UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS
DEPARTAMENTO DE CIÊNCIAS CONTÁBEIS E ATUARIAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS CONTÁBEIS

GUSTAVO HENRIQUE COSTA SOUZA

ANÁLISE DA RELAÇÃO ENTRE A TRANSPARÊNCIA DA INTELIGÊNCIA ARTIFICIAL E A TOMADA DE DECISÕES GERENCIAIS

RECIFE
2023

GUSTAVO HENRIQUE COSTA SOUZA

ANÁLISE DA RELAÇÃO ENTRE A TRANSPARÊNCIA DA INTELIGÊNCIA ARTIFICIAL E A TOMADA DE DECISÕES GERENCIAIS

Tese apresentada ao Programa de Pós-Graduação em Ciências Contábeis da Universidade Federal de Pernambuco, como requisito para obtenção do título de Doutor em Ciências Contábeis.

Orientador: Prof. Cláudio de Araújo Wanderley, Ph.D.

Coorientador: Prof. Dr. Andson Braga de Aguiar.

RECIFE
2023

Catalogação na Fonte
Bibliotecária Ângela de Fátima Correia Simões, CRB4-773

S729a Souza, Gustavo Henrique Costa
Análise da relação entre a transparência da inteligência artificial e a tomada de decisões gerenciais / Gustavo Henrique Costa Souza. – 2023.
116 folhas: il. 30 cm.

Orientador: Prof. Dr. Cláudio de Araújo Wanderley, Ph.D. e Coorientador: Prof. Dr. Andson Braga de Aguiar.
Tese (Doutorado em Ciências Contábeis) – Universidade Federal de Pernambuco, CCSA, 2023.
Inclui referências e apêndices.

1. Inteligência artificial. 2. Transparência de tecnologia. 3. Contabilidade gerencial. I. Wanderley, Cláudio de Araújo (Orientador). II. Aguiar, Andson Braga de (Coorientador). III. Título.

657 CDD (22. ed.) UFPE (CSA 2023– 065)

GUSTAVO HENRIQUE COSTA SOUZA

ANÁLISE DA RELAÇÃO ENTRE A TRANSPARÊNCIA DA INTELIGÊNCIA ARTIFICIAL E A TOMADA DE DECISÕES GERENCIAIS

Tese submetida ao Corpo Docente do Programa de Pós-Graduação em Ciências Contábeis da Universidade Federal de Pernambuco.

Aprovada em: 19/06/2023.

BANCA EXAMINADORA:

_____
Prof. Cláudio de Araújo Wanderley, Ph.D. (Orientador)
Universidade Federal de Pernambuco

_____
Prof. Dr. Vinícius Gomes Martins (Examinador Interno)
Universidade Federal de Pernambuco

_____
Prof. Dr. Giuseppe Trevisan Cruz (Examinador Interno)
Universidade Federal de Pernambuco

_____
Prof. Dr. Edgard Bruno Cornachione Junior, Ph.D. (Examinador Externo)
Universidade de São Paulo

_____
Prof. Dr. José Carlos Tiomatsu Oyadomari (Examinador Externo)
Universidade Mackenzie

À Virgem de Fátima.
Dedico.

AGRADECIMENTOS

A Cristo Rei e Sua Mãe, a Bem-Aventurada Virgem Maria.

A minha esposa, Elayne, pelo suporte, encorajamento e amor que dia após dia me tem dado.

Aos meus Pais, apoiadores incondicionais de todos os meus projetos, a quem quero e devo cada vez mais dar orgulho.

Ao meu orientador, Prof. Cláudio de Araújo Wanderley, Ph.D., pelo fundamental suporte que me deu em todas as etapas do desenvolvimento desta pesquisa, e a quem aprendi a admirar como profissional e como estudioso.

Ao meu coorientador, Prof. Dr. Andson Braga de Aguiar, que foi uma verdadeira bússola a me guiar nos caminhos da metodologia experimental, e cujo tranquilidade e segurança representaram para mim um indescritível apoio.

Ao Prof. Flávio da Cunha Rezende, Ph.D., do Departamento de Ciências Políticas, que me fez ver a produção de conhecimento científico a partir de um prisma totalmente novo: suas aulas foram os minutos mais empolgantes de todo este processo de formação que chamamos de doutorado!

Aos amigos, mestres e doutores na área de tecnologia, que muitas vezes sanaram minhas dúvidas e me ajudaram a desbravar esse universo novo no qual resolvi me embrenhar: Claudemir Pacheco, João Tiago, Maicon Herverton, Glauco Martins, Henrique Pereira e Marcos Miguel.

Ao ilustre corpo docente do PPGCC/UFPE, em especial aos professores Vinícius Martins e Luiz dos Anjos: as valorosas contribuições que me deram marcaram profundamente o texto final da minha tese.

A todos os servidores, terceirizados e estagiários da Secretaria do PPGCC/UFPE, que com simplicidade, discrição e receptividade, acolheram minhas demandas e cuidaram dos processos burocráticos com os quais me deparei ao longo do curso.

Ao Banco de Brasil, instituição onde trabalho e que desde o primeiro momento, independente do tema da minha pesquisa, me proporcionou tempo e recursos para conduzir com tranquilidade e determinação os meus estudos.

Aos meus nobres colegas de turma: Vanessa Janiszewski, Paulo Leal, Leandro Lopes, Sheila Kataoka e Marcelo Ribeiro, os quais – cada um a seu modo – contribuíram grandemente para o meu aprendizado e me proporcionaram um convívio inesquecivelmente prazeroso.

Aos meus colegas de trabalho, com os quais compartilhei as alegrias e as tristezas desta empreitada, e que sempre acreditaram – mais que eu próprio – nas minhas capacidades.

A todos e todas que, direta ou indiretamente, me auxiliaram nesta caminhada e cujos nomes, embora não citados aqui, estão (e estarão eternamente) gravados na minha memória e tatuados no meu coração: MUITO OBRIGADO!

“É justo que muito custe o que muito vale.”
(Santa Teresa d’Ávila)

“O heroísmo do trabalho está em acabar cada tarefa.”
(São Josemaría Escrivá)

RESUMO

A tecnologia, notadamente aquela relativa às ferramentas e modelos de inteligência artificial (IA), tem afetado e até redefinido o papel dos seres humanos no tocante ao processo decisório das organizações. Neste contexto, para um(a) gestor(a) abdicar das próprias impressões e experiências pessoais, e delegar a um modelo de IA a decisão que antes lhe cabia, é necessário que este modelo seja transparente e adequado à referida decisão. Dados os diferentes graus de transparência da IA e os distintos tipos de decisão gerencial, a presente tese doutoral propõe a seguinte questão de pesquisa: **a aderência dos gestores às recomendações de um modelo de IA é afetada pela transparência do referido modelo e pelo tipo de decisão envolvida?** A pesquisa adiciona argumentos à literatura sobre a chamada *Human-Computer Interaction* e, indiretamente, agrega nuances argumentativas que reforçam a discussão sobre a confiança na inteligência artificial. Para responder à questão de pesquisa, realizou-se um *Artefactual Field Experiment* com *design* experimental do tipo *between-participants* 2 x 2. As variáveis independentes foram a transparência (manipulada em dois níveis: alta *versus* baixa) e o tipo de decisão (também manipulada em dois níveis: operacional *versus* estratégica). A variável dependente foi a percepção dos gestores – manifesta no grau de aderência deles às recomendações feitas pelo modelo de IA do *case*. Utilizando o *software* Survey Monkey, aplicou-se a pesquisa a funcionários da área tática de uma instituição financeira no Brasil. Todos os participantes tinham participação e/ou poder de gestão, e foram aleatoriamente alocados entre os grupos que representaram as quatro condições experimentais. Obtiveram-se 102 respostas válidas. A análise de covariância (ANCOVA) das respostas obtidas permitiu concluir que a relação entre Transparência e Aderência é de natureza inversa – e não direta – isto é: mais Transparência implica menos Aderência. Assim, o efeito principal da Transparência sobre a Aderência, embora existente, aponta para uma direção oposta à da previsão teórica formulada. Além disso, verificou-se que Decisão não é capaz, por si só, de alterar de forma estatisticamente significativa a percepção dos gestores – e, consequentemente, não provoca uma modificação substantiva na sua Aderência. Por fim, ao examinar a interação entre as variáveis Transparência e Decisão, constatou-se que esta interação, de fato, ocorre e afeta de forma significativa a Aderência dos gestores às recomendações do sistema de inteligência artificial. Entretanto, este efeito apresentou sinal contrário ao da predição teórica realizada. Assim, não é possível afirmar que o efeito positivo da alta transparência sobre a Aderência é maior em decisões operacionais que em decisões estratégicas. Estes achados têm uma tripla implicação: primeiramente, contribuem para os estudos sobre *Explainable Artificial Intelligence*, aprimorando o entendimento sobre o impacto da transparência da IA no processo decisório; em segundo lugar, fornecem um subsídio teórico sobre como, no *design* de *recommender systems,* o tipo de decisão interage com a transparência da IA para orientar as decisões dos gestores; e, por fim, oferecem um *insight* sobre como o papel da confiança na tecnologia afeta a aderência dos gestores às recomendações feitas por modelos de inteligência artificial.

Palavras-chave: Inteligência artificial; Transparência; Decisão; Contabilidade gerencial; Experimento.

ABSTRACT

Technology, notably that related to AI tools and models, has affected and even redefined the role of human beings in the decision-making process of organizations. In this context, for a manager to give up his/her own personal impressions and experiences, and delegate to an AI model the decision that previously belonged to him/her, it is necessary that this model be transparent and adequate to that decision. Given the different degrees of transparency of AI and the different types of managerial decision, this doctoral thesis proposes the following research question: is managers' adherence to the recommendations of an AI model affected by the transparency of that model and by the type of decision involved? The research adds arguments to the literature on the so-called Human-Computer Interaction and, indirectly, adds argumentative nuances that reinforce the discussion about trust in artificial intelligence. To answer the research question, an Artefactual Field Experiment was carried out with a 2 x 2 between-participants experimental design. The independent variables were transparency (manipulated in two levels: high versus low) and the type of decision (also manipulated in two levels: operational versus strategic). The dependent variable was the managers' perception – manifested in their degree of adherence to the recommendations made by the case's AI model. Using Survey Monkey software, the survey was applied to employees in the tactical area of a financial institution in Brazil. All participants had participation and/or management power, and were randomly allocated among groups that represented the four experimental conditions. 102 valid responses were obtained. Analysis of covariance (ANCOVA) of the responses obtained allowed us to conclude that the relationship between Transparency and Adherence is of an inverse nature – and not direct – that is: more Transparency implies less Adherence. Thus, the main effect of Transparency on Adherence, although existing, points in the opposite direction to the formulated theoretical prediction. In addition, it was found that Decision is not able, by itself, to change in a statistically significant way the perception of managers – and, consequently, does not cause a substantive change in their Adherence. Finally, when examining the interaction between the variables Transparency and Decision, it was found that this interaction, in fact, significantly affects the Adherence of managers to the recommendations of the artificial intelligence system. This effect, however, showed the opposite sign to the theoretical prediction made. Thus, it is not possible to state that the positive effect of high transparency on Adherence is greater in operational decisions than in strategic decisions. These findings have a triple implication: first, they contribute to studies on Explainable Artificial Intelligence, improving understanding of the impact of AI transparency on the decision-making process; secondly, they provide a theoretical subsidy on how, in the design of recommender systems, the type of decision interacts with the transparency of AI to guide managers' decisions; and finally, they offer insight into the role of trust in technology in affecting managers' adherence to recommendations made by artificial intelligence models.

Keywords: Artificial intelligence; Transparency; Decision; Management accounting; Experiment.

LISTA DE FIGURAS

LISTA DE GRÁFICOS

LISTA DE QUADROS

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

CG – Contabilidade Gerencial

CGMA – Chartered Global of Management Accountant

FEBRABAN – Federação Brasileiros de Bancos

HCI – *Human Computer Interaction*

IA – Inteligência Artificial

IoT – *Internet of Things*

TCLE – Termo de Consentimento Livre e Esclarecido

XAI – *Explainable Artificial Intelligence*

VD – Variável Dependente

VI – Variável Independente

SUMÁRIO

1 INTRODUÇÃO

1.1 CONTEXTO E QUESTÃO DE PESQUISA

O xadrez e o pôquer requerem dos seus jogadores decisões. Decisões ponderadas, que visam antecipar as próximas jogadas do adversário; decisões cautelosas – para que delas não resultem derrotas; e decisões inovadoras, de preferência, para que o adversário seja tomado de surpresa. No tabuleiro e na mesa, as cartas e as peças se movimentam em um duelo silencioso, onde informações (e até o blefe!) são geradas, analisadas e confrontadas continuamente.

A diferença, porém, entre os dois jogos, é que no xadrez as peças estão à mostra. O jogador e o seu adversário veem as mesmas peças, enxergam perfeitamente cada movimento do outro porque tudo no tabuleiro é feito às claras. O xadrez é transparente. Já no pôquer, o jogo acontece nas mãos de cada jogador (o qual deve se esmerar em esconder aos demais participantes suas cartas e, sobretudo, sua estratégia). As cartas que estão à mesa ou não são conhecidas ou são itens de descarte. Isto significa dizer que, no pôquer, a regra é a ocultação, o desconhecimento, a penumbra – e não a transparência.

A metáfora do xadrez e do pôquer ilustra e instiga a reflexão sobre como o elemento transparência influencia as decisões humanas. Se no xadrez cada jogador visse apenas as suas peças (ou, quiçá, nem isso) e ainda assim tivesse que realizar os movimentos, como se portaria? Que parâmetros ele iria utilizar para tomar suas decisões? E se no pôquer tudo fosse revelado? Que consequências isso traria à dinâmica do jogo? Caso se tratasse da rodada inicial ou da rodada final os jogadores decidiriam da mesma forma (considerando que na inicial há espaço para correção de eventuais erros e na final já não há essa possibilidade)? Estas alterações hipotéticas nas regras de transparência/ocultação afetariam a dinâmica (e a própria natureza) dos jogos e forçaria os jogadores a rever suas estratégias.

Assim como para os jogadores, o papel da transparência e de sua influência sobre as decisões também é um ponto sensível para os diversos atores do universo empresarial – contadores, administradores, *controllers* etc. –, os quais constroem seus posicionamentos e consolidam suas decisões ante a abundância/carência de informações (MEYER, 1998). A transparência e a ocultação de informações também podem afetar aqueles que decidem as "jogadas" no âmbito organizacional e, em decorrência disso, o próprio "jogo negocial" pode sofrer alterações. Por isso, é oportuno compreender como os dilemas relativos à transparência estão configurados no mundo atual e como isso repercute nas organizações, nos profissionais e nas decisões destes.

Na era digital, a tecnologia – notadamente aquela relativa às ferramentas e modelos de inteligência artificial (IA) – tem influenciado e até redefinido o papel dos gestores no tocante ao processo decisório das organizações (ATHEY; BRYAN; GANS, 2020; BRANDS; HOLTZBLATT, 2015). Os diversos efeitos da transparência da IA sobre as decisões dos gestores têm sido abordados recorrentemente pela recente pesquisa acadêmica (ALUFAISAN *et al.*, 2021; COECKELBERGH, 2020; SCHEMMER *et al.*, 2022). Transparência, antes de mais nada, aqui entendida como a propriedade que permite ao usuário de um sistema compreender a lógica interna e o processo de raciocínio do referido sistema. (CHIEN *et al.*, 2014)

Na área contábil há quem enfatize a utilização da IA no processamento dos dados e afirme que "diferentes ângulos sobre pesquisas futuras poderiam revolucionar a aplicação da IA na contabilidade" (LOSBICHLER; LEHNER, 2021, p. 365). Um destes ângulos, o qual é abordado na presente tese, é o da colaboração homem-máquina – que, entre outras coisas, visa conciliar o papel do ser humano e da IA para que, desta interação, fluam as decisões informações e decisões (BANSAL *et al.*, 2019; DIEDERICH *et al.*, 2022; REN; BAO, 2020; SCHEMMER *et al.*, 2022; WILSON; DAUGHERTY, 2018).

Os impactos desta inserção da IA na contabilidade, é claro, têm sido examinados e avaliados a partir das mais diversas tônicas e acentos: desde uma perspectiva pragmática e focada que busca apresentar "tarefas e registros contábeis que podem ser delegados a IA" (PETKOV, 2020, p. 99) até uma visão mais ampla e genérica que busca mostrar o "potencial disruptivo e transformador" desta nova tecnologia (MARRONE; HAZELTON, 2019, p. 677).

Em contabilidade gerencial (CG), especificamente, muito da atual discussão gira em torno das transformações que a IA provoca/provocará sobre o papel dos contadores gerenciais (BRANDS; HOLTZBLATT, 2015; CGMA, 2019; LAWSON, 2019), sobre como ela pode afetar as práticas e métodos da CG (DING *et al*., 2020; PETKOV, 2020), e sobre as próprias modificações que as organizações experimentam quando introduzem a IA nos seus processos e atividades (NIELSEN, 2020; SHI, 2019). Enquanto uns destacam vantagens e melhorias que a IA proporciona (DING *et al*., 2020; SHI, 2019), outros apontam seus desafios e limitações (GOTTHARDT *et al.*, 2020; LEHNER *et al*., 2022; LOSBICHLER; LEHNER, 2021).

Uma destas supostas limitações da IA seria a capacidade de decidir. Em tese, apenas os humanos teriam a habilidade de sopesar interesses, ponderar entre custos e benefícios, e avaliar circunstâncias e conveniências para, então, formular e tomar uma decisão (PHILLIPS-WREN, 2012). À medida que se desenvolvem, porém, os modelos de inteligência artificial confrontam esta premissa e, adentrando à esfera da contabilidade gerencial, põem à disposição algoritmos

que oferecem diagnósticos, predições e prescrições cada vez mais eficazes que aqueles fornecidos por contadores, *controllers* e gestores (APPELBAUM *et al.*, 2017). Por meio da IA, decisões consideradas complexas são objetivadas, esquematizadas e simplificadas (WEIMEI, 2021). Com a IA, decisões costumeiramente morosas ganham celeridade com apenas um clique ou comando (SHRESTHA; BEN-MENAHEM; VON KROGH, 2019; TRUNK, BIRKEL; HARTMANN, 2020). Erros e inconsistências que antes confundiam a análise dos decisores (e que às vezes passavam desapercebidamente) agora são mapeados e filtrados por modelos de inteligência artificial capazes de tratar milhões de dados em segundos (LI; ZHENG, 2018) – algo que, até que se prove o contrário, os humanos não são capazes de realizar por si sós.

Assim, a colaboração da IA para a tomada de decisões gerenciais contábeis cresce dia após dia, consolidando seu papel no processo decisório das organizações, ao mesmo tempo em que se depara com a desconfiança e a falta de compreensão dos gestores sobre seus métodos e critérios e sobre a sua assertividade e aplicabilidade a determinadas decisões e contextos específicos (VĂRZARU, 2022).

Em um cômputo geral, remanescem dúvidas sobre se este suporte advindo da IA deixa os humanos suficientemente seguros para tomar decisões com base nas informações e sugestões geradas pela IA (LARKIN; DRUMMOND; ÁRVAI, 2022). Também não se sabe ao certo se o grau de transparência da IA e a natureza da decisão envolvida atuariam, conjunta ou isoladamente, para influenciar de maneira significativa a percepção, a análise e o posicionamento dos decisores (FELZMANN *et al*., 2020; HANNA; LEMON; SMITH, 2019).

Assim, a presente tese doutoral propõe a seguinte questão de pesquisa: **a aderência dos gestores às recomendações de um modelo de IA é afetada pela transparência do referido modelo e pelo tipo de decisão envolvida?**

1.2 OBJETIVOS

O problema que norteia a presente pesquisa é comumente chamado *black-box problem* (ADADI; BERRADA, 2018; CASTELVECCHI, 2016; LOYOLA-GONZALEZ, 2019; VON ESCHENBACH, 2021; ZEDNIK, 2021). Ele emerge de uma dualidade paradoxal que consiste no seguinte: **i)** há modelos de inteligência artificial cujo funcionamento é mais difícil de ser explicado e compreendido, mas que apresentam altos níveis de acurácia. Isto é: são sistemas ditos opacos, nebulosos, menos transparentes, mas que produzem *outputs* de grande precisão; **ii)** há modelos de inteligência artificial menos complexos, cuja lógica interna é fácil de se explicar e de se entender, mas cujos resultados são menos precisos, carecem de refinamento. O

confronto desta dupla constatação constitui uma intrigante realidade que se desdobra em diversas frentes de pesquisa sobre os efeitos que o nível de transparência da IA produz ou pode produzir.

