martínez. Dilemas éticos en el escenario de la inteligencia artificial. **Economía y Sociedad**, [s. l.], v. 25, n. 57, p. 93-109, 2020. Disponível em: <https://www.scielo.sa.cr/pdf/eys/v25n57/2215-3403-ey-25-57-93.pdf>. Acesso em 17 mar. 2023.
- BICKLEY, Steve J.; TORGLER, Benno. Cognitive architectures for artificial intelligence ethics. **AI & SOCIETY**, [s. l.], p. 1-19, 2022. Disponível em: <https://link.springer.com/article/10.1007/s00146-022-01452-9>. Acesso em: 12 mar. 2023.
- BORGES, Natália; JANISSEK-MUNIZ, Raquel. O uso da Inteligência Artificial no Foresight: status e potencialidades. *In: IFBAE Congresso do Instituto Franco-Brasileiro de Administração de Empresas*, 11, 2022, França. **Anais eletrônicos [...]**. Rennes, IFBAE, 2022. Disponível em: https://www.researchgate.net/publication/361825907_O_uso_da_Inteligencia_Artificial_no_Foresight_status_e_potencialidades. Acesso em: 15 mar. 2023.
- BRITES, Elise Eleonore. Dilemas Éticos e a possibilidade do Desvio de Conduta. *In: JUSBRASIL*. [s. l.], 2020. Disponível em: <https://elisebrites.jusbrasil.com.br/artigos/870731969/dilemas-eticos-e-a-possibilidade-e-do-desvio-de-conduta>. Acesso em 23 fev. 2023.
- BRITZ, G. **Freie entfaltung durch selbstdarstellung**. 1. ed. Tübingen: Mohr Siebeck, 2007.

BRUSSEAU, James. AI human impact: toward a model for ethical investing in AI-intensive companies. **Journal of Sustainable Finance & Investment**, New York, p. 1-28, 2021. Disponível em: <https://www.tandfonline.com/doi/epdf/10.1080/20430795.2021.1874212?needAccess=true&role=button>. Acesso em: 11 mar. 2023.

BUCHANAN, B. G. (2019). Artificial intelligence in finance. **The Alan Turing Institute**. Seattle, p. 1-45, 2019. Disponível em: https://www.turing.ac.uk/sites/default/files/2019-04/artificial_intelligence_in_finance_-_turing_report_0.pdf. Acesso em: 19 jan. 2023

CALVO, Patrici. The ethics of Smart City (EoSC): moral implications of hyperconnectivity, algorithmization and the datafication of urban digital society. **Ethics Inf. Technol**, [S. l.], v. 22, p. 141–149, 2020. Disponível em: <https://link.springer.com/article/10.1007/s10676-019-09523-0>. Acesso em: 17 mar. 2023

CAO, L. AI in Finance: A Review. **SSRN**, Australia, v. 0, n. 0, p. 1–47, 2020. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3647625. Acesso em: 17 jan. 2023

CARREÑO, R. *¿Qué sucede con las plataformas digitales latinoamericanas?. In: LinkedIn*, [S. l.], 2019. Disponível em: <https://www.linkedin.com/pulse/qu%C3%A9-sucede-con-las-plataformas-digitales-ren%C3%A9-carre%C3%B1o/?originalSubdomain=es>. Acesso em: 20 mar. 2023

CHARNIAK, E.; MCDERMOTT, D. **A Bayesian Model of Plan Recognition**. Massachusetts: Addison-Wesley, 1985.

CHAGAS, Edgar Thiago De Oliveira. **Deep Learning e suas aplicações na atualidade**. Revista Científica Multidisciplinar Núcleo do Conhecimento. Ano 04, Ed. 05, Vol. 04, pp. 05-26 Maio de 2019. ISSN: 2448-0959. Disponível em: <https://www.nucleodoconhecimento.com.br/administracao/deep-learning>. Acesso em: 04 abr. 2023

CHENG, Xusen *et al.* The good, the bad, and the ugly: Impact of analytics and artificial intelligence-enabled personal information collection on privacy and participation in ridesharing. **European Journal of Information Systems**, [s. l.], v. 31, n. 3, p. 339-363, 2022. Disponível em: https://www.tandfonline.com/doi/abs/10.1080/0960085X.2020.1869508?casa_token=nUINtv3tjdMAAAAA%3A2FrOqfcX5y9PfcRLWmgyN484Gamr1Pz3a6odatOqWNSalcSH4HOLxCWLbvMZ9r-5xkZLWmUE4RU2&journalCode=tjis20. Acesso em: 21 fev. 2023.