Além disso, a literatura destaca que as decisões não são todas iguais (nem no que diz respeito ao seu objeto, nem em termos de abrangência, tampouco em termos de risco) (HARRISON; PELLETIER, 1995; MISNI; LEE, 2017; NOORAIE, 2008) e, portanto, ao examinar fatores que supostamente têm efeitos sobre o processo decisório, convém avaliar separadamente decisões de naturezas distintas.

Adicionalmente, ao estudar o efeito de diferentes graus de transparência sobre diferentes tipos de decisão, torna-se imperativo examinar o confronto destes aspectos em vista da possibilidade de uma interação entre eles. Assim, o objetivo geral desta pesquisa é de **investigar se a aderência dos gestores às recomendações de um modelo de IA é afetada pela transparência do referido modelo e pelo tipo de decisão envolvida.**

Para alcançar tal objetivo foram constituídos os seguintes objetivos específicos:

- Verificar o efeito da transparência do modelo (alta *versus* baixa) de IA sobre a aderência dos gestores às recomendações do referido modelo.
- Verificar o efeito do tipo de decisão (operacional *versus* estratégica) sobre a aderência dos gestores às recomendações do modelo de IA.
- Testar a aderência dos gestores às recomendações do modelo de IA nas diferentes interações entre transparência do modelo (alta *versus* baixa) e tipo de decisão (operacional *versus* estratégica).

1.3 JUSTIFICATIVA

A dicotomia entre ver-se diante do aprimoramento da tecnologia, que assume os papéis e tarefas dos humanos sem pedir licença, e – ao mesmo tempo – sentir-se diante de uma espécie de "caixa-preta" (quase impenetrável e de difícil compreensão, como são alguns modelos de IA) é uma tensão bastante acentuada e ainda carente de exploração científica (ALBUQUERQUE *et al.* 2019; FREY; OSBORNE, 2017; TAYLOR, 2016).

As opções de aplicação da IA contam hoje com o suporte de elementos que, décadas atrás, eram inexistentes ou embrionários: *internet*, redes sociais, dispositivos de

reconhecimento de voz, além da própria evolução das ferramentas de *cognitive computing*[1] e da ideia de *machine learning*[2]. Tais elementos situam a pesquisa ora proposta em um universo distinto daquele em que se desenvolveram os estudos anteriores (BRADY, 1984; FETZER, 1990; HORVITZ; BREESE; HENRION, 1988).

Além disso, abdicar das próprias impressões e experiências pessoais para delegar a um modelo de IA um papel que antes cabia a um ser humano exige que este modelo (e o produto final que ele oferece) inspire confiança em quem delega (MCKNIGHT *et al.*, 2011). Por outro lado, se a lógica subjacente ao modelo não for suficientemente transparente para que seus usuários o entendam e avaliem, a construção desta confiança fica prejudicada (CHIEN *et al*., 2014; MADSEN; GREGOR, 2000; YAGODA; GILLAN, 2012). Assim, entender que a relação entre os diferentes graus de transparência da IA e os distintos tipos de decisão é permeada pela confiança abre espaço para que se examine o papel desta última nessa relação e seu eventual impacto sobre a aderência dos gestores às recomendações de um modelo de IA.

Nos meios profissionais, o conflito entre apropriar-se das ferramentas de inteligência artificial e o temor de ser por elas substituído gera inquietações recorrentes sobre quais ocupações são/serão suficientemente resistentes (ou resilientes) a ponto de sobreviver ao avanço tecnológico que o século XXI experimenta (TAYLOR, 2016; ZUBOFF, 1988). No fundo, a agitação parece ser a respeito de sob que condições a inteligência humana poderá ser dispensada e suplantada pela inteligência artificial.

Em contabilidade, esta realidade não é diferente: há preocupação sobre até quando tarefas contábeis de ordem repetitiva, burocrática e protocolar (atribuídas aos chamados *bean counters*) continuarão a demandar força humana (BALDVINSDOTTIR *et al.*, 2009; FRIEDMAN; LYNE, 1997; GRANLUND; LUKKA, 1997). Modelos de IA já são capazes de executar muitas destas tarefas – com mais velocidade e menos erros (DING *et al*., 2020; HOFFMAN, 2017). Mesmo entre os contadores que desempenham o papel de *business partners* há temor em lançar mão da tecnologia quando se sabe que a responsabilidade pela gestão e as consequências das decisões continuam a recair sobre os humanos (e não sobre as máquinas e programas) (LEITNER-HANETSEDER *et al.* 2021; MARRONE; HAZELTON, 2019).

---

[1] *Cognitive computing* "visa desenvolver um mecanismo coerente, unificado e universal inspirado nas capacidades da mente. Mais que reunir uma coleção de soluções fragmentadas, em que diferentes processos cognitivos são construídos individualmente por meio de soluções independentes, busca-se implementar um sistema unificado da teoria computacional da mente" (MODHA *et al.*, 2011, p. 62).

[2] *Machine learning* é a área da ciência computacional que lida com "o *design* de programas que podem aprender regras a partir dos dados, adaptar-se às mudanças e melhorar o desempenho com base na experiência" (BLUM, 2007, p. 1).

Até mesmo o papel revisional de auditores e *controllers*, bem como a função analítica e decisória de contadores gerenciais, estão postos em xeque: as modernas técnicas de *business analytics*[3] tornam no mínimo subsidiária a função destes profissionais (LEITNER-HANETSEDER *et al.*, 2021). Uma tecnologia capaz de filtrar erros e inconsistências praticamente imperceptíveis a "olho nu", trabalhar com uma enormidade de dados (prescindindo, inclusive, da técnica de amostragem), encontrar padrões de significado nestes dados, e desenvolver novas soluções para problemas complexos faz com que os tradicionais papéis de gestão em contabilidade sejam diuturnamente contestados e ameaçados (OESTERREICH *et al.*, 2019).

Diante desta transformação, é compreensível que gestores encarregados de tomar decisões contábeis façam suas escolhas não apenas a partir dos *outputs* que lhes são oferecidos pela contabilidade, mas também a partir da relação que eles próprios constroem/têm com a tecnologia que gerou aqueles *outputs* (OESTERREICH *et al.*, 2019). Este debruçar-se sobre a percepção que os decisores têm com relação à tecnologia que lhes dá suporte contribui para que se possa entender melhor o processo de transformação digital que atinge em cheio a contabilidade (CGMA, 2019).

Em alguns casos, como foi o caso do trabalho de Zhang e Nauman (2018), investigações relativas à IA são tratadas a partir de um ponto de vista conceitual. Além disso, a pesquisa em IA por vezes se volta para a lógica da aplicação eficiente de uma ferramenta específica, simulando se determinada máquina é capaz ou não de agir tão eficazmente quanto um humano (ERNEST *et al*., 2016). Trata-se de uma análise do desempenho intrínseco a determinado dispositivo. Diferentemente destes estudos, a presente pesquisa centraliza-se não sobre um dispositivo específico, mas sim sobre o usuário, suas percepções e decisões.

Esta preocupação com a perspectiva do usuário recebeu uma importante contribuição mediante o estudo de Lang (2018). O autor investiga como a utilização e manipulação de dados (estruturados e não-estruturados) gerados a partir de ferramentas de IA afetou o julgamento (e a confiança) dos usuários em tomadas de decisões contábeis. O foco no tipo de dados empregados e nas intervenções humanas sobre eles, porém, fez com que aspectos como a natureza da decisão contábil envolvida e sua interação com a percepção dos usuários sobre a IA fossem deixados de lado – oportunizando, assim, espaço para discutir tais questões na presente tese.

---

[3] "*Business Analytics* é uma aplicação/subconjunto específico dentro de *Analytics,* que potencializa as ferramentas, técnicas e princípios deste último para desenvolver soluções para os mais complexos problemas de negócios" (DELEN; RAM, 2018, p. 2).

Assim, a aplicação da IA à tomada de decisões gerenciais pode transformar a contabilidade como ferramenta de suporte à decisão. Discutir isso, examinando os pormenores relativos à transparência do processo da IA e testando os reflexos relativos à caracterização da decisão envolvida, é a proposta da presente tese.

1.4 CONTRIBUIÇÃO DO ESTUDO

A pesquisa sobre a transparência da inteligência artificial ajuda a entender como a informação (oriunda de diversas fontes, transmitida por vários canais, e "empacotada" de diferentes modos) influencia os processos, as decisões e até a performance das organizações – e dos gestores que nelas atuam. Trata-se de um tema de relevo, frequentemente visitado pela literatura a partir de distintas abordagens (HALL, 2010; LAUDON; LAUDON, 2011; LI; MCLEOD JR; ROGERS, 1993; SCHOLTEN *et al.,* 2007; STANCIU; TINCA, 2017; TAYLOR, 1975; UNGSON; BRAUNSTEIN; HALL, 1981). As discussões relacionadas a este tema, aliás, deram causa a um ramo próprio de investigação científica: a pesquisa sobre *Explainable Artificial Intelligence* (XAI) (ADADI; BERRADA, 2018; ANGELOV *et al.* 2021; ARRIETA *et al.* 2020; DORAN; SCHULZ; BESOLD, 2017; GUNNING *et al.* 2019; PREECE, 2018; SAMEK; MÜLLER, 2019). Diferenciando-se dos demais estudos por examinar a interação entre a transparência da IA e os diferentes tipos de decisão gerencial, a tese ora apresentada evidencia as consequências desta interação sobre os gestores (e seus posicionamentos) e incrementa, assim, o conhecimento da academia com relação a esta matéria.

Como consequência direta desta primeira contribuição, esta pesquisa também adiciona argumentos à literatura sobre os direcionadores da tomada de decisão. O uso de modelos de inteligência artificial para subsidiar e orientar decisões tem sido cada vez mais frequente (LU *et al.*, 2015; NUNES; JANNACH, 2017; VULTUREANU-ALBIŞI; BĂDICĂ, 2021) e, por isso, pesquisadores de diversas áreas têm tratado a IA como uma ferramenta de *decision-aid* (MIN, 2010; SHARMA *et al.*, 2022). Entretanto, não está claro, até o momento, sob que condições características intrínsecas aos modelos de IA (como a transparência) influenciam a percepção dos gestores direcionando efetivamente suas decisões. A presente tese contribui com o esclarecimento desta questão.

Em específico, na esfera da contabilidade gerencial, o que se constata é que decisões classicamente difíceis como comprar *versus* alugar e contratar *versus* terceirizar já podem ser tomadas com o suporte de inteligência artificial (BANERJEE, 2020; GEETHA; BHANU, 2018;

LI *et al.*, 2021). A oferta de uma reflexão sobre em que condições gestores percebem este suporte de IA como sendo, efetivamente, um auxílio para a resolução de dilemas gerenciais contábeis representa uma contribuição significativa, necessária e útil aos que vivenciam tais situações. Ao utilizar participantes que desempenham, de fato, cargos de gestão e de apoio à gestão, o ambiente de pesquisa se aproxima ao máximo daquele onde de fato ocorre/pode ocorrer o fenômeno estudado – o que confere robustez e validade ao contributo prático aqui ofertado.

Por fim, ao abstrair do aspecto técnico-funcional relacionado a um dispositivo ou robô específico e focar na visão dos gestores sobre modelos de IA em geral (aplicados a diferentes tipos de decisão e com diferentes graus de transparência), este trabalho contribui para que melhor se compreenda a percepção dos decisores com relação à confiabilidade dos sistemas de recomendação baseados em IA, além de proporcionar a geração de *insights* sobre a cooperação homem-máquina.

1.5 ESTRUTURA DO TRABALHO

Esta tese encontra-se subdividida em seis capítulos. O primeiro diz respeito a aspectos introdutórios, tais como: a contextualização do problema de pesquisa, o oferecimento de justificativas teóricas e metodológicas para sua realização, a definição dos objetivos do estudo, a explanação sobre a contribuição que o trabalho oferece à comunidade acadêmica e à sociedade em geral, e – por fim – este esclarecimento sobre a própria organização do texto da tese.

O capítulo dois versa de maneira objetiva sobre os temas-chave que sedimentam este estudo, quais sejam: IA aplicada à área de negócios, transparência da IA, e utilização da IA para fins de tomada de decisão gerencial. Toda a parte conceitual relativa a estes tópicos é desenvolvida com base em estudos seminais, artigos publicados em revistas com alto fator de impacto, e demais produções acadêmicas tanto da área de contabilidade quanto da área de tecnologia pura. O capítulo culmina com a apresentação da base teórica sobre a qual se desenvolvem as hipóteses do estudo.

No terceiro capítulo, o *design* da pesquisa e os procedimentos metodológicos são explicados em detalhes. Tratando-se de um experimento, o capítulo traz os procedimentos para seleção dos participantes, os esclarecimentos sobre a aleatorização na alocação dos participantes entre os grupos experimentais, as características do instrumento de coleta de dados, a apresentação das variáveis (dependente, independentes e de controle), e os ditames

protocolares de ordem ética. Além disso, ele traz explicações sobre como os dados coletados foram analisados e exibe os resultados obtidos a partir da análise estatística destes dados.

O capítulo quatro apresenta os resultados dos testes de hipóteses e de outras análises estatísticas a que foram submetidos os dados obtidos. A discussão destes dados e resultados é feita no quinto capítulo, no qual se avalia o impacto dos achados de pesquisa (frente à literatura) na perspectiva de justificar inferências e levantar novos *insights* de pesquisa para trabalhos futuros.

Em seguida, o sexto e último capítulo traz a conclusão do estudo diante de tudo o que foi encontrado e analisado. Nele, a questão de pesquisa é respondida de maneira clara e fundamentada, os objetivos são retomados e avaliados (à luz dos achados), e as hipóteses formuladas são aceitas ou rejeitadas, de acordo com os resultados obtidos. Referências, apêndices e anexos vêm em seguida.

2 REVISÃO DA LITERATURA

2.1 INTELIGÊNCIA ARTIFICIAL NOS NEGÓCIOS

No século XVIII, a primeira revolução industrial se deu pela introdução das máquinas a vapor ao processo produtivo. A inovação tecnológica de então permitiu ganhos de eficiência acima do padrão da época (SCHWAB, 2017). A força gerada a partir do vapor era visível e podia ser fisicamente explicada e facilmente compreendida. Em seguida, no século XIX, a energia elétrica provocou uma nova revolução, ainda mais intensa que a primeira, porque a novidade não estava mais restrita ao ambiente fabril ou às locomotivas: ela havia invadido a casa das pessoas, facilitando rotinas e trazendo conforto (DATHEIN, 2003). No século XX, a terceira revolução industrial foi marcada pelo advento da *internet* e dos computadores pessoais (PAULO, 2019).

Neste caminho que vem sendo traçado e percorrido há pelo menos trezentos anos, a novidade que interpela o século XXI e que justifica a alusão a uma nova revolução industrial (a quarta, portanto) é a integração das diversas tecnologias à vida diária (Schwab, 2017). A "*internet* das coisas" (na sigla, em inglês, *Internet of Things* [IoT]) faz com que dispositivos tecnológicos estejam tão acessíveis e incorporados ao cotidiano que as pessoas e empresas os utilizem até mesmo sem perceber (CONTI *et al.*, 2018; ROSE; ELDRIDGE; CHAPIN, 2015).

À guisa de exemplo, ferramentas de busca, assistentes virtuais, *chatbots* e atendimentos telefônicos automatizados fazem parte da vida de pessoas que podem não fazer ideia do que vem a ser algoritmos e inteligência artificial. É impensável que alguém utilize uma prensa a vapor, uma lâmpada elétrica, ou mesmo um laptop sem se dar conta da existência e da presença do dispositivo tecnológico diante de si. Entretanto, é possível reclamar de um produto adquirido com defeito em um chat sem perceber que, por trás daquele atendimento, não está uma pessoa e sim uma máquina. Ou seja: na era digital, é possível estar perante a tecnologia, fazer uso dela, e ainda assim não notá-la. E mais: as tecnologias adquiriram tal autonomia e desfaçatez que em muitos casos não é necessário nenhum tipo de intervenção, complementação ou validação humana (MICROSOFT, 2022). Não por acaso, um engenheiro da computação (funcionário de uma grande empresa de tecnologia) cogitou, recentemente, que uma dada ferramenta de inteligência artificial tinha ganho vida própria (BBC, 2022).

Na era da chamada "4ª Revolução Industrial" (SCHWAB, 2017) a geração de informações é frequente, intensa e veloz. Embora a sistematização das informações já venha sendo discutida em estudos anteriores, fica cada vez mais nítida a dificuldade humana em lidar

com dados tão volumosos e diversificados (CHEN; ZHANG, 2014). Esta reflexão remete aos questionamentos que o matemático britânico Alan Turing se fez em meados do século passado sobre a possibilidade de se desenvolver uma inteligência artificial (IA) capaz de processar informações e executar tarefas tanto quanto – ou melhor que! – os humanos (TURING, 1950).

A ideia de que máquinas ou, mais especificamente, computadores podem adquirir e desenvolver determinadas habilidades por meio de um processo cognitivo similar ao raciocínio humano (LEGG; HUTTLER, 2007; MINSKY, 1990; SARMAH, 2019; TURING, 1950) é um dos pilares do esforço e do progresso tecnológico atual. Pragmaticamente, "fabricar" inteligência (IKEDA, 2020; KATZ, 2020) tornou-se, na era da revolução digital nos negócios, uma questão de sobrevivência.

O conceito de IA permanece bastante complexo e não consensual (LEGG; HUTTLER, 2007). Apesar desta inexistência de uniformidade conceitual, a discussão teórica desta temática conseguiu avançar consideravelmente. Na presente tese, adotaremos o conceito de inteligência artificial de Neil (2020, p. 6): "habilidade de máquinas tomarem decisões e aprenderem de maneira similar aos humanos [...] refere-se a um sistema computacional que pode cumprir tarefas que normalmente exigiriam a inteligência humana".

Tecnicamente, estes sistemas que mimetizam o agir humano o fazem por meio de diferentes estratégias de *machine learning*, o que permite classificá-los em: sistemas supervisionados, sistemas não supervisionados, sistemas semissupervisionados, sistemas de *reinforcement*, e sistemas de *deep learning* (NIELSEN, 2020). O propósito de utilização dos dados, o tipo de regra empregado no tratamento deles, e o nível de assertividade dos *outputs* permitem que o usuário diferencie e escolha entre estas estratégias a que parecer mais apropriada ao seu banco de dados, às suas possibilidades e conveniências.

O funcionamento, propriamente dito, destes sistemas depende de um conjunto de regras, procedimentos e etapas que constituem o chamado algoritmo (HILL, 2016). Com a IA, os algoritmos deixaram de ser apenas um padrão de ação para o raciocínio das máquinas. Eles se consolidaram como influenciadores do comportamento humano, geradores de (re)interpretações, e até difusores e mantenedores de preconceitos e distorções (ASCARZA; ISRAELI, 2022). O mercado vende algoritmos prontos, adaptáveis a diversos sistemas e propósitos, enquanto prossegue a busca pelo algoritmo perfeito (DENWATTANA; GETTA, 2001; SHAH, 2012).

A competição internacional pelo domínio e aplicação de novas tecnologias tem se intensificado e levado alguns países a investir "pesadamente" em IA (CGMA, 2019).

Executivos destacam a IA como uma das mais importantes ferramentas tecnológicas para as suas companhias (FORBES INSIGHTS, 2017).

A IA pode proporcionar ganhos significativos de produtividade, especialmente para empresas situadas em mercados emergentes (STRUSANI; HOUNGBONON, 2019). Pesquisas destacam o aumento da eficiência dos negócios e processos a partir do emprego da IA. O argumento, segundo alguns, explica que isso está atrelado ao fato de que os modelos e ferramentas de IA são mais ágeis e menos sujeitos a erros que os humanos – além de serem mais versáteis e abrangentes que outros sistemas (CHOWDHURY; SADEK, 2012; JARRAHI, 2018).

O redesenho promovido pela introdução de inovações tecnológicas no mercado gera uma dupla consequência: por parte das organizações, há a expectativa de que a IA otimize a produção, armazenagem, processamento e utilização dos dados (NORDLANDER, 2001; WAMBA-TAGUIMDJE, 2020); e por parte dos contadores gerenciais, há o desafio de reinventar-se profissionalmente para não ser "engolido pela" IA (BRANDS; HOLTZBLATT, 2015).

Entre 2015 e 2019, o Brasil constou como um dos cinco países que mais contrataram pessoal especializado em IA e adquiriram ferramentas de IA (PERRAULT *et al*., 2019). No segmento de instituições financeiras, por exemplo, a inteligência artificial tem estado no *top ranking* das ferramentas tecnológicas que mais recebem investimentos por parte destas entidades (Federação Brasileiros de Bancos [FEBRABAN], 2019). No setor financeiro, aliás, a utilização da IA está bastante centralizada no atendimento a clientes. Previsões dão conta de que em menos de uma década o relacionamento entre instituições financeiras e sociedade tende a ser unicamente digital (SENA, 2020). A máquina, de acordo com tal predição, não seria apenas o meio que viabiliza a interação de dois seres humanos: ela própria passa a ser um polo interativo capaz de construir diálogo e fornecer respostas.

Por isso, dentre outros fatores, a importância da IA decorre do fato de ela ser apontada como um dos *drivers* tecnológicos do processo de mudança de papéis e competências profissionais que a sociedade contemporânea vivencia (LEOPOLD; RATCHEVA; ZAHIDI, 2016; OESTERREICH *et al*., 2019). Ela é importante, sobretudo, porque exerce, e tende a continuar exercendo, significativo impacto sobre as áreas de economia e negócios (DIRICAN, 2015).

Não obstante os benefícios da IA sejam amplamente destacados por especialistas em tecnologia dentro e fora das universidades, as desvantagens, fragilidades e pontos críticos

relacionados à IA também têm se agigantado (BROUSSARD, 2018; KHANZODE; SARODE, 2020; TENG, 2019) e são cada vez mais alardeados.

As críticas mais comuns dizem respeito a um progresso, prometido pelos desenvolvedores de IA, que nem sempre se efetivou (GÜNGOR, 2020; MANNES, 2020; ZHANG; NAUMAN, 2018). O *marketing* no setor de tecnologia, em matéria de IA, parece ter vendido mais ideias que produtos, é o que sustentam alguns (BROUSSARD, 2018). O anúncio de grandiosos avanços que depois se revelaram "mais do mesmo", levou ao descrédito da IA em certos círculos (BROUSSARD, 2018; DREYFUS, 1992).

Há também quem afirme que a IA não é propriamente inteligência (SKALFIST; MIKELSTEN; TEIGENS, 2019) posto que, filosoficamente, a inteligência é a capacidade de imaginar, abstrair, criar, aprender, fazer conexões lógicas de forma autônoma. (GARDNER, 1983; NEISSER *et al.*, 1996). Nenhuma destas competências se verifica na IA sem que o ser humano, de algum modo, a programe para tal. Nesse contexto de questionamento, o estudo da IA enquanto modelo, ferramenta, conceito e produto mostra-se pertinente e necessário.

Uma IA capaz de "pensar e agir por si só", isto é, de forma autoconsciente, é considerada IA forte; ao passo que uma IA subsidiada e supervisionada por humanos, que apenas automatiza (de forma mecânica e não autoconsciente) a resolução de problemas com base em scripts elaborados por estes humanos, é considerada IA fraca. No presente estudo, será adotada a perspectiva da IA fraca – tendo em vista que a hipótese forte da IA ainda não se encontra contemplada, integrada e efetivada no contexto empresarial.

O Comitê Diretivo de Inteligência Artificial da Stanford University (Estados Unidos da América) conduziu uma pesquisa a nível global e afirmou que entre 1998 e 2018 o volume de artigos em geral (revisados por pares), em matéria de IA, cresceu mais de 300% – representando 3% das publicações em periódicos, e 9% dos artigos publicados em conferências (PERRAULT *et al*., 2019). De acordo com Gray *et al.* (2014), no fim dos anos 1990 o volume de estudos nesta área diminuiu.

Passada esta fase de arrefecimento, houve aumento considerável dos trabalhos científicos, com destaque para os que examinaram a relação entre IA e sistemas de informações contábeis (SUTTON; HOLT; ARNOLD, 2016). Termos como "redes neurais", "*machine learning*" e "*expert systems*", conceitos basilares para a discussão acadêmica em torno da IA, figuraram bastante nas publicações científicas que buscavam entender a relação entre o bloco contabilidade-finanças-gestão e a crescente onda de inteligência artificial. Neste sentido, o International Journal of Intelligent Systems in Accounting, Finance and Management desempenhou um papel preponderante, e vanguardista, ao fomentar o estudo e divulgação desta

temática (O'LEARY, 1995). Na esteira deste aprofundamento da relação entre IA e contabilidade gerencial, Nielsen (2020, p. 1) afirma ter verificado que "muitas áreas clássicas e tópicos dentro de contabilidade gerencial e gestão de desempenho são candidatas naturais para [o estudo de] *machine learning* e inteligência artificial".

Assim, aos poucos a literatura em contabilidade gerencial foi se apropriando dos conhecimentos relativos à IA e percebeu tratar-se de uma nova era (BAIER, 2019; WEIMEI, 2020), que inauguraria uma série de mudanças – disruptivas e transformadoras – para as organizações, para os profissionais e para a própria contabilidade (MARRONE; HAZELTON, 2019), que provocaria alterações no tipo e na qualidade da matéria-prima da contabilidade (a informação) (HALEEM; RAISAL, 2016), que redefiniria o uso dos sistemas na contabilidade e a própria função contábil (MIRZAEY *et al*., 2017; PETKOV, 2020; WEIMEI, 2020).

Embora, em termos científicos, o conhecimento sobre a IA e suas aplicações à contabilidade gerencial tenha se ampliado, dentro das organizações as transformações não acontecem no mesmo momento nem na mesma velocidade: o uso de redes neurais, a reconfiguração do papel do contador gerencial em função da absorção de (parte) de suas funções pela IA, a adaptação a uma era em que tudo (ou quase tudo) é digital, são eventos que requerem tempo para maturação e, por isso, constituem um processo de transição (BAIER, 2019; MARRONE; HAZELTON, 2019; MIRZAEY *et al*., 2017; WEIMEI, 2020).

Assim é que, de acordo com uma pesquisa internacional conduzida pelo Chartered Global Management Accountant (CGMA, 2019), apenas 11% dos entrevistados alegaram usar robótica, e somente 5% deles dizem utilizar *cognitive computing*. O uso da IA em contabilidade gerencial, de acordo com os resultados desta pesquisa do CGMA, encontra-se em um estágio inicial, embrionário. Mesmo quando vencida a barreira dos custos de implantação, parece ser necessário adquirir confiança nesta tecnologia e na integridade dos dados por ela fornecidos para, só então, dar-lhe adesão (CGMA, 2019).

Em meio a este debate sobre as virtudes, defeitos e limitações da IA, a contabilidade gerencial tem buscado extrair da IA tantos benefícios quanto seja possível, esmerando-se em diminuir debilidades que eventualmente possam eclodir (LAWSON, 2019). Aceitar a introdução de tecnologias baseadas em IA, compreender a utilidade, a adequação e a conveniência do seu uso, e integrar efetivamente a IA às rotinas da contabilidade gerencial é um desafio em curso (VĂRZARU, 2022).

A contabilidade gerencial se destaca de outros ramos da contabilidade e da administração porque, grosso modo, procura a partir dos dados da contabilidade oferecer um suporte embasado para as decisões gerenciais (DWIPUTRA; MUSTIKASARI, 2021;

GARRISON; NOREEN; BREWER, 2013; MARION; RIBEIRO; 2017). É uma contabilidade que se enxerga e se pretende como preditiva (COKINS, 2013; COOPER; CROWTHER; CARTER, 2001). Entretanto, o advento dos inúmeros sistemas contábeis, fiscais, e gerenciais provocou uma enxurrada de dados de tal forma que prever – razoável e fundamentadamente – qualquer coisa se tornou uma tarefa por demais laboriosa. Neste sentido, há quem preveja que os relatórios financeiros produzidos pela contabilidade sofrerão alterações relevantes em decorrência da aplicação de IA (TÜREGÜN, 2019).

Alguns estudos sobre IA aplicada à área de contabilidade e gestão versam sobre os impactos de dispositivos específicos (KESUMA; SAIDIN; AHMI, 2016; RAYMOND; PARÉ, 1992), normalmente em um corte longitudinal (de "antes e depois"), sem, contudo, realizar um comparativo para dimensionar, de forma simultânea, o impacto de tecnologias de natureza distintas sobre uma mesma realidade ou problema.

A IA representa não apenas uma inovação incremental para o rol de ferramentas da contabilidade gerencial: ela é um elemento disruptivo, capaz de transformar efetivamente a realidade da contabilidade e dos contadores (MARRONE; HAZELTON, 2019). Não obstantes as limitações e obstáculos com os quais se depara (LOSBICHLER; LEHNER, 2021), a aplicação da IA à contabilidade gerencial é vista como elemento facilitador das decisões (CHONG; EGGLETON, 2003), capaz de melhorar as estimativas da contabilidade (DING *et al.*, 2020), capaz de (re)definir o perfil dos contadores gerenciais (BRANDS; HOLTZBLATT, 2015) e transformar a contabilidade a partir de dentro e adequando-a à era digital (CGMA, 2019; LAWSON, 2019; NIELSEN, 2020).

A questão é que o emprego de IA nos negócios, especialmente como ferramenta de suporte à contabilidade gerencial, apresenta uma série de especificidades que tornam o estudo desta interface entre tecnologia e contabilidade particularmente necessária e oportuna. O recente estudo de Zhang *et al.* (2023), por exemplo, mostra que a presença de vieses nos algoritmos que orientam decisões gerenciais pode semear dúvida em alguns *stakeholders* sobre a aplicação deste tipo de tecnologia em processos decisórios; os autores evidenciam também riscos éticos associados à adoção da IA, insegurança dos contadores gerenciais com relação à capacidade de a IA solucionar problemas, e a dificuldade deles em lidar com modelos de IA pouco transparentes.

Neste sentido, há quem advogue pela ideia de que o ponto central de qualquer modelo de IA é, de fato, explicar o seu funcionamento, tornando-o claro e compreensível (GUPTA, 2023). Segundo estes, desta forma seria mais fácil construir confiança na IA, catalisar a adoção de modelos de IA nas organizações, e estabelecer um parâmetro para avaliá-los. Esta ideia de

uma “explicação suficientemente transparente”, contudo, também tem nuances – as quais serão discutidas na próxima seção.

2.2 TRANSPARÊNCIA DA INTELIGÊNCIA ARTIFICIAL

Transparência refere-se à propriedade que permite ao usuário de um sistema compreender a lógica interna e o processo de raciocínio do referido sistema (CHIEN *et al.*, 2014). Em um modelo de IA qualquer, a transparência é composta e avaliada a partir de alguns aspectos que vale a pena dissecar para entender melhor, quais sejam: complexidade, previsibilidade, explicabilidade e compreensibilidade.

A complexidade da transparência, às vezes associada ao seu nível de sofisticação (DAVENPORT, 2018; HALEEM; RAISAL 2016; KELLY; SELFRIDGE, 1962; KESUMA; SAIDIN; AHMI, 2016), diz respeito ao modo como a informação é gerada, comunicada e exibida. *Designs* cuja transparência é considerada complexa estão associados a uma apresentação confusa de grandes e diversificados conjuntos de informações (ZHAO *et al.* 2018). A previsibilidade consiste na capacidade de gerar predições a partir de um processo transparente e contribui para a formação da confiança dos usuários nos modelos de IA (MUIR, 1994; MUIR; MORAY, 1996**)**.

A explicabilidade tornou-se, no ambiente dos estudos em IA, um tema relevante - que acabou constituindo-se em uma área de estudo própria (DOŠILOVIĆ; BRČIĆ; HLUPIĆ, 2018). Ela representa o polo complementar da compreensibilidade (ADADI; BERRADA, 2018). Enquanto esta última diz respeito à capacidade de uma tecnologia ser entendida, a primeira versa sobre a sua capacidade de ser explicada, explanada, detalhada (DORAN; SCHULZ; BESOLD, 2017). A compreensibilidade depende de um usuário com conhecimentos suficientes para acessar a lógica embutida na ferramenta. Já a explicabilidade é intrínseca à ferramenta e depende de uma estrutura lógica interna simples e intuitiva (idem).

Assim, pode-se dizer que um modelo é explicável, por exemplo, quando ele apresenta uma mensagem de erro cada vez que identifica uma inconsistência nos dados que o alimentam. Se tal mensagem aponta “erro *x* deve-se à inconsistência *y*”, significa que o sistema se explica (PREECE, 2018; ROSENFELD; RICHARDSON, 2019). A sua compreensibilidade, porém, dependerá do usuário que vai se deparar com a mensagem de erro: o sistema será tão mais compreensível quanto menos conhecimento for exigido do usuário para entender a conexão do erro *x* à inconsistência *y* (ADADI; BERRADA, 2018). Ou seja: quando o que foi comunicado ao usuário puder ser compreendido recorrendo-se o mínimo possível a argumentos difíceis e

estruturas lógicas complexas, se poderá dizer que o sistema é compreensível. Em certo sentido, portanto, a compreensibilidade se opõe à complexidade.

Assim, o pleito por transparência em IA não diz respeito simplesmente à exibição da estrutura interna dos algoritmos que estão por trás dela. Para além disso, esta transparência pressupõe a disposição de apresentar e explicar os dados e os processos, e a explicação – por sua vez – pressupõe a capacidade dos usuários compreenderem o modelo de IA com o qual estão lidando. Há, portanto, um encadeamento entre a apresentação do modelo (*transparency*), a explicação do modelo (*explainability*) e o entendimento do modelo (*comprehensibility*) (HERM *et al.*, 2021).

A transparência é por vezes questionada quando diz respeito a informações de natureza sensível ou estratégica (TAMMINEN, 2022). Não convém, por exemplo, que uma organização – sob pretexto de ser transparente – divulgue os seus planos de lançar no mercado um novo produto: isto a fragilizaria do ponto de vista da concorrência. Então, entre os dilemas relacionados à transparência está a questão de o quê, quando e como deve-se apresentar certos conteúdos (YAMPOLSKIY, 2020). Atrelado a isto está um outro ponto nevrálgico: saber qual é o benefício em ser transparente – se é que há algum. Embora haja quem argumente que a transparência, dependendo de como e para quem seja apresentada, pode influenciar de forma significativa os rumos de uma decisão (DARGNIES; HAKIMOV; KÜBLER, 2022), isto não pode ser tomado taxativamente como benefício inconteste.

Pensando no dilema do pôquer *versus* xadrez, anteriormente mencionado, é preciso perguntar-se se o fato de ter uma IA mais transparente (como o xadrez, em que todas as peças estão à mostra no tabuleiro) traz algum benefício para a organização ou se, pelo contrário, a expõe para além do necessário. Da mesma forma, também é oportuno saber se uma IA menos transparente (cuja discrição recorda o pôquer, em que as cartas são mantidas sob custódia sigilosa nas mãos dos jogadores) seria o melhor caminho para preservar vantagens competitivas ou se, pelo contrário, manter certas informações na penumbra traria algum tipo de prejuízo (porque, talvez, alimentasse desconfianças sobre a organização e sobre a tecnologia por ela adotada). Este tipo de dilema justifica quem sustenta que não há evidências conclusivas de que a transparência impacte de modo significativo a assertividade e a precisão de uma decisão com suporte de IA (ALUFAISAN *et al.,* 2021).

A transparência tem graus. A comunidade científica internacional que lida com tecnologia chamou de modelos *black box* todas as propostas que contêm uma função matemática complexa (como máquina de vetor de suporte e redes neurais) e que precisam de um entendimento profundo de funções de distância e representação de espaço, algo difícil de

ser explicado e entendido por usuários simples (LOYOLA-GONZALES, 2019). Em um modelo *black box* os *inputs* são geralmente volumosos, diversificados, e sua classificação se dá por meio de interações aleatórias que, uma vez realizadas, formam tantas e tão distintas redes neurais que se torna difícil classificar, ordenar e compreender tais dados segundo a lógica comumente utilizada pela mente humana (ADADI; BERRADA, 2018). Em outras palavras: o processo como um todo é bastante obscuro uma vez que as interações entre os dados e a geração dos *outputs* depende muito de uma massa de informações que não é totalmente conhecida nem foi previamente moldada/condicionada.

Nestes modelos, o nível de complexidade da sua estrutura é grande, inexistem regras de programação preconcebidas, e as associações entre os dados (milhões deles, em uma base não completamente conhecida) são completamente aleatórias (ADADI; BERRADA, 2018). Eles não explicam como e por que chegaram a uma recomendação ou decisão específica (NIELSEN, 2020). De fato, a natureza dos modelos *black box* permite poderosas predições, mas que não podem ser diretamente explicadas (ADADI; BERRADA, 2018). Tendo em vista que eles não explicam "como e porque chegaram a uma determinada decisão ou recomendação. Isto força muitos usuários a dizer que 'o algoritmo me fez fazer isto'" (NIELSEN, 2020, p. 13) – o que gera incerteza e falta de controle, prejudicando a construção de confiança e colocando os gestores em uma situação delicada, na qual precisam tomar decisões a partir de modelos que não entendem e de cuja interpretação não estão seguros (MCKINGHT; CARTER; CLAY, 2009; NIELSEN, 2020).

Os modelos do tipo *white box*, por sua vez, são baseados em padrões, regras e árvores de decisão. Eles, geralmente, podem ser compreendidos por seus usuários já que fornecem um modelo mais próximo da linguagem humana (LOYOLA-GONZALES, 2019). Eles são, via de regra, menos complexos (do ponto de vista da sua lógica interna), contam com regras pré-estabelecidas pelo programador, e geram decisões a partir de associações entre dados conhecidos (ADADI; BERRADA, 2018). Normalmente, esses modelos proporcionam um bom equilíbrio entre precisão e explicabilidade (LOYOLA-GONZALES, 2019).

Nos modelos *white box* os dados são selecionados obedecendo a uma determinada ordem, classificados a partir de argumentos lógicos claros e específicos, e as regras de associações entre eles são ditadas por um *script* previamente elaborado (em geral um algoritmo que já é bastante utilizado entre os desenvolvedores de sistemas). Em resumo, se conhece todo o processo, do início ao fim (e por isso se pode dizer que há mais transparência) (NIELSEN, 2020). O Quadro 1 resume os principais aspectos que distinguem os modelos *white box* e *black box*.

Quadro 1 – Caracterização dos modelos de IA de acordo com a transparência do processo

Aspecto	*White Box*	*Black Box*
Dados	Conhecidos e quantificáveis	Não totalmente conhecidos e em volume imenso
Regras	Estabelecidas pelo programador	Definidas aleatoriamente a partir de associações entre os dados
Lógica	Básica (se; então)	Funções matemáticas, vetores e redes neurais
Complexidade	Baixa	Alta
Vantagem	Facilmente compreensível por parte do usuário comum	Previsões com elevado grau de precisão

Fonte: adaptado de Adadi e Berrada (2018), Loyola-Gonzales (2019) e Nielsen (2020).

De acordo com Loyola-Gonzáles (2019), há uma tendência de migração dos modelos *black box* para os modelos *white box*. Sobretudo em áreas como saúde e finanças. Essa tendência decorre de uma demanda do mercado por modelos que sejam, sim, acurados mas concomitantemente explicáveis e compreensíveis – para, assim, proporcionar mais segurança a quem quer que tome decisões com base nos *outputs* de tal modelo. Desta forma, diariamente se impõe o desafio de tentar construir o modelo perfeito, ao mesmo tempo em que se amplia paulatinamente a discussão sobre as consequências de adotar/manter este ou aquele modelo (LOYOLA-GONZÁLES, 2019).

Além de *black box* e *white box*, a literatura destaca os modelos *gray box*, como um ponto intermediário entre os dois primeiros (BOHLIN, 2006; KROLL, 2000; OUSSAR; DREYFUS, 2001). Importante é notar, porém, que a gradação *black – gray – white* não deriva exatamente da quantidade de informação que o modelo apresenta quando gera seus *outputs*. A gradação ocorre em função da codificação interna do modelo, isto é, do modo como os *inputs* são tratados no interior do modelo (BOHLIN, 2006).