DE MELLO, Mariana Rodrigues Gomes; DA SILVA CAMILLO, Everton; DOS SANTOS, Beatriz Rosa Pinheiro. Big data e inteligência artificial: aspectos éticos e legais mediante a teoria crítica. **Complexas Revista de Filosofia Temática**, Belém, v. 3, n. 1, p. 50-60, 2018. Acesso em: 18 fev. 2023.

DELOITTE the ai dossier by deloitte institute. **Deloitte**. [S. l.], 2021. Disponível em: <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/deloitte-analytics/us-ai-institute-ai-dossier-full-report.pdf>. Acesso em 03 mar. 2023.

DIGNUM, V. **Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way**. 1. ed. [S. l.]: Springer, 2019.

DISTRITO Report. **Inteligência artificial report**. [S. l.], 2021. Disponível em: <https://materiais.distrito.me/mr/inteligencia-artificial>, Acesso em 23 fev. 2023.

DONEDA, Danilo Cesar Maganhoto *et al.* Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal. **Pensar Revista de Ciências Jurídicas**, Fortaleza, v. 23, n. 4, p. 1-17, 2018. Disponível em: <https://ojs.unifor.br/rpen/article/view/8257>. Acesso em: 21 fev. 2023.

DONEPUDI, Praveen Kumar. AI and Machine Learning in Banking: A Systematic. **Asian Journal of Applied Science and Engineering**, [s. l.], v. 6, n. 3, p. 157-162, 2017. Disponível em: https://www.researchgate.net/profile/Praveen_Donepudi/publication/349859207_AI_and_Machine_Learning_in_Banking_A_Systematic_Literature_Review/links/6044553c4585154e8c80ae56/AI-and-Machine-Learning-in-Banking-A-Systematic-Literature-Review.pdf?_sg%5B0%5D=started_experiment_milestone&origin=journalDetail. Acesso em: 19 fev. 2023.

El Qadi, A. *et al.* Feature contribution alignment with expert knowledge for artificial intelligence credit scoring. **Signal, Image and Video Processing**, [s. l.], v. 17, p. 427–434, 2023. Disponível em: <https://link.springer.com/article/10.1007/s11760-022-02239-7>. Acesso em: 17 mar. 2023

FINANCIAL Stability Board. Artificial Intelligence and Machine Learning in Financial Services: Market Developments and Financial Stability Implications. **FSB Financial Stability Board**. [S. l.], p. 1-41, 2019. Disponível em: <https://www.fsb.org/wp-content/uploads/P011117.pdf>. Acesso em:

FLORIDI, Luciano. Information ethics, its nature and scope. **SIGCAS Comput. Soc.**, New York, v. 36, n. 3, p. 21-36, 2016. DOI:10.1145/1195716.1195719.

FLORIDI, Luciano; TADDEO, Mariarosaria. Introduction: What is data ethics? **Phil. Trans. R. Soc.**, Londres, v. 374, n. 2083, 2016. DOI: 10.1098/rsta.2016.0360.

GARCIA, Ana Cristina. Ética e inteligencia artificial. **Computação Brasil**, [s. l.], n. 43, p. 14-22, 2020. Disponível em: <https://sol.sbc.org.br/journals/index.php/comp-br/article/view/1791>. Acesso em: 20 fev. 2023.

Gokul, B. Artificial Intelligence in Financial Services. **Sansmaran Management Research Journal**, Mumbai, v. 8, n. 1, p. 3-5, 2018. Disponível em:

<https://www.proquest.com/openview/3a206bafb8a5ea4f8b24fd0c8e4178a8/1.pdf?pq-origsite=gscholar&cbl=2046061>. Acesso em:

GOMES, D. dos S. Inteligência Artificial: conceitos e aplicações. **Revista Olhar Científico**, [s. l.], v1, n. 2, p. 234-246, 2010. Disponível em: https://www.professores.uff.br/screspo/wp-content/uploads/sites/127/2017/09/ia_intro.pdf. Acesso em: 17 fev. 2023.