Na presente tese, o modelo *white box* descrito no cenário que foi apresentado aos participantes da pesquisa seguiu a estratégia dos chamados sistemas supervisionados. A razão disto é que em sistemas supervisionados os modelos partem de um conjunto de dados classificados e treinados (por humanos) e, com base nisso, ao receber novos dados, eles são capazes de prever o comportamento destes novos dados, bem como de fornecer *feedback* disso aos seus usuários (NASTESKI, 2017; NIELSEN, 2020). Assim, a participação humana no tratamento dos dados, a presença do fator preditivo, e o oferecimento de uma resposta objetiva ao usuário são elementos que se alinham à proposta de uma tomada de decisão gerencial relativa a orçamento com suporte de um modelo de alta transparência.

Por outro lado, o modelo *black box* que consta do cenário apresentado aos participantes desta pesquisa seguiu a estratégia dos sistemas de *deep learning*. Tais sistemas se valem de um amplo conjunto de dados, tanto qualitativos como quantitativos, para associar os dados, formar redes neurais e aprender a partir de sua própria experiência. Isto lhes proporciona resultados

mais acurados (NIELSEN, 2020; SHINDE; SHAH, 2018) se comparados às outras abordagens de *machine learning*, mas ofusca a conexão entre *inputs* e *outputs* – que não pode ser verificada de forma clara nem estabelecida de forma direta. Desta forma, ao vincular o modelo *black box* à estratégia de *deep learning* acentua-se a sua diferença em relação ao modelo *white box* intensificando o dilema entre mais transparência ou mais acurácia.

2.3 DECISÕES GERENCIAIS COM SUPORTE DE INTELIGÊNCIA ARTIFICIAL

Um dos papéis da contabilidade gerencial – talvez o mais frequente – é fornecer subsídios de qualidade para a tomada de decisão (CHONG; EGGLETON, 2003; GHANBARI; VASELI, 2015; HALL, 2010; NOVIANTY, 2015). Fundamentar-se em dados e ferramentas confiáveis é, portanto, um pressuposto para que se construam soluções adequadas aos problemas que se apresentam, orientações alinhadas aos objetivos estratégicos, e predições bem ponderadas e assertivas.

É nesse ponto, portanto, que a IA tangencia a questão da decisão: a habilidade de tomar decisões – antes restrita à capacidade humana de avaliar e sopesar – agora compete a um algoritmo que, de modo mais ágil e eficaz que o ser humano, consegue lidar com bilhões de dados, variáveis e critérios (SAGIROGLU; SINANC, 2013). Os modelos de IA tomam decisões e estas decisões são tomadas de forma mais veloz, mais precisa e mais racional do que qualquer humano seria capaz (BARRAT, 2013). Desta constatação, porém, derivam preocupações sobre quem se responsabiliza no caso de uma decisão errada tomada por uma máquina, e também sobre qual a amplitude do poder decisório que deve ser delegado a estas máquinas (que operam com uma valoração moral diferente da humana – algo que parece importante considerar em certas decisões). Decisões clínicas, decisões militares, decisões amorosas, são exemplos de áreas cinzentas que motivam calorosos debates sobre a necessidade, a conveniência e a efetividade de usar a IA em tais casos.

No campo das decisões gerenciais uma das principais polêmicas consiste em convencer gestores de que, embora não se saiba como, algoritmos complexos e de certo modo incompreensíveis (MARCUS; DAVIS, 2019), de fato, conseguem ser efetivos e bastante assertivos ao orientar decisões.

Um outro ponto bastante debatido quando se trata de tomada de decisões em negócios empresariais diz respeito aos riscos atrelados às decisões (BĂRBUȚĂ-MIȘU *et al.*, 2019; KOCHER; POGREBNA; SUTTER, 2008; MANNES, 2020; RIABACKE, 2006; SINGH, 1986). Longe de propor uma estratégia onde o risco seja nulo, a intervenção dos modelos

tecnológicos – notadamente aqueles originários da IA – visa dar um nível mais elevado de tecnicidade às decisões (CEPNI, 2019). O cerne da proposta de tomar decisões com suporte de IA, portanto, não está em "não assumir riscos" (tampouco em "não errar") e sim em balizar as decisões por critérios técnicos, condizentes com a estratégia empresarial, e ponderados a partir das múltiplas variáveis que tangenciam o objeto da decisão. Este argumento gira em torno da ideia de que o ser humano – quando posto diante de um conjunto grande de dados e critérios – faz um esforço cognitivo mais complexo e elaborado para poder analisar, avaliar e se posicionar (SWELLER, 1988). A utilidade da IA, portanto, estaria em fazer este papel considerando objetivamente os dados e ponderando os critérios com uma facilidade e rapidez superiores à capacidade humana (CUI *et al*., 2018).

A pretensa objetividade das decisões tomadas com suporte da IA contrasta com a subjetividade que circunda e permeia o ambiente decisório (LEAVY; O'SULLIVAN; SIAPERA, 2020). Um simples erro por parte de um programador, por exemplo, poderia causar danos significativos a uma empresa que tem a IA como sua principal ferramenta decisória. Um profissional (ou conjunto de profissionais) de tecnologia que, por qualquer razão, se corrompa pode programar decisões erradas para propositadamente gerar prejuízos à organização em que atua. Uma invasão *hacker* poderia promover alterações em um sistema de IA de modo a enviesar seus *outputs* e, por conseguinte, as decisões deles derivadas. Em suma, os riscos decorrentes da ação humana na interação com a IA naturalmente existem – e até demandam, no entender de alguns, controle e regulação – mas não diferem essencialmente dos riscos de outros modelos e ferramentas (BAUGUESS, 2017).

Dispor de muitos dados não garante a geração de informação de qualidade, tampouco assegura a obtenção de *insights* inovadores e úteis (LUFTMAN; BEN-ZVI, 2010). Para que as decisões sejam coerentes, fundamentadas e assertivas é preciso que haja um ajuste entre os tomadores de decisão e os diversos sistemas de informação que os apoiam (ALLES, 2015). Neste sentido, os modelos de IA se propõem a integrar o processo decisório facilitando a análise de grandes volumes de dados e trazendo à tona a melhor decisão possível de acordo com as informações disponíveis (ALUFAISAN *et al*., 2021; HORVITZ; BREESE; HENRION, 1988; PHILLIPS-WREN, 2012; POMEROL, 1997; RADERMACHER, 1994; SCHEMMER *et al.*, 2022). Desempenham, grosso modo, papel semelhante ao de outros sistemas de suporte à decisão só que com tecnologia mais sofisticada. Diferentemente de outros sistemas, porém, os modelos de IA não apenas tratam os dados para gerar informações, mas também são capazes de orientar objetivamente o julgamento dos gestores indicando-lhes qual a decisão mais adequada a determinada circunstância (LU *et al*., 2015; NUNES; JANNACH, 2017).

Importante notar que as decisões são condicionadas por fatores externos e internos (SHATILO, 2019). Há um ambiente e uma circunstância específica para cada decisão, de tal forma que a análise de conjuntura é um desafio constante para os decisores. Estes, aliás, com seu conjunto único de experiências, conhecimentos e percepções, constituem eles próprios também uma importante variável na configuração decisória (NOORAIE, 2008). Por isso, na presente tese, a tomada de decisão não é um processo estritamente objetivo e puramente lógico: trata-se de uma dinâmica contextualizada e permeada pelo componente humano.

As decisões diferem, entre outras coisas, quanto ao seu alcance. Anthony (1965) as distinguiu em três níveis: operacionais, táticas e estratégicas. As decisões operacionais, em geral, são tomadas por gerentes operacionais e dão conta de eventos surgidos durante a execução das tarefas (IMANE; DRISS, 2017). Elas podem estar relacionadas ao controle/gestão dos estoques, à alocação da força de trabalho, à manutenção de equipamentos, à revisão de gargalos na produção etc. (DHAMIJA; BAG, 2020; SHIVAKUMAR, 2014). Uma decisão operacional visa promover mudanças em uma determinada função/tarefa/processo de modo a facilitar sua execução e aumentar sua efetividade (TURNER, 2003). Por sua própria natureza e característica rotineira, este tipo de decisão tem impacto no curto prazo (IMANE; DRISS, 2017; MISNI; LEE, 2017; SHIVAKUMAR, 2014).

Por seu turno, as decisões estratégicas versam sobre o *design* de produtos e serviços, a gestão da qualidade, o *layout* dos processos, o nível de automação envolvido, a gestão da cadeia de suprimentos etc. (DHAMIJA; BAG, 2020). Normalmente tomadas pelo alto escalão (HARRISON; PELLETIER, 2000), este tipo de decisão fala de como a empresa se relaciona com seus *stakeholders*, de como ela constrói (ou não) pontes com o mundo externo a ela (IMANE; DRISS, 2017). Por isso, tais decisões são mais sensíveis, no sentido de que um deslize ou erro qualquer pode prejudicar a imagem da organização e/ou sua competitividade ante a concorrência, causando grandes danos (WARD; DURAY, 2000). Em geral, as decisões estratégicas projetam-se no longo prazo (HARRISON; PELLETIER, 2000; IMANE; DRISS, 2017). Todas estas noções acerca dos elementos caracterizadores das decisões estratégicas já haviam sido sumarizadas no trabalho de Harrison e Pelletier (1995), de modo que a literatura esparsa que se seguiu apenas reescreve e reforça os pontos outrora elencados por estes autores.

O Quadro 2 apresenta algumas das principais diferences entre as decisões operacionais e as decisões estratégicas:

Quadro 2 – Caracterização dos tipos de decisão

Aspecto	Decisão Operacional	Decisão Estratégica
Reflexo temporal	Curto prazo	Longo prazo
Instância competente	Gerentes operacionais	Alto escalão
Frequência	Rotineira	Esporádica
Objeto	Estoques, processos, patrimônio, finanças, pessoal etc.	*Design* de produtos serviços; gestão de qualidade; posicionamento de mercado etc.
Objetivo	Garantir a continuidade das operações	Garantir a competitividade da organização
Foco	Interno	Externo

Fonte: adaptado de Dhamija e Bag (2020), Harrison e Pelletier (1995, 2000), Imane e Driss (2017), Misni e Lee (2017), Nooraie (2008), Shivakumar (2014), Turner (2003) e Ward e Duray (2000).

As decisões táticas não serão objeto de reflexão e análise na presente tese, tendo em vista que – em uma perspectiva de organograma – o nível tático, responsável por tomar esse tipo de decisão, difere muito de empresa para empresa no que concerne às suas atribuições e poderes. Consequentemente, as decisões tomadas em nível tático variam bastante entre as organizações, dificultando a construção de um caso experimental capaz de capturar a essência geral deste tipo de decisão.

2.4 DESENVOLVIMENTO DE HIPÓTESES

Mesmo com todo o aprimoramento tecnológico dos últimos anos, um elemento tem sido descrito como privativo da espécie humana: tomar decisões (GONZÁLEZ-MENDOZA, CAÑIZARES-ARÉVALO; CARDENAS-GARCÍA, 2022; PHILLIPS-WREN, 2012; TRONCO *et al.,* 2019). Argumenta-se que a capacidade de avaliar componentes subjetivos e circunstanciais, que normalmente permeiam o processo decisório, é algo essencialmente humano. Há uma percepção de que "máquinas não podem pensar nem sentir" (BIGMAN; GRAY, 2018, p. 21). Em decorrência disso, há uma aversão a atribuir certas decisões a uma máquina – ainda que seja tecnicamente possível fazê-lo (BIGMAN; GRAY, 2018). Outro argumento assinala que apenas os humanos, e não as máquinas, podem ser considerados agentes responsáveis por suas decisões. Logo, em sentido estrito, só os humanos estão aptos a conduzir um processo de tomada de decisão. (COECKELBERGH, 2020; GUNKEL, 2020).

Para conciliar esta suposta exclusividade dos humanos no que compete à tomada de decisões com o desempenho (cada vez mais célere, preciso e eficiente) da inteligência artificial, fala-se no estabelecimento de uma relação simbiótica entre a força de trabalho humana e os sistemas de IA (JARRAHI, 2018; WILSON; DAUGHERTY, 2018). A complementaridade, a interação, a colaboração entre seres humanos e modelos de inteligência artificial tem sido

destacada em diversos trabalhos (AMERSHI *et al.,* 2019; BANSAL *et al.*, 2019; REN; BAO, 2020; WILSON; DAUGHERTY, 2018).

Este encadeamento de ideias, propostas e fatos tem estimulado pesquisadores a estudar os chamados *recommender systems* (JANNACH, 2017; LARKIN; DRUMMOND; ÁRVAI, 2022; LU *et al.*, 2015; NUNES; VULTUREANU-ALBIŞI; BĂDICĂ, 2021), cuja concepção estabelece que cabe à IA recomendar o que seria a decisão tecnicamente mais adequada, e ao ser humano cabe avaliar e acatar (ou não) a sugestão que recebeu da IA. Estes sistemas materializam a *Human-Computer Interaction* (HCI), fundamento teórico básico para quem visa aprofundar o entendimento sobre as particularidades da relação homem-máquina (BANSAL; KHAN, 2018; DIEDERICH *et al.* 2022; PUSTEJOVSKY; KRISHNASWAMY, 2021; REN; BAO, 2020).

O estudo da HCI ajuda a compreender, atualizar e (re)definir os papéis que os humanos e os computadores desempenham quando cooperam em prol da realização de uma tarefa, além de esclarecer eventuais consequências decorrentes desta interação. Por definição tem-se que a *Human-Computer Interaction* é "a área de interseção entre a psicologia e as ciências sociais, por um lado, e a ciência da computação e a tecnologia, por outro. Pesquisadores da HCI analisam e projetam interfaces tecnológicas para usuários específicos" (CARROLL, 1997, p. 501).

A HCI pressupõe que inteligência artificial pode assumir diferentes perspectivas – a depender do grau de autonomia/consciência que a ela se atribua em relação ao fator humano. Tradicionalmente, estas perspectivas (chamadas "hipóteses") são duas: hipótese fraca (ou restrita) e hipótese forte (ou geral) (SEARLE, 1980).

A hipótese fraca da IA, que se valida e se expande a cada nova ferramenta ou modelo desenvolvido, contesta algumas ideias comuns, mas não necessariamente verdadeiras, a respeito da IA. A primeira delas é de que ao utilizar dispositivos formulados com essa lógica, a consequência natural seria a instauração da autonomia das máquinas (com consequente eliminação da autonomia humana).

Pesquisas dão conta de que a plena autonomia das máquinas não se dá de forma imediata, e sim gradual; e também não se dá da mesma forma em todas as áreas profissionais: há áreas em que o esforço tecnológico da IA poderia substituir facilmente o trabalho humano, e outras em que ainda não se consegue vislumbrar tal substituição. Frey e Osborne (2017) assinalaram, por exemplo, que o ofício de escrivão e também o de atendente de telemarketing têm grande probabilidade de ser exercidos inteiramente por máquinas; ao passo que a atividade

de um cirurgião, de um profissional de relações públicas ou de um designer são pouco suscetíveis a migrar do humano para o computador.

Defensores da hipótese fraca da IA discordam da hipótese forte por questões filosóficas (como o próprio conceito de "inteligência"), mas também apresentam questões de ordem prática que põem em xeque a plena autonomia das máquinas. Tópicos relacionados a: **i)** confiança na ferramenta tecnológica (SCHAEFER *et al.*, 2016); **ii)** estabilidade de emprego (VENKATESH; BALA, 2008); **iii)** limites inerentes ao sistema de informações (LOSBICHLER; LEHNER, 2021); i**v)** maturidade digital dos usuários (KANE *et al*., 2017); **v)** consequências de um eventual *enforcement* legal/normativo em matéria de tecnologia (GAVIRIA, 2020); **vi)** descrença quanto ao benefício marginal gerado com a IA (GLIKSON; WOOLLEY, 2020); e muitos outros fatores dificultam a concretização da hipótese forte e permanecem em aberto, sem que haja uma investigação que avalie o impacto conjunto deles sobre a autonomia.

Convém notar que autonomia de máquina, tal como enunciada pela hipótese forte, é uma sistemática em que "decisões (em resposta a entradas ou sinais externos de qualquer complexidade) são tomadas dentro do sistema e não envolvem tomada de decisão humana" (NORRIS; PATTERSON, 2019, p. 1). Não por acaso, autonomia e decisão andam lado a lado em muitos trabalhos (CORDESCHI, 2013; WALLACH, 2008), sendo a autonomia normalmente expressa como a capacidade, a liberdade e o poder para decidir.

O interesse por como gestores tomam suas decisões, ou seja, a investigação de seus "gatilhos" decisórios, passa pela investigação minuciosa da visão ontológica deles(as), isto é, pela forma como percebem e encaram o mundo à sua volta (HOLSAPPLE; JOSHI, 2004; SINGH; TRIPATHI; JARA, 2014). Os seus posicionamentos, pareceres e decisões estão vinculados à sua percepção da realidade, e refletem a consolidação das suas experiências passadas, da sua situação atual, e das projeções que adotam em relação ao futuro (AMASON; MOONEY, 1999; BASI, 1998; LAMOND, 2005). No contexto digital que o século XXI experimenta, é natural que esta percepção não prescinda das inovações tecnológicas que cotidianamente emergem e interpelam o homem contemporâneo (FERREIRA; MONTEIRO, 2021; PHILLIPS-WREN, 2012; SCHEMMER *et al*., 2022).

Dentre estas inovações, a inteligência artificial, suas ferramentas e modelos, como realidade que permeia e adentra cada vez mais o universo empresarial, é aqui aventada como gatilho capaz de alterar a percepção do(a) gestor(a) e, por conseguinte, afetar a decisão final dele(a). Evidentemente, a inferência da IA como gatilho decisório está circunscrita a um cenário e a determinados aspectos da decisão (PHILLIPS-WREN, 2012; POMEROL, 1997).

Ademais, se por um lado a participação humana na tomada de decisões se impõe e faz acontecer as coisas, por outro ela imprime às decisões (em maior ou menor grau) o viés do decisor – que nem sempre se harmoniza com os interesses circunstantes (FERNANDES; SCHNORRENBERGER; RENGEL, 2020; GARCIA *et al.,* 2012; TRONCO *et al.,* 2019).

O humano não somente acrescenta o seu viés à decisão, como também o inocula nos algoritmos da inteligência artificial (que se tornam, a partir de então, eles mesmos enviesados e propensos a perpetuar seu viés) (ASCARZA; ISRAELI, 2022). A preocupação, portanto, com os elementos que podem influenciar o decisor humano é atual e justificada – quer estes elementos estejam presentes em percepções anteriores (preconceitos) do decisor, quer façam parte da natureza intrínseca da decisão a ser tomada, quer se constituam como parte da lógica de processamento dos dados pela IA.

Neste sentido, a transparência é o elemento que, na presente tese, denota a característica (inata ao modelo de IA) potencialmente capaz de induzir um decisor a tomar esta ou aquela decisão, a aderir ou não a uma recomendação feita pelo modelo de IA (CRAMER *et al.*, 2008; LANG, 2018; NUNES; JANNACH, 2017; VULTUREANU-ALBIŞI; BĂDICĂ, 2021).

O desdobrar teórico que leva à construção da percepção dos gestores se baseia na ideia de que eles tomam suas decisões (aderindo ou não às recomendações da IA) a partir de uma ponderação entre o tipo de decisão envolvida e o grau de transparência do modelo de IA que lhe fez a recomendação. Esta ponderação decorre e reflete duas considerações específicas: **i)** as decisões são tipificadas segundo as suas características intrínsecas; e **ii)** diferentes graus de transparência inspiram diferentes níveis de confiança no modelo de IA.

2.4.1 Transparência e a aderência à IA em decisões gerenciais

Modelos de IA têm diferentes níveis de transparência quanto à sua lógica interna (ADADI; BERRADA, 2018; LARSSON; HEINTZ, 2020; LOYOLA-GONZALES, 2019). Esta diferença no nível de transparência pode repercutir na percepção dos gestores sobre o modelo de IA e, consequentemente, no posicionamento deles sobre as recomendações geradas pela IA. Em outros termos: diferentes níveis de transparência implicam em diferentes percepções e, por conseguinte, em diferentes decisões (LICHT; LICHT, 2020; ZERILLI *et al.,* 2018).

Makarius *et al.* (2020) evoca a ideia de que é improvável que um funcionário que não compreende a IA que está usando consiga fazer uso dela de modo a agregar valor à empresa. Por isso, “antes de IA ser introduzida na organização, gestores precisam ajudar os empregados

a atribuir sentido aos sistemas de inteligência" (MAKARIUS *et al.,* 2020, p. 267). Os autores também apresentam três tipos de abordagem comuns em estudos sobre IA: a abordagem cognitiva, a relacional e a estrutural. A proposta desta tese alinha-se a este entendimento de Makarius *et al.* (2020) e atende ao apelo do mesmo na medida em que aborda a inteligência artificial contemplando tanto a sua dimensão cognitiva (relacionada à tomada de decisões), como a sua dimensão relacional (concernente à confiança), como ainda a sua dimensão estrutural (concernente à transparência).

A transparência da IA ajuda a consolidar a percepção do usuário sobre a confiabilidade da IA (CHIEN *et al.*, 2016; GLIKSON; WOOLEY, 2020). Um processo fluido e nítido, um funcionamento claro e compreensível, dão ao usuário de um modelo de IA a segurança necessária para pautar suas decisões de acordo com as sugestões deste modelo (ou a validá-las quando for instado a fazê-lo) (CRAMER, 2008; MADSEN; GREGOR, 2000; VULTUREANU-ALBIŞI; BĂDICĂ, 2021).

A utilidade da transparência da IA tem sido corroborada por diferentes estudos (ALUFAISAN *et al.*, 2021; SCHEMMER *et al.*, 2022), a maior parte deles vinculando a transparência à construção de confiança nas informações e recomendações da IA (LARSSON; HEINTZ, 2020; SHIN, 2021; VON ESCHENBACH, 2021). Lang (2018), por exemplo, argumenta que quando usuários de determinado modelo de IA conseguem compreender sua lógica e avaliam que ela é consistente do ponto de vista cognitivo, passam a confiar mais no suporte que tal modelo oferece à tomada de decisão. Venkatesh e Bala (2008), por sua vez, alegam que quando os modelos de IA são – entre outros aspectos – compreensíveis, os usuários avaliam isto de forma positiva e melhoram a sua percepção quanto à qualidade e utilidade do modelo.

Assim, partindo do pressuposto de que confiança na IA é importante e se constrói com base na transparência para com os usuários, emergiu na área de ciências da computação um ramo específico de estudo atrelado à transparência da IA: *Explainable Artificial Intelligence* (XAI) (MESKE *et al.,* 2022). O argumento central da XAI é o de que a inteligência artificial deve ser capaz de ser explicada e compreendida (ADADI; BERRADA, 2018). Para que decisões não sejam tomadas às cegas, a inteligência artificial, quando atua como suporte decisório, precisa ser apresentada em detalhes e entendida minudentemente (ALUFAISAN *et al.*, 2021; LOYOLA-GONZALES, 2019).

A XAI materializa o vínculo entre transparência da IA, confiança na IA e tomada de decisões com suporte de IA. Compreender esta conjugação de fatores é pertinente e oportuno porque a literatura a este respeito já alertou que o excesso de confiança (*overtrust*) na tecnologia

pode levar a decisões descuidadas (com o descumprimento de procedimentos prévios à tomada de decisão) (CHIEN *et al*., 2014; GLIKSON; WOOLLEY, 2020; HANCOCK *et al*., 2011; MILLER, 2015; SUTTON; HOLT; ARNOLD, 2016); a falta ou quebra da confiança (*distrust*) pode inibir ou dificultar o uso de determinada tecnologia (CHIEN *et al*., 2014; HANCOCK *et al*., 2011); e, por fim, uma pré-disposição excessiva à confiança por parte dos decisores (*initial trust*) pode conduzir a um indesejado viés de automação nas decisões (LEWIS; SYCARA; WALKER, 2018; MILLER, 2015) etc.

A relação entre a confiança e a transparência (via XAI) tem sido abordada por alguns autores: ora de forma isolada (SHIN, 2021; VON ESCHENBACH, 2021), ora aplicada aos já mencionados *recommender systems* (C8RAMER *et al.*, 2008; VULTUREANU-ALBIŞI; BĂDICĂ, 2021; WANG *et al.,* 2010). Nesta última perspectiva, a aderência – isto é – o agir de acordo com a recomendação do sistema (CAVANAGH, 2010) é fundamental para que se avaliem os efeitos e a utilidade da transparência da IA na tomada de decisões (ALUFAISAN *et al.*, 2021; CRAMER *et al.*, 2008).

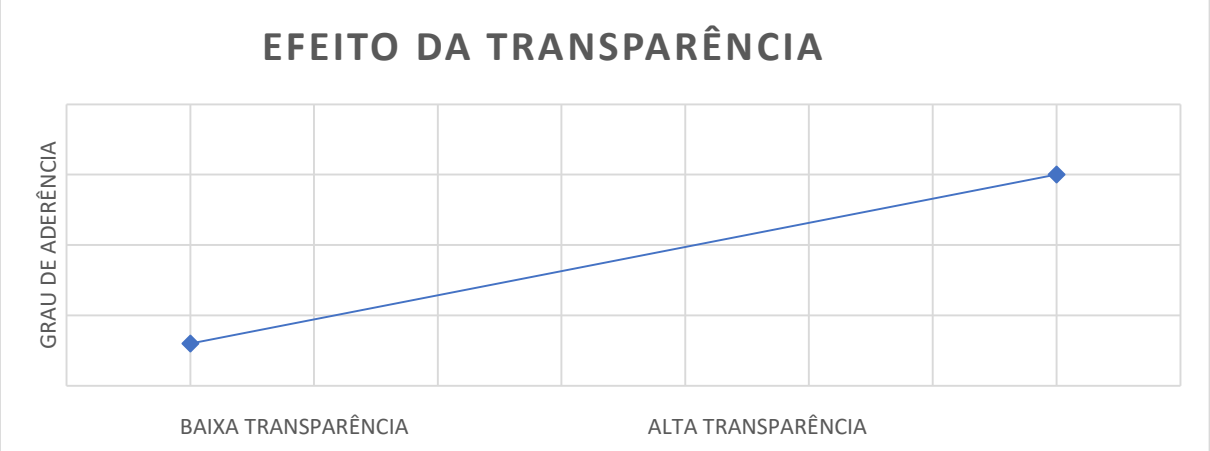
O tríplice estudo experimental de Lang (2018) destacou que características dos modelos de tecnologia, atreladas à natureza dos *inputs* que eles recebem, influenciam a formação da confiança dos seus usuários e acabam por impactar o julgamento e as decisões dos mesmos. O autor concluiu também que a similaridade entre a forma como os usuários e os sistemas "raciocinam" incrementa a percepção de utilidade que os primeiros têm em relação aos últimos. Entretanto, Lang (2018) tratou de sistemas de suporte mas não de sistemas de recomendação, não focou especificamente em modelos de IA, e abordou diferentes níveis de transparência a partir da forma de apresentação dos dados de *input* (e não a partir da lógica intrínseca aos modelos).

Assim, a concatenação dos argumentos sobre o papel da XAI, as implicações da confiança na IA, e a utilidade dos *recommender systems* sugere que mais transparência produz mais confiança e isto reverbera em maior aderência às recomendações do modelo de IA. Desta forma, a construção da primeira hipótese do presente estudo pode ser assim formulada:

H1: A aderência dos gestores às recomendações do modelo de IA é maior quando a transparência do modelo é alta do que quando é baixa.

Graficamente, esta hipótese poderia ser assim ilustrada (Figura 1):

Figura 1 – Predição do efeito da transparência sobre a aderência dos gestores



Fonte: elaboração própria.

2.4.2 Tipos de decisão a aderência à IA em decisões gerenciais

As decisões não têm todas o mesmo escopo, a mesma gravidade e o mesmo impacto (RADERMACHER, 1994). Por isso, a literatura acadêmica debruçou-se fartamente sobre a natureza das decisões e estabeleceu uma tipologia, por assim dizer, segundo a qual as decisões caracterizam-se como operacionais, táticas ou estratégicas (IMANE; DRISS, 2017). O caráter da decisão, portanto, é fator preponderante para que se avalie que metodologia deve ser empregada para tomar determinada decisão.

Em geral, os sistemas tradicionais de suporte à decisão são específicos (porque produzidos/customizados para tratar um tipo de problema e/ou orientar um tipo de decisão) e abertos (porque facilmente permitem o acesso a como os dados são organizados e processados). Watson (2017) apresenta diversas gerações de sistemas de apoio à decisão, descrevendo suas particularidades e destacando que a evolução deles ao longo do tempo aponta para os sistemas de IA lastreados em *cognitive computing*. Ao trazer isso para o universo da IA, em que o volume de dados costuma ser imenso (LEYER *et al*., 2020), que vai rodar a partir de uma fórmula lógico-matemática (algoritmo), cuja linguagem não é de uso comum e cuja tradução para o vernáculo por vezes não é fácil, a reticência, a desconfiança e a insegurança com relação aos modelos de IA para suporte à decisão começa naturalmente a surgir no decisor (ADADI; BERRADA, 2018; LARSSON; HEINTZ, 2020).

Ao examinar os modelos de IA que dão suporte às decisões, não raro os estudos apresentam modelos que suprem melhor determinada decisão e que podem não funcionar bem em outro contexto decisório (CHIEN *et al.,* 2014; KOLASINSKA; LAURIOLA; QUADRIO, 2019). É elucidativo o exemplo oferecido por Chien *et al*. (2014, p. 36), ao recordar que em

uma “grande amostra de pilotos, uma profissão tão especializada e regulamentada, [verificou-se que] a cultura nacional exerce uma influência significativa sobre a atitude e o comportamento [dos pilotos]”. Significa dizer que a ferramenta ou modelo de IA com o qual os profissionais trabalham não é um determinante absoluto de suas intenções ou ações. O contexto também pode influenciar positiva ou negativamente o seu comportamento e decisão final.

Confrontar o tipo de decisão envolvida com a percepção do usuário decisor sobre a IA pode evitar dois comportamentos: i) que decisores deixem de seguir as recomendações da IA em razão de preconceitos e traumas anteriores relacionados a IA ou à tecnologia em geral (MORAY, 1992; LEWIS; SYCARA; WALKER, 2018; LEE); ii) que apliquem “cegamente” as recomendações da IA, sem a devida reflexão e ponderação quanto objeto específico da decisão (CHIEN *et al*., 2014; GLIKSON; WOOLLEY, 2020; HANCOCK *et al.*, 2011; MILLER, 2015; SUTTON; HOLT; ARNOLD, 2016). Com esta tônica, Chien *et al.* (2014) alertaram que os humanos tanto podem deixar de utilizar a tecnologia quando seria apropriado utilizar, quanto podem aceitar suas recomendações e ações quando não seria o mais adequado a ser feito.

Esta classificação das decisões não apenas aponta para as características que as distinguem, mas também para o nível de risco associado a cada decisão. A percepção de risco está associada a um gerenciamento de risco particularizado e a decisões gerenciais casuísticas (RENN, 1998). Neste sentido, as decisões operacionais (tidas como corriqueiras e de menor impacto) representam menor grau de risco ao negócio como um todo; ao passo que as decisões estratégicas – mais abrangentes e impactantes – são mais arriscadas (sobretudo porque, ao considerar o horizonte de longo prazo, acabam se baseando em informações limitadas, indisponíveis ou incertas) (YANG; HAUGEN, 2015).

Desta forma, o decisor que utiliza a IA como ferramenta de suporte à decisão faz, ao mesmo tempo, uma avaliação do risco da situação-problema e do risco de pautar suas decisões pelo que orienta a IA (ARAÚJO *et al.,* 2020; DUNEGAN; DUCHAN; BARTON, 1992; MANNES, 2020; PASSAT; MERTENS, 2019). O estudo de Larkin, Drummond e Árvai (2021) forneceu um exemplo dessa situação ao mostrar que, analisando separadamente decisões sobre finanças e decisões sobre saúde, as pessoas (decisores) apresentam preferências distintas quanto a acolher recomendações oriundas de especialistas humanos ou recomendações oriundas de sistemas de IA. A percepção de risco delas muda conforme o tipo de decisão em questão e isso afeta a sua receptividade a recomendações e a sua decisão final.

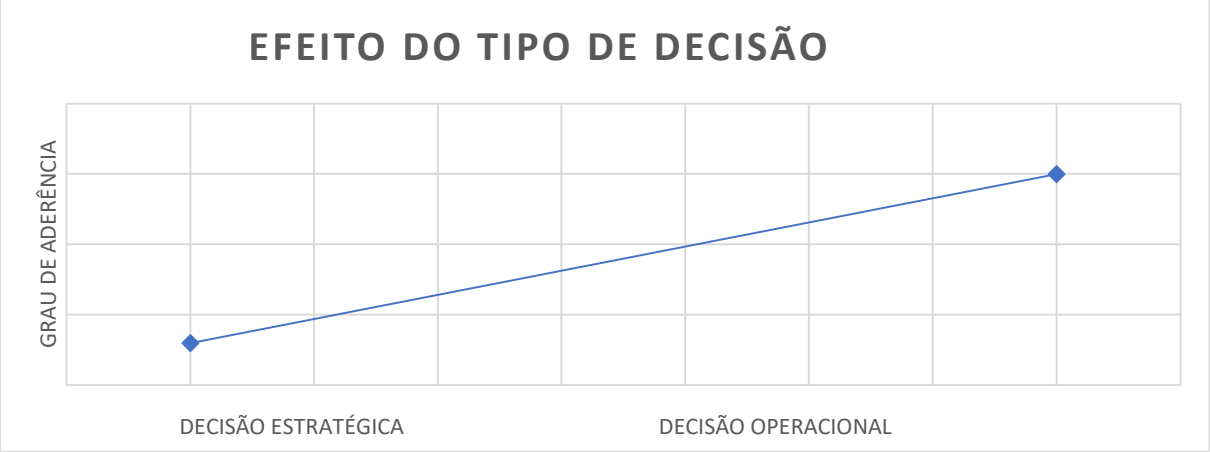
Assim, pode-se razoavelmente supor que em decisões mais corriqueiras e menos impactantes (operacionais) os gestores estão mais dispostos a aceitar as recomendações da IA

– pois eventuais erros, imprecisões ou inconsistências nesta decisão estão inseridos em um plano de baixo risco e reduzido impacto; ao passo que, em decisões mais sensíveis e abrangentes (estratégicas), estes gestores são mais reticentes em acatar as recomendações da IA – uma vez que este tipo de decisão envolve riscos elevados e tem um amplo impacto sobre a organização. Assim, foi elaborada a segunda hipótese deste estudo:

H2: A aderência dos gestores às recomendações do modelo de IA é maior quando o modelo se aplica a decisões operacionais do que quando se aplica a decisões estratégicas.

A representação gráfica de H2 foi ilustrada na Figura 2, abaixo:

Figura 2 – Predição do efeito do tipo de decisão sobre a aderência dos gestores



Fonte: elaboração própria.

2.4.3 Interação entre transparência e tipo de decisão na aderência à IA em decisões gerenciais

Ao retomar estas duas hipóteses preliminares, pode-se intuir que: **i)** quando a transparência do modelo de IA é alta, ela inspira potencialmente mais confiança nos gestores e, por isso, a aderência deles às recomendações do modelo de IA é maior; **ii)** quando se trata de decisões operacionais, os gestores potencialmente aderem mais facilmente às recomendações da IA, pois consideram que este tipo de decisão envolve menores riscos. A consequência destas duas hipóteses preliminares é que, em cenários de alta transparência, as recomendações de um modelo de IA aplicável a decisões operacionais podem obter maior aderência dos gestores em comparação às demais possibilidades de combinação entre transparência e tipo de decisão.

Em decisões operacionais (menor risco), elevar o nível de transparência do modelo pode deixar os gestores mais confortáveis e seguros com as recomendações dele; ao passo que em decisões estratégicas, dado o elevado nível de risco e todas as demais pressões a que estão associadas, o incremento na transparência do modelo de IA pode não ser suficiente para aproximar os gestores da recomendação da IA.

A alta transparência pode possibilitar que o decisor entenda o funcionamento da IA. Isso ajudaria a mitigar uma eventual resistência pessoal à IA, podendo reduzir a desconfiança em relação a ela, distensionando a percepção de risco (Mercado *et al.*, 2016). Falando especificamente dos sistemas de recomendação, esta percepção é condizente com os resultados de Cramer *et al.* (2008, p. 491), segundo quem – no contexto de decisões estratégicas - "a transparência aumentou a aceitação das recomendações". Entretanto, não há evidências de que – em decisões de naturezas distintas – a alta transparência gere a mesma curva de aceitação das recomendações da IA.

Em decisões operacionais, os gestores podem manifestar alguma aderência às recomendações do modelo de IA menos transparente porque não vislumbram grandes riscos e prejuízos caso a decisão se mostre equivocada posteriormente; ao passo que em decisões estratégicas, os gestores são mais resistentes a aderir às recomendações do modelo de IA menos transparente porque sabem (ou imaginam) que qualquer decisão incorreta pode ter consequências muito negativas, em diversos sentidos, e causar danos graves. O fato de saber que sua decisão terá impacto em toda a organização, que ela poderá determinar o cumprimento (ou não) dos objetivos organizacionais, e que ela repercutirá no longo prazo, torna o gestor mais cauteloso e exigente para dar sua anuência a quaisquer recomendações ou sugestões em matéria de decisões estratégicas. Este comportamento cauteloso, bem como a própria heurística e vieses dos gestores em decisões estratégicas, foi destacado por Busenitz e Barney (1997) e, posteriormente, corroborado por Young, Arthur Jr e Finch (2000) – que elencaram o temor de ser mal avaliado e a responsabilidade por algo como sendo um dos preditores do comportamento/decisão dos gestores.

Este tema da responsabilidade (*accountability*) se conecta ao aspecto cognitivo da tomada de decisão. A necessidade de avaliar circunstâncias, a tentativa de racionalizar incertezas, a ponderação sobre reflexos temporais, tudo isso são atividades cognitivas (RADERMACHER, 1994) que, ao exigir a intelecção humana, impõem sobre os ombros do decisor o peso da responsabilidade pelas escolhas feitas. A *accountability* tem relação direta com a formação do julgamento dos decisores sobre as situações (LIBBY; SALTERIO; WEBB, 2004) e é apresentada como um dos fatores que influenciam na aderência dos gestores a novas

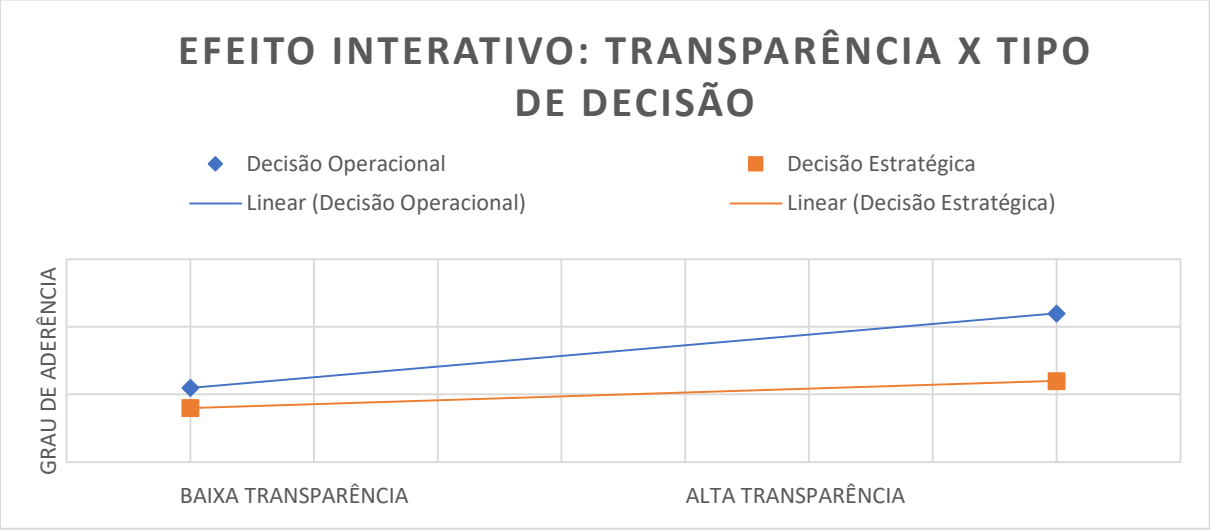
tecnologias (RYAN; BERGIN; WELLS, 2018). A ideia, já mencionada, de que "o algoritmo me fez fazer isto" (NIELSEN, 2020, p. 13), embora compreensível sob o ponto de vista puramente operacional de quem age em obediência ao fluxo que uma máquina orienta, não isenta o decisor de responder por seus atos. Assim, em decisões simples e corriqueiras (operacionais) ele assume mais facilmente o risco de seguir o algoritmo, ao passo que em decisões mais complexas e de longo alcance (estratégicas), ele se mantém mais reticente e conservador em seu posicionamento quando diante das recomendações de uma máquina.