GOMES, I. S.; CAMINHA, I. O. Guia para estudos de revisão sistemática: uma opção metodológica para as Ciências do Movimento Humano. **Revista Ensaio**, Porto Alegre, v. 20, n. 01, p. 395-411, 2014. Disponível em: <http://www.seer.ufrgs.br/Movimento/article/viewFile/41542/28358>. Acesso em: 8 out. 2022.

GONZÁLEZ, Mario Orestes Aguirre; TOLEDO, José Carlos de. A integração do cliente no processo de desenvolvimento de produto: revisão bibliográfica sistemática e temas para pesquisa. **Production**, Natal, v. 22, p. 14-26, 2012. Disponível em: <https://www.scielo.br/j/prod/a/Lg5tRKqW6QNwcW4QdNGJhyk/abstract/?lang=pt>. Acesso em: 17 mar. 2023.

GONZÁLEZ-ESTEBAN, Elsan, CALVO, Patrici. Ethically governing artificial intelligence in the field of scientific research and innovation. **Heliyon**, [s. l.], v. 8., p. 1-9, 2022. Disponível em: <https://repositori.uji.es/xmlui/handle/10234/197407>. Acesso em:

HALPIN, L., DANNEMILLER, D. Artificial intelligence: The next frontier for investment management firms. **Deloitte**. [S. l.], p. 1-15, 2019. Disponível em: <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/financial-services/artificial-intelligence-investment-mgmt.pdf>. Acesso em:

HEIZENBERG, Jorgen; JUDAH, Saul; DUNCAN, Alan D. What are the essential roles for data and analytics? **Gartner Research**. [S. l.], 16 ago. 2022. Disponível em: <https://www.gartner.com/en/doc/what-are-the-essential-roles-for-data-and-analytics>. Acesso em 3 mar. 2023.

HENTZEN, Janin Karoli *et al.* Artificial intelligence in customer-facing financial services: a systematic literature review and agenda for future research. **International Journal of Bank Marketing**, [s. l.], v. 40, n. 6, p. 1299-1336, 2022. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/IJBM-09-2021-0417/full/pdf?title=artificial-intelligence-in-customer-facing-financial-services-a-systematic-literature-review-and-agenda-for-future-research>. Acesso em: 15 mar. 2023.

HUANG, Changwu *et al.* An overview of artificial intelligence ethics. **IEEE Transactions on Artificial Intelligence**, [s. l.], p. 1-21, 2022. Disponível em: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9844014>. Acesso em: 10 mar. 2023.

JAREK, K., MAZUREK, G. Marketing and artificial intelligence. **Central European Business Review**, [S. l.], v. 8, n. 2, p. 46–55. DOI <https://doi.org/10.18267/j.cebr.213>.

JOHNSON, Kristin; PASQUALE, Frank; CHAPMAN, Jennifer. Artificial intelligence, machine learning, and bias in finance: toward responsible innovation. **Fordham L. Rev.**, [s. l.], v. 88, p. 499-529, 2019. Disponível em: https://heinonline.org/HOL/Page?handle=hein.journals/flr88&div=20&g_sent=1&casa_token=ztkiTkqt_tcAAAAA:FGuBDxZzkok3Nav8kByNTQd5osMD-mUzTDNfzAJ2v3dCk7BSmwdaoSu0m59SqarnAn-ZMwlydQ&collection=journals#. Acesso em: 10 mar. 2023.

KAYA, Orçun. Artificial intelligence in banking: a lever for profitability with limited implementation to date. **Deutsche Bank Research**. Frankfurt, p. 1-9, 2019. Disponível em: https://www.dbresearch.com/PROD/RPS_EN-PROD/PROD000000000495172/Artificial%20intelligence%20in%20banking%3A%20A%20lever%20for%20profitability%20with%20limited%20implementation%20to%20date.pdf?undefined&reaload=~JutPeH5W~WUW00H0Md53s2AX7Nend4xSYPIPrQZ68pK0jj3Znu0bGwXcpe9WBII. Acesso em:

LAMB, D. The uses of analysis: rhetorical analysis, article analysis, and the literature review. **Academic writing tutor**, 2013.