Resumidamente: embora a alta transparência em geral aumente a aceitação das recomendações da IA, espera-se que, em decisões operacionais, esse aumento ocorra de forma mais acentuada que em decisões estratégicas. Deste entendimento, portanto, emana a terceira hipótese da presente pesquisa:

H3: O efeito positivo da alta transparência sobre a aderência dos gestores às recomendações do modelo de IA é maior para decisões operacionais do que para decisões estratégicas.

É possível visualizar o efeito que H3 sugere por meio da Figura 3, a seguir:

Figura 3 – Predição do efeito interativo entre transparência e tipo de decisão sobre a aderência dos gestores



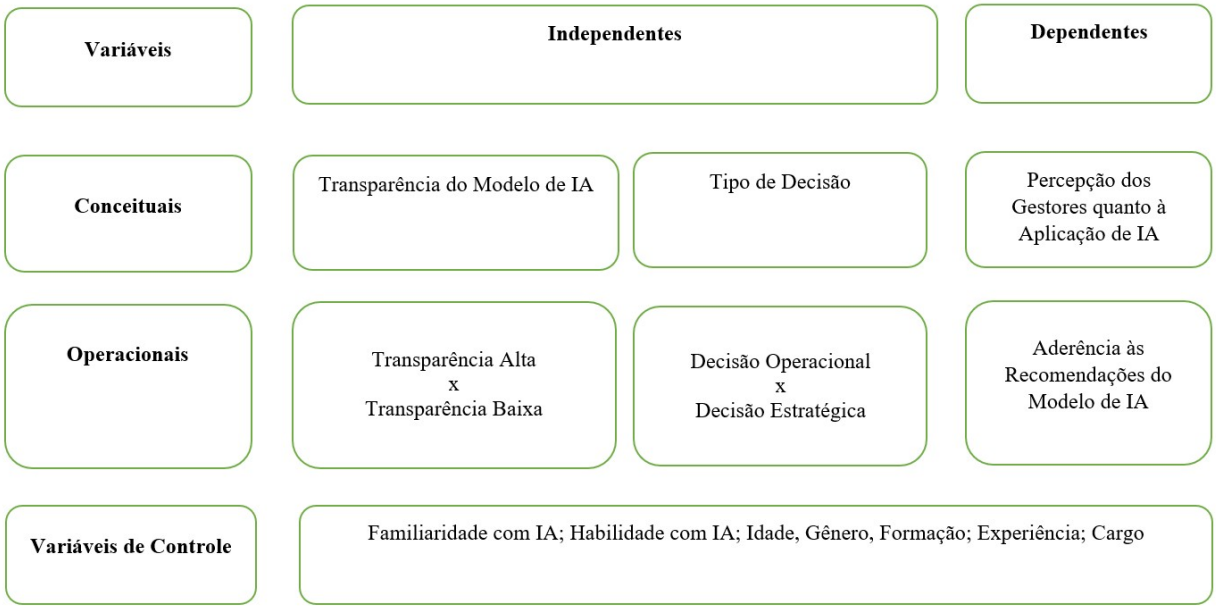
Fonte: elaboração própria.

3 PROCEDIMENTOS METODOLÓGICOS

3.1 CARACTERÍSTICAS GERAIS DO DESENHO DE PESQUISA

O *design* experimental foi o *between-participants* do tipo 2 x 2, ou seja: cada uma das duas variáveis testadas foi manipulada em dois níveis. Para este *design* não se fez necessário realizar a tarefa experimental de forma escalonada, isto é, em etapas. Portanto, cada participante teve apenas uma oportunidade de ler o *case* e responder a tarefa experimental. Este procedimento evitou que os participantes identificassem os propósitos específicos da pesquisa e acabassem por modular suas respostas em função disso. Para melhor visualização de todas as variáveis adotadas, foi construída a *Libby Box* a seguir:

Figura 4 – Esquematização das variáveis do estudo (*Libby Box*)



Fonte: elaboração própria.

As variáveis conceituais são os pilares teóricos a partir dos quais se sustenta toda a argumentação desenvolvida. A conjunção de tais variáveis foi rastreada a partir da confrontação entre diversos autores e estudos, conforme expresso no capítulo 2.

A variável Transparência foi operacionalizada em dois níveis: transparência alta e transparência baixa. O modelo de IA do tipo *Black Box*, pelas razões já anteriormente expostas, foi apresentado como parâmetro de modelo de baixa transparência. Já o modelo *White Box* representou a alta transparência da IA.

A variável Tipo de Decisão também foi operacionalizada em dois níveis: decisões operacionais e decisões estratégicas. A distinção entre estas decisões e os elementos específicos que caracterizam uma e outra foram apresentados no tópico 2.3. A decisão inserida em cada cenário, hipotética, respeita as características identificadas na literatura.

A variável dependente, que captura o efeito de manipulações nas variáveis independentes, é a percepção dos gestores quanto à aplicação da IA. Esta percepção se manifesta tanto na propensão (%) dos participantes em concordar com a recomendação do modelo de IA, quanto no valor ($) que eles atribuem à decisão a que foram submetidos. A propensão dá um indicativo da reação inicial dos participantes frente a uma proposta de utilização da IA, e o valor materializa a decisão deles diante desta possibilidade.

A operacionalização das variáveis independentes e da dependente encontra-se descrita nos tópicos 3.4 e 3.5, respectivamente; e os aspectos e justificativas específicos relacionados às variáveis de controle encontram-se descritos no tópico 3.7. Em geral, as opções metodológicas adotadas nesta tese foram reunidas e explicadas no Apêndice A.

A fim de eliminar possíveis vieses e evitar uma interpretação errônea dos resultados, também foi construído um quadro de tratamento de vieses (vide Apêndice C). Não foram oferecidos incentivos monetários, nem quaisquer brindes atrelados à participação na pesquisa.

3.2 PROTOCOLO ÉTICO

O Termo de Consentimento Livre e Esclarecido (TCLE), formalidade necessária na condução de pesquisas no Brasil, foi inserido e disponibilizado na primeira página do instrumento de coleta de dados. Após apresentar a proposta de estudo e o pesquisador que a conduz, o(a) participante foi questionado sobre sua voluntariedade, de forma que só poderia prosseguir acessando e respondendo ao experimento se a questão fosse respondida afirmativamente, o que manifestava sua adesão ao TCLE. Por oportuno, é mister destacar que o TCLE ora empregado seguiu os parâmetros e exigências estabelecidos pela Universidade Federal de Pernambuco (vide: https://www.ufpe.br/cep/manual-e-modelos) em congruência com as orientações da Comissão Nacional de Ética em Pesquisa.

3.3 PARTICIPANTES

Os participantes foram gestores, consultores e assessores, aleatoriamente alocados entre os grupos experimentais, que atuam em uma organização financeira brasileira. A decisão de

trabalhar com os profissionais deste tipo de organização fundamenta-se na intensidade e na velocidade da transformação digital que o setor financeiro tem vivenciado no Brasil e no mundo: só em 2021 o setor financeiro no Brasil investiu mais de 30 bilhões em tecnologia, sendo a inteligência artificial um dos tópicos prioritários de destinação de recursos (FEBRABAN, 2022). Assim, estas instituições mostram-se como campo adequado para a observação e depuração de fenômenos relacionados à introdução de IA no dia a dia de uma organização e, também, no seu processo decisório.

A população de participantes está lotada em 161 unidades táticas da empresa parceira da pesquisa (gerências regionais, gerências especializadas e superintendências) espalhadas por todo o país. Dentro desta estrutura organizacional, a área tática toma as decisões operacionais de larga escala, e também subsidia – com informações e pareceres – as decisões estratégicas que serão tomadas pelas diretorias, vice-presidências, e demais unidades e gerências do alto escalão. Desta forma, as unidades táticas tangenciam ambos os tipos de decisão abordados na presente pesquisa – o que ratifica sua adequação aos propósitos do experimento ora realizado.

O total de gestores lotados nestas unidades é de 13.544 pessoas (base: agosto/2022). Embora haja nestas unidades os mais diversos cargos, apenas os(as) gerentes, consultores(as) e assessores(as) são público-alvo do presente estudo, tendo em vista que apenas estes detêm poder/influência sobre as decisões da organização. Isto resulta em uma amostra de 2.245 pessoas. As atribuições dos cargos das unidades táticas são estabelecidas em normativo interno da instituição, válido em todo país. Assim, onde quer que se localize a unidade a que está vinculado(a) o(a) participante, as atribuições que lhe dizem respeito são as mesmas de qualquer um(a) outro(a) que desempenhe o mesmo cargo.

Visando minimizar falhas na execução da tarefa experimental, foram realizados três pré-testes (pilotos) por meio dos quais voluntários dispuseram-se a participar e conceder um *feedback* sobre o tempo de realização da tarefa, a facilidade (ou não) da navegação no instrumento de pesquisa, a linguagem empregada na descrição do cenário, a coerência (ou não) das questões formuladas, e sobre quaisquer outras impressões que tiveram com relação à sua participação no estudo. Os comentários foram utilizados para aprimorar a coleta de dados, de modo que a versão final do instrumento de pesquisa efetivamente distribuída aos participantes incorporou as adaptações e melhorias decorrentes destes pré-testes.

Os pré-testes também visaram averiguar eventuais falhas operacionais na ferramenta de coleta (no caso, o *software* Survey Monkey), avaliar a clareza das questões e informações prestadas, checar a saliência da manipulação das variáveis independentes, e por fim, em uma perspectiva mais qualitativa, colher impressões e sugestões de melhoria.

A Tabela 1 fornece mais detalhes acerca destes pré-testes:

Tabela 1 – Detalhamento dos pré-testes

Pré-Teste	Quantidade de participantes	Período de realização
1	20	Abril/2022
2	10	Maio/2022
3	13	Junho/2022

Fonte: dados da pesquisa.

Uma vez finalizada a coleta de dados, foram obtidas 239 respostas. Para consolidar a amostra final, entretanto, partiu-se dos 128 participantes que 0responderam acertadamente as questões de verificação da manipulação (vide tópico 3.8). Deste total, foram excluídas 26 pessoas que – ao atribuir valor à decisão solicitada na tarefa experimental – ficaram aquém ou além da faixa estabelecida no próprio cenário da tarefa (que variava de R$ 440.000,00 a R$ 485.000,00). A amostra final, portanto, contou com 102 participantes.

A partir da distribuição aleatória destes 102 participantes, verificou-se que os grupos experimentais ficaram assim formados (de acordo com a condição experimental a que foram submetidos), conforme Tabela 2:

Tabela 2 – Grupos experimentais

Condição experimental (COND)	Número de participantes (N)	Característica da COND
1	22	Transparência ALTA com decisão ESTRATÉGICA
2	25	Transparência ALTA com decisão OPERACIONAL
3	30	Transparência BAIXA com decisão ESTRATÉGICA
4	25	Transparência BAIXA com decisão OPERACIONAL

Fonte: dados da pesquisa.

A aleatorização na formação dos grupos experimentais foi bem-sucedida tendo em vista que a média de respostas a cada uma das variáveis sociodemográficas não diferiu de forma significativa entre as condições experimentais (Tabela 3).

Tabela 3 – Resultado dos testes de aleatorização

Variável	p-valor
Gênero	0.9467
Idade	0.6910
Formação	0.3300
Cargo	0.9014
Experiência	0.2023
Habilidade com IA	0.6786
Familiaridade com IA	0.4855

Fonte: dados da pesquisa.

Pelas respostas obtidas, apurou-se que os participantes eram predominantemente do sexo masculino (70,53%) e tinham em média 42 anos de idade. A maior parte (88,42%) indicou ser pós-graduado. Apenas dois respondentes informaram não ter graduação de nível superior. Embora sejam diversas as áreas de formação acadêmica dos participantes, três áreas concentraram mais de três quartos das respostas: ciências humanas (45,16%), ciências exatas (18,28%) e ciências sociais aplicadas (15,05%).

Embora, como já dito, todos os respondentes exerçam funções associadas à gestão tática da organização, a distribuição de frequências dos cargos foi bastante dispersa. Neste ínterim, tiveram discreta predominância os cargos de Gerente de Cobrança e Reestruturação de Ativos (21,05%), Assessor (21,05%) e Consultor (24,21%). Quanto ao tempo de cargo, 85,26% dos participantes informaram que ocupam seu posto atual há não mais que 5 anos. Entretanto, a maioria dos participantes (57,89%) tem entre 6 e 15 anos de experiência em cargos de gestão, e 80,00% dos participantes têm entre 6 e 20 anos de tempo de serviço na organização. Sete pessoas optaram por não responder às questões sociodemográficas e dois participantes optaram por não responder às questões sobre Habilidade e Familiaridade com IA. A Tabela 4 a seguir descreve de forma sumarizada os dados sociodemográficos dos participantes. Para as variáveis de natureza contínua, apresenta-se – além do número de participantes por categoria e da frequência percentual em 0relação ao todo – a média e o desvio-padrão.

Tabela 4 – Sumário de dados sociodemográficos dos participantes

	Variável		**n**	**%**	**Média (desvio-padrão)**
	Gênero	Masculino	67	70,53	-
		Feminino	28	29,47	
	Idade	-	-	-	41,8 (8,32)
Formação	Grau de Escolaridade	Ensino Médio/Graduação	11	11,58	-
		Pós-Graduação	84	88,42	
	Área de Formação	Ciências Exatas	17	18,28	-
		Ciências Humanas	42	45,16	
		Ciências Sociais Aplicadas	14	15,05	
		Outras	20	21,51	
	Cargo	Gerente	48	50,53	-
		Superintendente/Assessor	24	25,26	
		Consultor	23	24,21	
Experiência	Tempo no Cargo	-	-	-	3,37 (2,88)
	Tempo como Gerente	-	-	-	8,30 (6,02)
	Tempo na Empresa	-	-	-	15,94 (5,41)
	Grau de Habilidade com ferramentas de Inteligência Artificial	Nenhum	24	24,00	–
		Básico	23	23,00	
		Principiante	34	34,00	
		Intermediário ou mais	19	19,00	
	Familiaridade com a Inteligência Artificial	Nada ou pouco familiar	26	26,00	–
		Moderadamente familiar	59	59,00	
		Muito ou extremamente familiar	15	15,00	

Fonte: dados da pesquisa.

3.4 MANIPULAÇÃO DAS VARIÁVEIS INDEPENDENTES

O cenário apresentado a cada um dos participantes forneceu um breve conjunto de informações acerca da instituição financeira fictícia que estaria requerendo deles o preenchimento da tarefa experimental. A tarefa experimental (ver Apêndice D) levou em média 5 minutos para ser realizada.

A construção de uma situação hipotética alusiva ao tema específico da investigação, o interesse pelo julgamento subjetivo do participante, e a garantia de que não há respostas certas ou erradas são elementos que costumam estar presentes nos instrumentos de pesquisa que se propõem a validar propostas experimentais (HARTMANN; MAAS, 2010; HIRST; KOONCE; VENKATARAMAN, 2007; LEE; MORAY, 1992). Por isso, ao desenhar os cenários adequados aos propósitos específicos desta tese, estes elementos foram inseridos na ferramenta de coleta de dados.

Nos cenários, valendo-se de uma classificação comumente atribuída aos modelos de IA (LOYOLA-GONZALEZ, 2019), o modelo cujo processo utiliza uma lógica *White Box* (como regressão linear ou árvores de decisão) foi considerado de "alta transparência"; ao passo que o modelo que faz uso de uma lógica *Black Box* (como redes neurais) foi rotulado como de "baixa transparência". Assim, ao apresentar o modelo de alta transparência, o texto do cenário do instrumento de pesquisa faz referência a um sistema cujas "regras de cálculo utilizadas para prever gastos são estabelecidas por programadores" e cujo "funcionamento [...] depende de escolhas humanas pré-definidas" (vide Apêndice D). Por outro lado, ao manipular o cenário para apresentar o modelo de baixa transparência, o texto é adaptado e aponta para um sistema cujas "regras de cálculo utilizadas para prever gastos são estabelecidas pelo próprio sistema" e cujo "funcionamento [...] depende de associações aleatórias entre os dados que o alimentam" (vide Apêndice D).

Foram inseridas figuras para que, além da explicação dada no texto do cenário, ficasse evidente a distinção lógica entre os modelos. Para certificar-se de que tais figuras foram úteis no sentido de incrementar a saliência da manipulação realizada, foi feita uma consulta após a qual 6 (seis) voluntários respondentes informaram ao pesquisador suas impressões sobre as referidas ilustrações. Em geral, sinalizaram que as figuras efetivamente os ajudaram a compreender o texto. Apenas os dois que responderam por meio do celular reportaram que, embora conseguissem visualizar as figuras, seria melhor que elas fossem maiores. Isto posto,

aumentou-se a diagramação das figuras e, em uma nova consulta, realizada em seguida com outros seis voluntários, respondendo via celular, não houve mais solicitações neste sentido.

Em geral, a elaboração das figuras visou evidenciar que: **i)** no modelo tradicional o orçamento se baseava no histórico de gastos dos anos anteriores e era projetado com o auxílio do Microsoft Office Excel; **ii)** no modelo *white box* (denominado "Sistema Forest") o orçamento era fruto de uma multiplicidade de informações (internas e externas) que eram processadas a partir das regras fixadas em uma árvore de decisão por um ser humano; **iii)** no modelo *black box* (denominado "Sistema Spider") o orçamento era fruto das mesmas informações oferecidas ao Sistema Forest, porém processadas de forma diferente: sem intervenção humana, os dados se associavam aleatoriamente até formar redes neurais capazes de determinar qual deveria ser o orçamento para o período seguinte. A versão final de cada uma das figuras pode ser visualizada no Apêndice D.

No que concerne à manipulação da variável Tipo de Decisão, o instrumento de pesquisa também se valeu das características que a literatura correlata elenca para diferenciar as decisões operacionais das decisões estratégicas (DHAMIJA; BAG, 2020; HARRISON; PELLETIER, 1995, 2000; IMANE; DRISS, 2017; MISNI; LEE, 2017; NOORAIE, 2008; SHIVAKUMAR, 2014; TURNER, 2003; WARD; DURAY, 2000). No cenário apresentado ao participante, são assinalados três aspectos referentes à tipologia das decisões: o objeto da decisão; o horizonte de tempo ao qual ela se aplica; e a esfera/nível a que pertence o detentor do poder decisório dentro da organização. Objetivamente: para referir-se a uma decisão operacional o texto do cenário destaca que a responsabilidade pela decisão compete a "gestor(a) encarregado(a) da parte de suprimentos e estocagem", o qual toma "decisões de curto prazo sobre aquisições, reposições e alienações de bens móveis" (Vide Apêndice D). Doutro lado, para fazer menção a uma decisão de natureza estratégica, o texto do cenário menciona que a decisão está sendo atribuída a "Diretor(a) de Infraestrutura", a quem compete, "como membro do Alto Escalão da organização, decidir sobre investimentos na estrutura física da organização, com vistas à expansão dos negócios no longo prazo" (Vide Apêndice D).

3.5 MENSURAÇÃO DA VARIÁVEL DEPENDENTE

A tarefa experimental consistiu em responder a dois quesitos: o primeiro foi elaborado para mensurar o quanto os participantes estariam dispostos a concordar com as recomendações

do modelo de IA, partindo de uma escala de 0 a 100[4]. A captura de percepção mediante esta escala centesimal é frequentemente utilizada em outros estudos experimentais sobre tomada de decisão (CHENGALUR-SMITH; BALLOU; PAZER, 1999; HARTMANN; MASS, 2010; ISHIZAKA; SIRAJ, 2018; LIBBY; SALTERIO; WEBB, 2004). Ela se aplica a decisões com forte carga de subjetividade que, por isso mesmo, não podem ser materializadas em uma opção decisória binária nem em uma escala reduzida (como a *Likert Scale*).

O segundo quesito, nos moldes do que fizeram Bonollo e Zhang (2018) e Lang (2018), consistiu em colocar o decisor frente a um determinado valor para que possa arrematar um número – ratificando ou retificando o que lhe está sendo apresentado. O desvio entre o valor inicialmente recomendado e o valor final que o participante fixou em sua decisão é o objeto que o pesquisador irá apreciar.