LEE, Joseph. Access to finance for artificial intelligence regulation in the financial services industry. **European Business Organization Law Review**, [s. l.], v. 21, p. 731-757, 2020. Disponível em: <https://link.springer.com/article/10.1007/s40804-020-00200-0>. Acesso em: 13 mar. 2023.

LINDE, K.; WILLICH, S. N. How objective are systematic reviews? differences between reviews on complementary medicine. **Journal of the royal society of medicine**, [S. l.], v. 96, p. 17–22, 2003.

LUDERMIR, Teresa Bernarda. Inteligência Artificial e Aprendizado de Máquina: estado atual e tendências. **Estudos Avançados**, Recife, v. 35, p. 85-94, 2021. Disponível em: <https://www.scielo.br/j/ea/a/wXBdv8yHBV9xHz8qG5RCgZd/?format=pdf&lang=pt>. Acesso em: 24 fev. 2023.

MAHDAVI, Mehrzad; KAZEMI, Hossein. It's All About Data: How to Make Good Decisions in a World Awash with Information. **The Journal of Financial Data Science**, [s. l.], v. 2, n. 2, p. 8-16, 2020. Disponível em: <https://jfds.pm-research.com/content/2/2/8.short>. Acesso em: 10 mar. 2023.

MARTINS, José Eduardo Figueiredo de Andrade. Dilemas Éticos e Jurídicos do Uso da Inteligência Artificial na Prática Jurídica. **Centro de Investigação de Direito Privado**, [s. l.], v. 7, p. 919-952, 2021. Disponível em: https://www.cidp.pt/revistas/rjlb/2021/4/2021_04_0919_0952.pdf. Acesso em: 10 mar. 2023.

MCCARTHY, John. What is artificial intelligence. [Entrevista cedida a] Computer Science Department. Stanford: Stanford University, 12 nov. 2007. Disponível em: <https://www.diochnos.com/about/McCarthyWhatisAI.pdf>. Acesso em: 18 fev. 2023.

MCCULLOCH, W.S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **Bulletin of Mathematical Biophysics**, [S. l.], v. 5, p. 115-33, 1943. DOI <https://doi.org/10.1007/BF02478259>.

MIKALEF, Patrick; GRUPTA, Manjul. Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. **Information & Management**, v. 58, n. 3, p. 1-20, 2021. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0378720621000082>. Acesso em:

MISHKIN, Frederic S. **The Economics of Money, Banking and Financial Markets**. 1. ed. Edinburgh Gate: Pearson, 2012. *E-book*. Disponível em: <https://unitimesofficial.files.wordpress.com/2020/06/the-economics-of-money-banking-and-financial-markets-11th-edition-global-edition.pdf>. Acesso em: 12 mar. 2023.

MOGAJI, Emmanuel; NGUYEN, Nguyen Phong. Managers' understanding of artificial intelligence in relation to marketing financial services: insights from a cross-country study. **International Journal of Bank Marketing**, [s. l.] v. 40, n. 6, p. 1272-1298, 2022. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/IJBM-09-2021-0440/full/pdf?title=managers-understanding-of-artificial-intelligence-in-relation-to-marketing-financial-services-insights-from-a-cross-country-study>. Acesso em: 16 mar. 2023.

MOMO, Fernanda da Silva *et al.* Gestão do conhecimento na área de sistemas de informações: uma revisão no contexto brasileiro. **Perspectivas em Gestão & Conhecimento**, João Pessoa, v. 9, n. 1, p. 138-153, 2019. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/112419>. Acesso em: 21 mar. 2023.