No caso em tela, portanto, em termos de operacionalização estatística, a percepção dos gestores foi medida a partir da diferença entre a recomendação de valor dada pelo sistema tradicional (R$ 485.000) e o valor da resposta que cada participante atribuiu ao tomar sua decisão. A esta distância chamou-se aderência às recomendações da IA. Para evitar que valores extremos prejudicassem a mensuração desta aderência, estabeleceu-se um valor também para a recomendação feita pela IA: R$ 440.000. A resposta do participante, portanto, poderia oscilar dentro de uma faixa que tinha como patamar mínimo R$ 440.000 e como patamar máximo R$ 485.000. Assim, a diferença supramencionada [entre o valor sugerido pelo sistema tradicional e a resposta do participante], poderia ser de, no máximo, R$ 45.000, e quanto maior o valor desta diferença, mais o participante estaria se afastando da recomendação tradicional e aderindo à recomendação da IA.

Logo, a mensuração da aderência às recomendações da IA pode ser matematicamente enunciada da seguinte forma:

$$[A = 485.0000 – r]$$

Em que:

A = Aderência às Recomendações da IA;

[4] Como se verá adiante, no capítulo 4, as análises e comentários finais aos testes de hipóteses não levaram em consideração as respostas sobre esta probabilidade de concordância. A utilidade deste quesito foi apenas a de convalidar (ou não) os resultados obtidos a partir de outros parâmetros estatísticos – apresentados na sequência. Constatou-se, ao fim, que a intenção de concordar com as recomendações da IA nem sempre se confirma no momento da efetiva tomada de decisão – evidenciando uma dissociação entre o querer (ou não-querer) e o agir do participante [Simon (1987) já havia demonstrado e atribuído diversas causas a esta discrepância]. Apenas no que concerne à variável Tipo de Decisão a probabilidade de concordância com a IA comportou-se de modo congruente com a decisão efetivamente tomada.

r = Valor Indicado pelo Participante ao Tomar sua Decisão.

A congruência entre as respostas aos dois quesitos relativos à mensuração da variável dependente reflete e incrementa a validade interna das conclusões do experimento.

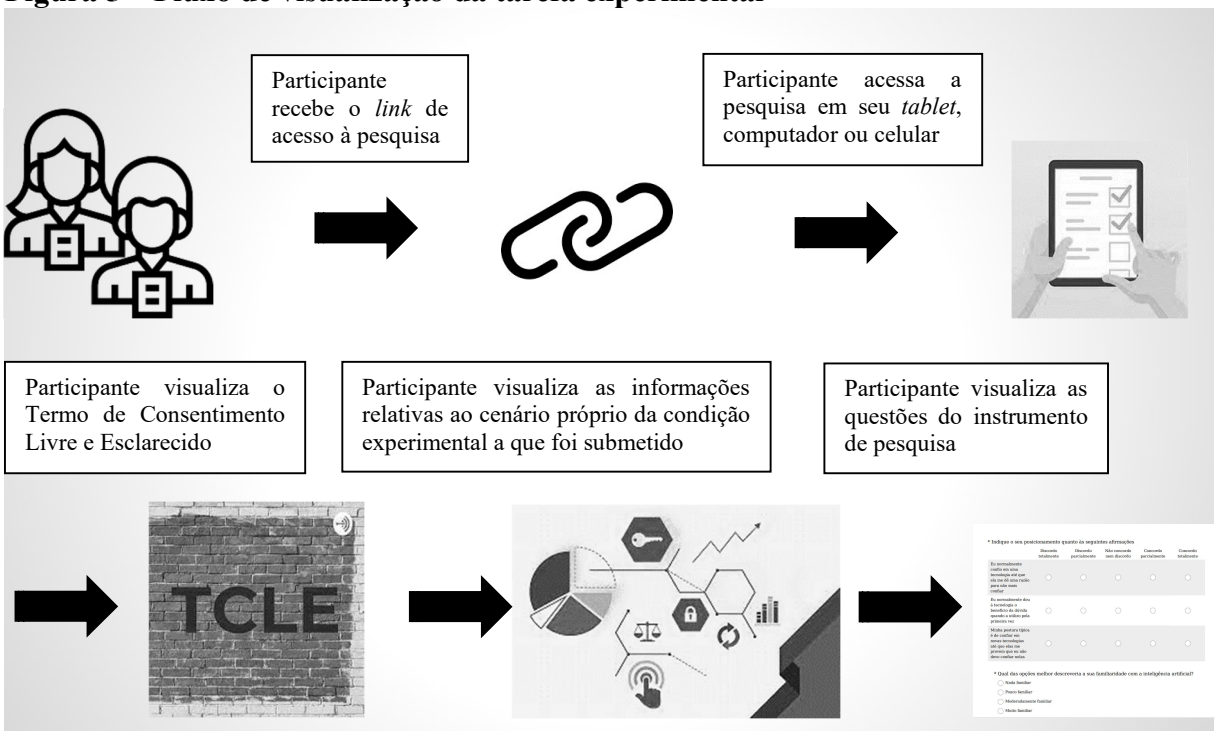
3.6 INSTRUMENTO DE PESQUISA

O instrumento utilizado para a coleta de dados consistiu em um formulário elaborado na plataforma Survey Monkey. O uso dessa plataforma permitiu que as respostas fossem coletadas mediante convite por *e-mail* e também permitiu que fossem apresentadas, de forma aleatória, diferentes versões de uma questão. Desta forma, foi possível agilizar o contato com o público-alvo ao mesmo tempo que se conseguiu randomizar, sem interferência do pesquisador, a formação dos grupos experimentais.

Para impulsionar a coleta de dados, foram feitos 5 disparos de *e-mail* (3 para os cargos de gestão e 2 para os cargos de consultoria e assessoria), com intervalo de cerca de 12 dias entre eles, contendo o *link* e o QR *Code* para acesso à pesquisa. A ferramenta do Survey Monkey não permite que um mesmo aparelho (seja ele celular, computador, tablet ou qualquer outro) responda à tarefa mais de uma vez. A aleatorização de qual cenário o respondente recebeu foi feita pela própria plataforma Survey Monkey, sem qualquer participação ou interferência do pesquisador, mediante aplicação da ferramenta “aleatorização de blocos”.

Do ponto de vista do participante, o acesso ao instrumento de pesquisa e à resposta à tarefa experimental se deu da seguinte forma: **i)** o(a) participante recebeu no seu *e-mail* institucional o *link* para acesso ao instrumento de pesquisa; **ii)** ao clicar no *link*, pôde acessá-lo no seu computador, celular ou tablet; **iii)** a primeira tela disponibilizada consistiu no Termo de Consentimento Livre e Esclarecido, ao fim do qual o(a) participante poderia optar por dar consentimento (prosseguindo assim com a navegação) ou negar consentimento (ocasião em que sua participação seria automaticamente encerrada); **iv)** tendo dado seu consentimento, a tela seguinte lhe traria um conjunto de informações contextuais, básicas e hipotéticas que constituíam o cenário específico da condição experimental a que foram submetidos; **v)** uma vez apresentado ao cenário, foram disponibilizadas a(o) participante as questões, sempre na mesma sequência, que seriam objeto de análise do pesquisador (o instrumento completo pode ser visualizado no Apêndice D). A Figura 5 ilustra este passo a passo:

Figura 5 – Fluxo de visualização da tarefa experimental



Fonte: elaboração própria.

3.7 VARIÁVEIS DE CONTROLE

As variáveis de controle, constituídas para garantir que os efeitos observados na variável dependente decorrem de manipulações feitas pelo pesquisador nas variáveis independentes e não de idiossincrasias presentes nos grupos experimentais, foram nomeadas, categorizadas e, a partir disso, testadas por meio de comparação entre as médias das suas categorias.

No presente estudo, as variáveis de controle “Familiaridade com a IA” e “Habilidade com a IA” foram extraídas e replicadas do estudo de Lang (2018). Com uma pergunta simples e direta para cada uma destas variáveis, oportunizou-se que o participante se autoavaliasse quanto a estes aspectos e, assim, fosse possível avaliar quanto das suas respostas à tarefa experimental poderia estar afetada por essa sua avaliação. No Apêndice D, a questão sobre a “Familiaridade com a IA” (Qual das opções melhor descreveria a sua familiaridade com a inteligência artificial?) e a sobre a “Habilidade com a IA” (Qual o seu grau de habilidade com ferramentas de inteligência artificial?) figuram imediatamente após os quesitos de verificação da manipulação.

As demais variáveis de controle, de cunho sociodemográfico, constituem o corpo de questões pós-experimentais que constam na última parte do Apêndice D. A Idade, o Gênero, o Cargo, e a Formação (depreendida a partir do Grau de Escolaridade e da Área de Formação)

são elementos frequentes em pesquisas de natureza similar à desta tese porque podem afetar a percepção dos respondentes – e por isso é preciso controlá-los. A experiência profissional (materializada em Tempo no Cargo Atual, Tempo em Cargos de Gestão e Tempo de Empresa), segundo Libby, Salterio e Webb (2004), também é uma variável de controle que figura em questionários de diversas modalidades de pesquisa e que tem sua razão de ser porque costuma impactar o *mindset* dos respondentes.

Com o propósito de testar todas estas variáveis de controle quanto a seu eventual efeito sobre a variável dependente, foram utilizadas diversas técnicas estatísticas: quando se tratou da variável Gênero, que possuía apenas duas categorias, foi utilizado o Test T; quando se tratou de variáveis categóricas com mais de duas categorias, fez-se uso da análise de variância (ANOVA); no caso das variáveis contínuas, o método empregado foi a Regressão Linear.

Uma vez realizados estes testes, à exceção do Gênero ($p = 0{,}0958$), as demais variáveis de controle de fato não apresentaram p-valor com significância estatística capaz de influenciar a mensuração da variável dependente. Especificamente, o que se observou é que as mulheres têm uma média de aderência às recomendações da IA superior à média de aderência dos homens. Dada a significância deste resultado, foi necessário incluir o "Gênero" como covariável quando da posterior realização dos testes de hipóteses e análises suplementares.

3.8 CHECAGEM DA MANIPULAÇÃO

De acordo com Libby, Bloomfield e Nelson (2002) é oportuno e necessário inserir questões no instrumento de coleta de dados para verificar se os participantes interpretaram corretamente as variáveis independentes, bem como para avaliar a eficácia da manipulação realizada por meio da tarefa experimental proposta. No caso presente, foram inseridas duas questões: uma para checar a manipulação da variável "Transparência" ("As regras que a ferramenta de inteligência artificial utiliza para prever valores são...") e a outra para checar a manipulação da variável "Tipo de Decisão" ("O tipo de decisão que me foi solicitada...").

As perguntas foram apresentadas ao(à) participante imediatamente após ele(a) ter respondido aos quesitos da tarefa experimental propriamente dita, isto é, depois de ele ter informado qual a sua probabilidade de seguir a recomendação da IA e de ter se posicionado quanto ao valor (financeiro) que atribuiu à decisão que lhe foi requerida.

Para a checagem da manipulação da variável Transparência, esperava-se que o participante que recebeu o cenário de alta transparência, identificasse as características desta condição e indicasse que as regras da ferramenta de IA que constava no cenário que lhe foi

apresentado haviam sido “estabelecidas por programadores e, portanto, dependem(iam) de escolhas humanas predefinidas”. Para a baixa transparência, a expectativa era de que o respondente assinalasse que tais regras foram “estabelecidas pelo próprio sistema e, portanto, dependem(iam) de associações aleatórias entre os dados”. Tais trechos foram retirados *ipsis litteris* do texto do cenário – o qual, como já dito, continha figuras para facilitar o entendimento do participante quanto ao funcionamento do modelo de IA.

Da mesma forma, a checagem da manipulação da variável Tipo de Decisão foi construída a partir de excertos do cenário. Nesse caso, dos participantes que foram incumbidos de tomar uma decisão operacional esperava-se que associassem tal decisão a uma característica de “curto prazo e (que) visa garantir a disponibilidade de bens móveis utilizados nas atividades cotidianas”. Em relação aos que receberam um cenário de decisão estratégica, a expectativa era de que caracterizassem tal decisão como sendo “de longo prazo e (que) visa garantir a expansão dos negócios em um cenário de concorrência”.

Considerando o total de respostas obtidas (239), 64,43% dos participantes identificaram corretamente a manipulação da Transparência, e 83,26% identificaram corretamente a manipulação do Tipo de Decisão. Diante disso, 111 participantes foram excluídos da amostra por responder de forma incorreta algum (ou ambos) os quesitos de *manipulation check*.

Em estudos experimentais similares, mesmo tendo realizado pré-testes para validar o instrumento de pesquisa, respostas incorretas podem ocorrer tanto por desatenção dos participantes, quanto por interpretação equivocada por parte deles. Para lidar com isso, costuma-se desprezar tais respostas e utilizar apenas as de quem identificou corretamente as manipulações (CHENG; COYTE, 2014; ROSE *et al*., 2014). Desta forma os resultados, a conclusão, e as eventuais inferências não ficam comprometidas. Como assinalou Oppenheimer, Meyvis e Davidenko (2009, p. 871): “eliminar os participantes que estão respondendo aleatoriamente [...] aumentará a relação sinal-ruído e, por sua vez, aumentará poder estatístico”. Este expediente foi usado por Lang (2018), que obteve apenas 28,00% de respostas corretas quanto à manipulação, por Coram e Wang (2020), que obtiveram 57,50% de respostas corretas quanto à manipulação, e também por Oppenheimer, Meyvis e Davidenko (2009), que obtiveram 54,00% de respostas corretas quanto à manipulação.

4 RESULTADOS

A análise estatística dos dados foi realizada utilizando-se o *software* estatístico Stata – versão 14.2. Utilizou-se a Análise de Covariância (ANCOVA) para comparar as médias das variáveis independentes (Transparência da IA e Tipo de Decisão), nos seus diferentes níveis, em relação à variável dependente (Aderência às Recomendações da IA). O Gênero foi utilizado como covariável em razão do impacto estatisticamente significante que ele causa na Aderência às Recomendações da IA. A variável Confiança na Tecnologia, adicionada e explanada em um tópico suplementar, foi objeto de uma análise de mediação segundo a Modelagem de Equações Estruturais. Para uniformizar a nomenclatura das variáveis, a partir de agora a variável Transparência da IA será grafada tão somente como Transparência, a variável Tipo de Decisão será grafada como Decisão, a variável Aderência às Recomendações da IA será grafada como Aderência, e a Variável Confiança na Tecnologia será grafada como Confiança.

4.1 ANÁLISE DESCRITIVA

A Tabela 5, abaixo, apresenta as estatísticas descritivas concernentes à Aderência nas diversas condições experimentais em que ela foi mensurada. O exame das médias e desvios a seguir oferece uma visão do comportamento da variável dependente nos diversos cenários e enfoques que esta pesquisa se propôs a examinar.

Tabela 5 – Análise descritiva da Aderência de acordo com as Condições Experimentais

Decisão		**Transparência**		
		Alta	**Baixa**	**Total**
Operacional	Média	35.600	36.400	36.000
	Desvio Padrão	15.365	13.864	14.489
	Número de Participantes	[25]	[25]	[50]
Estratégica	Média	26.591	36.667	32.404
	Desvio Padrão	16.912	11.396	14.727
	Número de Participantes	[22]	[30]	[52]
Total	Média	31.383	36.545	34.167
	Desvio Padrão	16.565	12.458	14.650
	Número de Participantes	[47]	[55]	[102]

Fonte: dados da pesquisa.

Considerando que a extensão da medida de Aderência compreendia desde a concordância total com a recomendação do sistema tradicional (Aderência = 0) até a plena concordância com a recomendação da IA (Aderência = 45.000), percebe-se, em primeiro lugar, que em todas as condições experimentais a média de Aderência está mais próxima da plena

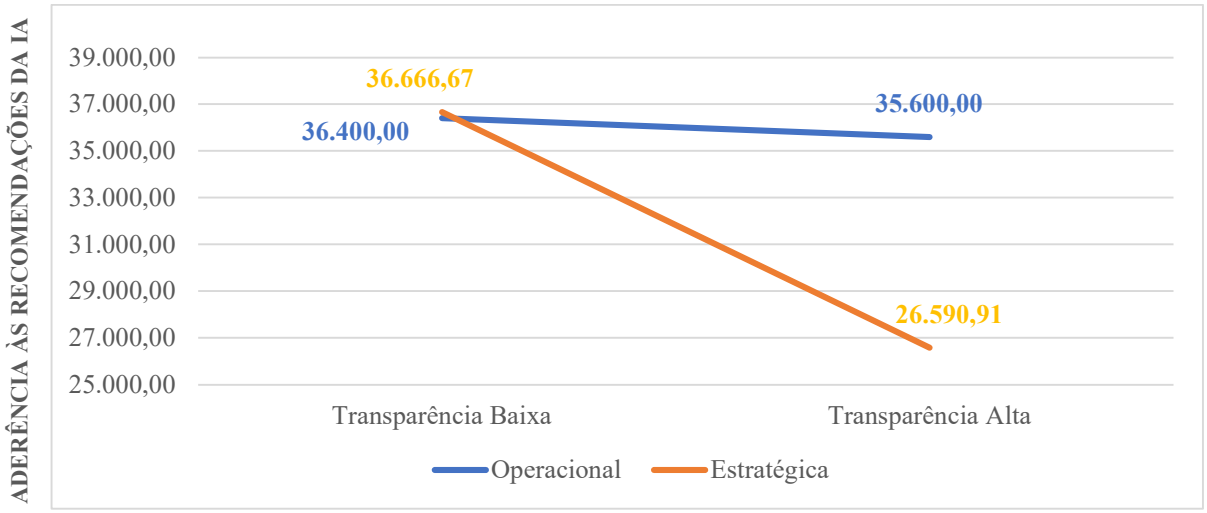
adesão à recomendação da IA do que do proposto pelo sistema tradicional. Tal constatação pode ser sinal de que há entre os participantes como um todo certa inclinação "pró-IA" ou, pelo menos, pouca resistência a ela.

Em segundo lugar, nota-se que o comportamento da variável Transparência aponta para uma conclusão que destoa da argumentação teórica desenvolvida no capítulo 2 e consignado na hipótese H1. De acordo com os dados obtidos, a média de Aderência em cenário de alta transparência ($\overline{x}$ = 31.283) foi menor que a média de Aderência na baixa transparência ($\overline{x}$ = 36.545).

Por outro lado, em linha com o que preconizou a hipótese H2, os gestores aderem mais às recomendações da IA quando estão diante de decisões operacionais ($\overline{x}$ = 36.000) do que quando estão diante de decisões estratégicas ($\overline{x}$ = 32.403).

Retomando as condições experimentais já apresentadas no capítulo anterior, nota-se que foi a condição experimental 1 (transparência alta com decisão estratégica) que apresentou o menor nível de Aderência ($\overline{x}$ = 26.590), enquanto que a condição 3 (transparência baixa com decisão estratégica) apresentou o maior nível de Aderência ($\overline{x}$ = 36.666). Este resultado, contrasta não apenas com o que foi postulado em H3, mas também com a ideia levantada em H1 (de que mais transparência leva a mais aderência). A predição da hipótese H3 era de que a maior aderência seria verificada na condição experimental 2 (transparência alta com decisão operacional) – fato que não se confirmou. O gráfico a seguir consolida e ilustra o comportamento da Aderência nos diversos cenários propostos na tarefa experimental:

Gráfico 1 – Média de Aderência de acordo com as Condições Experimentais



Fonte: dados da pesquisa.

4.2 TESTE DAS HIPÓTESES DO ESTUDO

A hipótese H1, desenvolvida no capítulo 2 da presente tese, previra que a Aderência seria maior quando a Transparência do referido modelo fosse alta do que quando fosse baixa. O objetivo desta hipótese era capturar o efeito da Transparência sobre a Aderência, supondo uma relação direta entre estas duas variáveis, ou seja: mais transparência, mais aderência; menos transparência, menos aderência.