MONARD, Maria Carolina; BARANAUKAS, José Augusto. **Aplicações de Inteligência Artificial: Uma Visão Geral**. São Paulo, Instituto de Ciências Matemáticas e de Computação de São Carlos, 2000. Disponível em: <https://dcm.ffclrp.usp.br/~augusto/publications/2000-laptec.pdf>. Acesso em:

MOURA, Luzia Menegotto Frick de. **Blockchain: perspectiva tecnológica para a administração pública no Brasil**. 2018. Trabalho de Conclusão (Bacharel em Administração) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2018. Disponível em: <https://www.lume.ufrgs.br/bitstream/handle/10183/185021/001078700.pdf?sequence=1&isAllowed=y>. Acesso em: 15 mar. 2023.

Mourelle, D. (2019). Amazon contra Microsoft: las tecnológicas entran en la industria de defensa. *In: El Orden Mundial*. [S. l.], 27 nov. 2019. Disponível em: <https://elordenmundial.com/amazon-microsoft-tecnologicas-industria-de-defensa/>. Acesso em:

NGUYEN, Duc Khuong; SERMPINIS, Georgios; STASINAKIS, Charalampos. Big data, artificial intelligence and machine learning: A transformative symbiosis in favour of financial technology. **European Financial Management**, [s. l.], v. 29, p. 517–548, 2023. Disponível em: <https://onlinelibrary.wiley.com/doi/epdf/10.1111/eufm.12365>. Acesso em: 14 mar. 2023.

NIELSEN, Richard P. Ethical and Political-Economic Dimensions and Potential Reforms of the Hybrid Leveraged, High Frequency, Artificial Intelligence Trading Model. **Business and Professional Ethics Journal**, [s. l.], v. 40, n. 2, p. 189-222, 2021. Disponível em: https://www.pdcnet.org/bpej/content/bpej_2021_0040_0002_0189_0222. Acesso em: 09 mar. 2023.

OLIVEIRA, A. B. S. **Métodos da pesquisa contábil**. 1. ed. São Paulo: Atlas, 2011.

PIOTROWSKI, Dariusz. Consumer perceived ethicality of banks in the era of digitalisation: The case of Poland. **Economics and Business Review**, [s. l.], v. 8, n. 1, p. 90-114, 2022. Disponível em: <https://sciendo.com/es/article/10.18559/ebr.2022.1.6>. Acesso em: 17 mar. 2023.

PRENTICE, C., NGUYEN, M. Engaging and retaining customers with AI and employee service. **Journal of Retailing and Consumer Services**, [S. l.], v. 56, n. 4, 2020. Disponível em: https://www.researchgate.net/publication/342355277_Engaging_and_retaining_customers_with_AI_and_employee_service. Acesso em:

PRUNKL, Carina. E. A. *et al.* Institucionalizar la etica en la IA atraves de requisitos de impacto mas amplios. **Nat Mach Intell**, [S. l.], v. 3, p. 104–110. 2021.

PWC Financial Services. Smart Money: AI transitions from fad to future of institutional investing. **PWC Financial Services**, November 2018. Disponível em: https://www.fa-mag.com/userfiles/ads_2019/Wavelength_WP_Oct_2019/PwC-FSI-WHITEpaper-Artificial-Intelligence-Investing.pdf. Acesso em:

RICH, E.; KNIGHT, K. **Artificial intelligence**. 2.ed. New York: McGraw-Hill, 1991.

RIZINSKI, Maryan *et al.* Ethically Responsible Machine Learning in Fintech. **IEEE Access**, [s. l.], v. 10, p. 97531-97554, 2022. Disponível em: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9869843>. Acesso em: 12 mar. 2023.

RUSSELL, S. J.; NORVIG, P. **Artificial Intelligence: A Modern Approach** 3. ed. [S. l.]: Pearson Education, 2010.

RUSSELL, S., HAUERT, S., ALTMAN, R., VELOSO, M. Ethics of artificial intelligence. **Nature**, [S. l.], v. 521, p. 415-416, 2015. Disponível em: <https://www.nature.com/articles/521415a>. Acesso em:

SAMPAIO, R.F.; MANCINI, M.C. Estudos de revisão sistemática: um guia para síntese criteriosa da evidência científica. **Revista Brasileira de Fisioterapia**, São Carlos, v. 11, n. 1, 2007, p. 83-89.

SCHWAB, Klaus. **The fourth industrial revolution**. Geneva: World Economic Forum, 2016. Disponível em: <https://www.weforum.org/about/the-fourth-industrial-revolution-by-klaus-schwab>. Acesso em:

SHAIKH, N.. The Financial Industry Needs to Start Planning for the Next 50 Years, Not the Next Five. *In: Harvard Business Review*. [S. l.], 17, jul. 2017. Disponível em: <https://hbr.org/2017/07/the-financial-industry-needs-to-start-planning-for-the-next-50-years-not-the-next-five>. Acesso em:

SHANTHI, R. & PAVITHRA, R. A study relating to the application and use of artificial intelligence in the banking sector. **International Journal of Advanced Science and Technology**, v. 29, n. 3, p. 797-802, 2020. Disponível em: <http://sersc.org/journals/index.php/IJAST/article/view/5897>. Acesso em:

SHAW, J. Artificial Intelligence and Ethics: Ethics and the dawn of decision-making machines. *In: Harvard Magazine*. [S. l.], jan-fev, 2019. Disponível em: <https://www.harvardmagazine.com/2019/01/artificial-intelligence-limitations>. Acesso em:

SICHMAN, Jaime Simão. Inteligência Artificial e sociedade: avanços e riscos. **Estudos Avançados**, [s. l.], v. 35, p. 37-50, 2021. Disponível em: <https://www.scielo.br/j/ea/a/c4sqqrthGMS3ngdBhGWtKhh/?format=html>. Acesso em: 22 fev. 2023.

SILVA, Amanda. **O uso de inteligência artificial no setor bancário e a experiência do consumidor**. 2020. Dissertação (Mestrado em Marketing) – Faculdade de Economia, Universidade do Porto, Porto, 2020. Disponível em: https://sigarra.up.pt/fep/pt/pub_geral.show_file?pi_doc_id=263764. Acesso em: 7 mar. 2023.

SONG, Chengyun *et al.* User abnormal behavior recommendation via multilayer network. **PLoS One**, China, v. 14, n. 12, p. 1-17, 2019. Disponível em: <https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0224684&type=printable>. Acesso em: 08 mar. 2023.

THE CALL. *In: Ren Aissance Foundation*. [S. l.], [2023?]. Disponível em: <https://www.romecall.org/the-call/>. Acesso em: 20 jan. 2023.

TIDJON, Lionel Nganyewou; KHOMH, Foutse. The Different Faces of AI Ethics Across the World: A Principle-To-Practice Gap Analysis. **IEEE Transactions on Artificial Intelligence**, [s. l.], v. 3, n. 6, p. 1-20, 2022. Disponível em: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9964285>. Acesso em: 13 mar. 2023.

VALDERRAMAS, Edgard. A ética como um dos desafios da Inteligência Artificial. In: **Anais do V Workshop sobre Aspectos Sociais, Humanos e Econômicos de Software**. SBC, 2020. p. 71-80. Disponível em: <https://sol.sbc.org.br/index.php/washes/article/view/11199>. Acesso em 21 fev. 2023.

VASSALI-HENRIQUE, Henrique; JANISSEK-MUNIZ, Raquel. O lado “sombrio” da inteligência artificial: uma revisão sistemática da literatura. XLVI Encontro da ANPAD - EnANPAD 2022 On-line - 21 - 23 de set de 2022 - 2177-2576 versão online. 16 p. Disponível em: <http://anpad.com.br/uploads/articles/120/approved/4639475d6782a08c1e964f9a4329a254.pdf>. Acesso em: 21 fev. 2023.

VIEIRA, Leonardo Marques. A problemática da inteligência artificial e dos vieses algorítmicos: caso COMPAS. In: Brazilian Technology Symposium, 2019, São Paulo. **Anais eletrônicos [...]**. São Paulo: Universidade Presbiteriana Mackenzie, 2019. Disponível em: <https://lcv.fee.unicamp.br/images/BTSym-19/Papers/090.pdf>. Acesso em: 18 fev. 2023.

WEBSTER, J., WATSON, R. T. Analyzing the past to prepare for the future: writing a literature review. **MIS Quarterly**, [s. l.], v. 26, n. 2, p. 13-23, 2002.