Ao aplicar ANCOVA para testar a hipótese H1[5], verificou-se que os resultados indicam que a Transparência afetou de forma significativa a Aderência (F = 3.96; p = 0.0496), – conforme descrito no Painel A da Tabela 6, a seguir:

Tabela 6 – Resultados dos Testes de Hipóteses

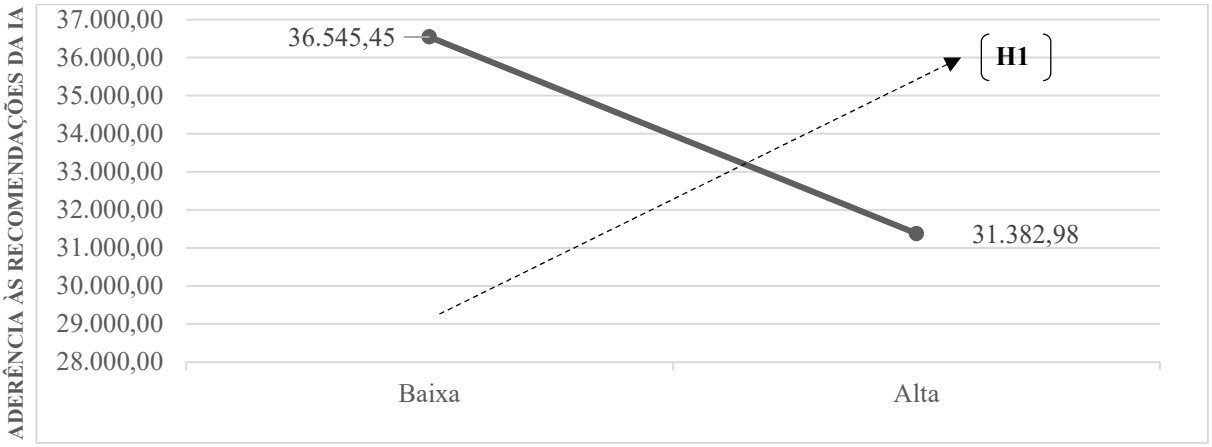
Fator	gl	SQ	F	p-valor
Painel A: Análise dos Efeitos Diretos e Interativo (variável dependente = Aderência)				
Modelo	4	6.170	3.02	0.0220
Transparência	1	8.099	3.96	0.0496**
Tipo de Decisão	1	5.157	2.52	0.1158
Transparência # Tipo de Decisão	1	6.627	3.24	0.0752***
Gênero	1	7.057	3.45	0.0665***
Residual	**90**	**2.045**		
Painel B: Análise dos Efeitos Simples (variável dependente = Aderência)				
Efeito simples do tipo de decisão na alta transparência	1	1.137	4.46	0.0407**
Efeito simples do tipo de decisão na baixa transparência	1	4732873.2	0.03	0.8654
Efeito simples da transparência na decisão operacional	1	1947529.9	0.01	0.9244
Efeito simples da transparência na decisão estratégica	1	1.449	7.35	0.0095*
Total	**94**	**2.221**		

*, ** e *** = significante ao nível de confiança de 1%, 5% e 10%, respectivamente.
gl= graus de liberdade; SQ = soma dos quadrados; F = frequência.
Fonte: dados da pesquisa.

Porém, contrariando a predição de H1, de acordo com os dados obtidos, no cenário de alta transparência a média de Aderência foi menor que a média no cenário de baixa transparência (vide Gráfico 2). Os resultados sugerem, portanto, que o efeito principal da Transparência sobre a Aderência dos gestores aponta para uma direção oposta à da previsão teórica formulada. Desta forma, rejeita-se a hipótese H1.

[5] O teste adotado para a apuração e análise do p-valor foi unicaudal. Isto porque as hipóteses estabelecidas foram direcionais, isto é, foram construídas na expectativa que determinado grupo apresente resultado superior/inferior a outro(s) grupo(s). Em experimentos do tipo 2 x 2 o uso deste tipo de teste aumenta o poder estatístico da análise (KACHELMEIER *et. al.,* 2016).

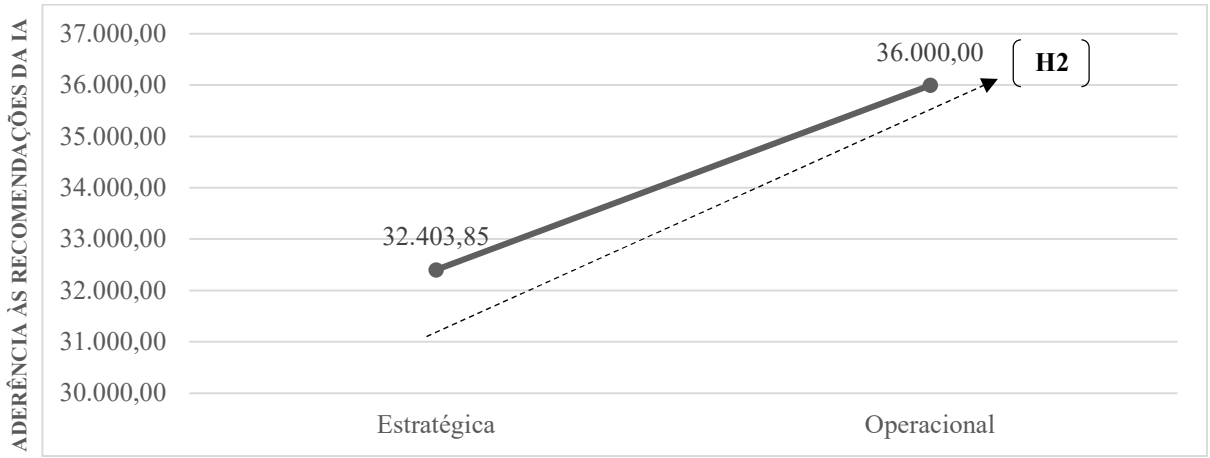
Gráfico 2 – Média de Aderência de acordo com a Transparência



Fonte: dados da pesquisa.

Quanto à hipótese H2, apesar de a Aderência no caso de decisões operacionais ter sido efetivamente maior que a média no caso de decisões estratégicas (vide Gráfico 3), o resultado do teste ANCOVA aplicado a esta hipótese não apresentou significância estatística para que tal hipótese fosse aceita. Os resultados indicam (vide Tabela 6, Painel A) que a Decisão não está associada de forma estatisticamente significativa (F = 2.52; p = 0.1158) com a Aderência.

Gráfico 3 – Média da Aderência de acordo com a Decisão



Fonte: dados da pesquisa.

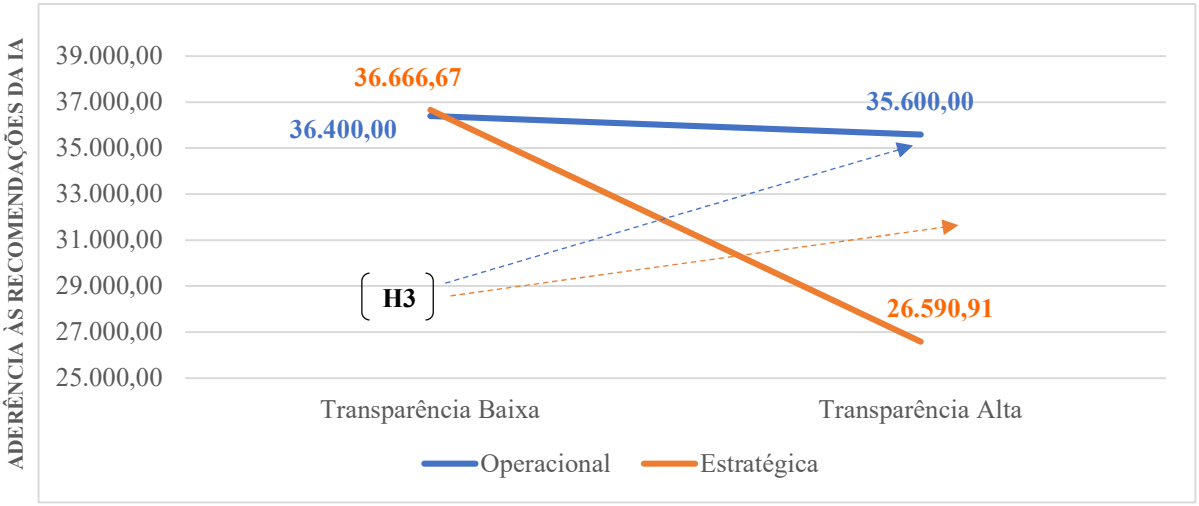
Em se tratando da hipótese H3, o que se constatou é que a interação entre Transparência Decisão afetou de forma significativa (F = 3.24; p = 0.0752) a Aderência (vide Tabela 6, Painel A). Entretanto, de acordo com os dados obtidos, o efeito da alta transparência sobre a Aderência foi negativo (isto é: quanto mais transparência, menos aderência).

Ao analisar os efeitos simples de cada nível das variáveis independentes (vide Tabela 6, Painel B), nota-se que a Decisão teve impacto significativo na Aderência em condições de alta transparência (F = 4.46; p = 0.0407); e a Transparência, por sua vez, afetou de forma significativa a Aderência quando aplicada a decisões estratégicas (F = 7.35; p = 0.0095). Estes resultados, se tomados em conjunto, fornecem indícios sobre as condições em que a interação entre as variáveis independentes pode afetar a variável dependente – algo que será objeto de um exame mais aprofundado quando da análise de H3.

Por outro lado, a análise de efeitos simples também fez notar que quando a Transparência é baixa a média de Aderência não apresenta diferenças estatisticamente significantes (F = 0.03; p = 0.8654) – quer seja em contexto de decisão operacional, quer seja em contexto de decisão estratégica. Além disso, em decisões operacionais percebe-se que a Transparência, seja alta ou baixa, não impacta de forma significativa a Aderência (F = 0.01; p = 0.9244).

Embora tal implicação tenha sido percebida tanto no contexto de decisões operacionais quanto no contexto de decisões estratégicas, este impacto negativo da alta transparência foi bem mais acentuado quando o cenário trazia uma decisão estratégica (vide Gráfico 4).

Gráfico 4 – Média de Aderência na interação entre Transparência e Decisão



Fonte: dados da pesquisa.

Os resultados do teste aplicado à H3, portanto, embora evidenciem o fenômeno da interação entre as variáveis independentes, e seu significativo impacto sobre a variável dependente, sugerem que a ponderação entre Transparência e Decisão leva a Aderência a um caminho distinto daquele apontado pela literatura e consignado em H3.

Assim, posto que as evidências do efeito interativo não seguem o mesmo sentido da previsão teórica desenvolvida, não é possível afirmar – como previra H3 – que o efeito positivo da alta transparência sobre a aderência dos gestores é maior em decisões operacionais que em decisões estratégicas.

4.3 ANÁLISE SUPLEMENTAR

O escopo teórico adotado na presente tese sugere que a Transparência é um fator que contribui efetivamente para que os gestores manifestem Aderência. Paralelamente, a vertente de estudos que trata da *Human-Computer Interaction* indica que, para que as sugestões e propostas da IA sejam seguidas, é preciso haver certo nível de confiança dos usuários na tecnologia em questão.

A Transparência costuma ser percebida como um indicador da confiabilidade dos modelos de inteligência (GILKSON; WOOLLEY, 2020; PREECE, 2018); e a confiabilidade – por sua vez – conta positivamente perante os gestores quando têm que decidir entre aderir ou não a uma recomendação de IA. É pertinente esclarecer: a confiabilidade é uma mensagem que o *design* do modelo quer transmitir, ao passo que a confiança é percepção que se origina no gestor quando recepciona tal mensagem. A confiabilidade, portanto, diz respeito ao modelo, enquanto a confiança diz respeito ao gestor. Assim, o encadeamento Transparência – Confiança – Aderência, amplamente referendado, ampara e justifica a argumentação ora apresentada sobre a mediação da variável Confiança.

A respeito do fator Confiança, a literatura apresenta um fenômeno chamado de *distrust* (desconfiança) e o vincula a um efeito nomeado como *disuse* (desuso) (CHIEN *et al.*, 2014; HANCOCK *et al.*, 2011). Isto é: a falta de Confiança está associada a certa resistência ao uso de determinados modelos tecnológicos. Além disso, se falta ao usuário a pré-disposição para confiar (*initial trust*) (confiança inicial), a adesão a uma nova tecnologia também fica comprometida (LEWIS; SYCARA; WALKER, 2018; MILLER, 2015). Por fim, se houver problemas na calibragem da Confiança (*trust calibration*) (calibragem da confiança), a decisão do usuário sofre certo enviesamento (à revés da nova tecnologia, naturalmente) (GLIKSON; WOOLLEY, 2020; LEWIS; SYCARA; WALKER, 2018).

Assim, uma experiência anterior negativa com a IA, por exemplo, ou mesmo uma percepção pessimista quanto à sua capacidade de subsidiar/orientar determinada decisão, pode afetar a disposição dos usuários em aceitar e utilizar um modelo de IA (ainda que ele seja altamente transparente), devido à quebra ou à falta de Confiança. Todos estes fenômenos,

quando observados sob o prisma da relação entre os humanos e as máquinas, dão conta de que o caminho que conecta a transparência de um modelo de IA à aderência que os humanos manifestam em relação a ele não é uma via expressa, mas sim uma via que perpassa necessariamente o quesito confiança.

Esta Confiança não se restringe a mero fator moderador da relação Transparência-Aderência, cuja presença em maior ou menor grau redundaria em maior ou menor aderência. Mesmo porque, como Trojanowski e Kułak (2017) mostraram, nem sempre confiar mais em uma tecnologia intensifica ou amplia o uso da mesma. A Confiança constitui-se, na verdade, em elemento fundamental de mediação porque vincula-se tanto à variável independente (Transparência) quanto à variável dependente (Aderência). Em relação à primeira, a confiança se apresenta como consequência: Transparência gera Confiança; e em relação à segunda, ela se apresenta como causa: Confiança gera Aderência. Portanto, sob este ponto de vista, a Confiança não é apenas um indutor ou redutor da Aderência (como se fora variável moderadora), mas sim uma conexão lógica essencial entre a Transparência e a Aderência (variável mediadora).

Neste sentido, suplementando o presente estudo experimental, foi feita uma análise de mediação da Confiança utilizando os mesmos grupos, respostas e dados obtidos na aplicação da tarefa experimental aqui proposta.

A Confiança foi mensurada a partir da escala desenvolvida e aplicada por Höddinghaus, Sondern e Hertel (2021) e abrange a fidúcia na tecnologia em geral. Entender esta variável a partir desta perspectiva evita que se associe equivocadamente o grau de (Des)Confiança dos participantes a uma suposta aversão / atração específica relacionada à IA, quando na verdade a (Des)Confiança de certos respondentes se manifesta em relação a quaisquer formas de tecnologia – e não apenas em relação à IA.

A mensuração da Confiança se deu a partir de três questões (“Eu normalmente confio em uma tecnologia até que ela me dê uma razão para não mais confiar”; “Eu normalmente dou à tecnologia o benefício da dúvida quando a utilizo pela primeira vez”; “Minha postura típica é de confiar em novas tecnologias até que elas me provem que eu não devo confiar nelas”) que se encontram no Apêndice D logo após o quesito que apura a “Habilidade com a IA”. Foi aplicada a técnica de Análise de Componente Principal às respostas aos quesitos supramencionados (codificados como conftec1, conftec2 e conftec3, conforme Apêndice E), a fim de consolidar uma medida única de confiança. A tradução da escala de Höddinghaus, Sondern e Hertel (2021), do inglês para o português, foi feita aplicando-se a técnica de *back translation* para garantir que o texto traduzido tivesse a maior fidelidade possível à ideia original da escala.

Isto posto, foi realizada a Análise de Variância (ANCOVA) para verificar o impacto das variáveis independentes sobre a Confiança (vide Tabela 7). Embora houvesse apenas duas categorias/níveis para cada variável independente, a inserção do Gênero como covariável fez com que a aplicação do Teste T fosse preterida em relação à ANCOVA. Assim, constatou-se, com relação à Transparência, que ela influencia de forma significativa (p = 0.012) a formação da Confiança, ao mesmo tempo em que se demonstrou que a alta transparência está associada a médias de confiança inferiores às apuradas na baixa transparência. Com respeito à Decisão, porém, não se verificou um impacto relevante (p = 0.560) sobre a Confiança.

Na sequência, o teste ANCOVA da Confiança entre as condições experimentais propostas (vide Tabela 7) mostrou que as diferenças de média entre as quatro condições foram estatisticamente significativas (p = 0.013). Especificamente: a condição experimental 1 (transparência alta com decisão estratégica) difere significativamente das demais condições e apresenta a menor média de Confiança e o maior desvio padrão.

Tabela 7 – Confiança de acordo com a Transparência, Decisão e Condição Experimental

Variável	Nível	N	Média (Desvio-Padrão)	p-valor
Transparência	Baixa	54	0.27 (0.17)	0.012**
	Alta	46	-0.31 (0.19)	
Tipo de Decisão	Operacional	48	0.11 (0.15)	0.560
	Estratégica	52	-0.10 (0.21)	
Condição Experimental	Transparência alta com decisão estratégica	22	-0.67 (0.33)	0.013**
	Transparência alta com decisão operacional	24	0.01 (0.21)	
	Transparência baixa com decisão estratégica	30	0.31 (0.26)	
	Transparência baixa com decisão operacional	24	0.21 (0.22)	

*, ** e *** = significante ao nível de confiança de 1%, 5% e 10%, respectivamente.
Fonte: dados da pesquisa.

Paralelamente, mediante aplicação da técnica de regressão linear, utilizando ainda a variável Condição Experimental como preditora de Confiança, foi confirmada a significância da correlação entre estas duas variáveis (p = 0.015), ao mesmo tempo em que se ratificou a condição experimental 1 como sendo a única cuja associação com a Confiança mostrou-se estatisticamente significante (p = 0.054)

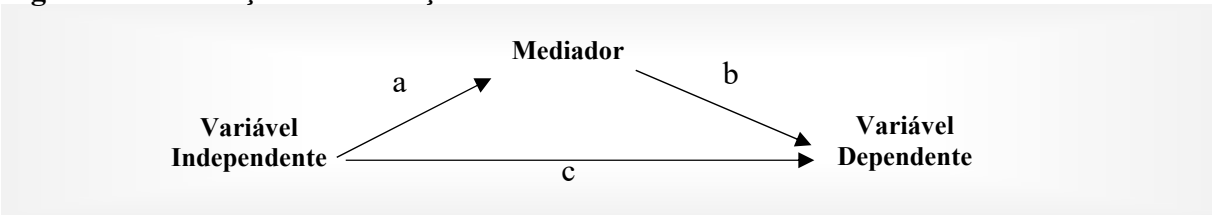
Na sequência, considerando que tanto a variável Confiança quanto a variável Aderência são de natureza contínua, foi feita uma análise, via regressão linear, para testar a força da associação entre estas variáveis. O resultado obtido sugere uma forte correlação entre elas (p = 0.000).

Em posse destes resultados preliminares (segundo os quais a Transparência e a Condição Experimental podem afetar a Confiança, e esta encontra-se fortemente associada à

Aderência), adotou-se a metodologia de Baron e Kenny (1986) para conduzir a análise da tríade Transparência-Confiança-Aderência, visando averiguar de modo sistemático a ocorrência de um fenômeno mediador.

De acordo com este método, a função mediadora é "um organismo ativo que intervém entre o estímulo e a resposta" (BARON; KENNY, 1986, p. 1776); ela é desempenhada, portanto, por uma "uma terceira variável, que representa o mecanismo gerador por meio do qual a variável independente focal é capaz de influenciar a variável dependente de interesse" (BARON; KENNY, 1986, p. 1173). A mediação pode ser assim ilustrada (Figura 6):

Figura 6 – Ilustração da mediação



Fonte: Baron e Kenny (1986, p. 1176).

Assim, ao testar (via regressão linear simples) a relação entre a variável independente (Transparência) e a variável supostamente mediadora (Confiança), notou-se que variações no nível da transparência afetaram de forma significativa ($p = 0.012$) a média de Confiança. Constitui-se, então, o primeiro indício de um possível fenômeno mediador.

Na sequência, ao regredir a Aderência tendo a Confiança como preditor, verificou-se a ocorrência de um efeito direto estatisticamente significante ($p = 0.000$), o qual atesta o efetivo impacto da variável mediadora sobre a variável dependente. O segundo elo da corrente de mediação ficou, assim, estabelecido.

Finalmente, ao regredir a Aderência utilizando a Transparência como preditor, constatou-se a produção de um efeito também significante do ponto de vista estatístico ($p = 0.066$). Além disso, utilizou-se também a Transparência e a Confiança, conjuntamente, como regressores da Aderência e a significância da Confiança persistiu ($p = 0.001$). Fechou-se assim, o ciclo da mediação. É possível, portanto, afirmar que, de acordo com os dados obtidos, a Confiança exerce, sim, a função de variável mediadora entre a Transparência e a Aderência.

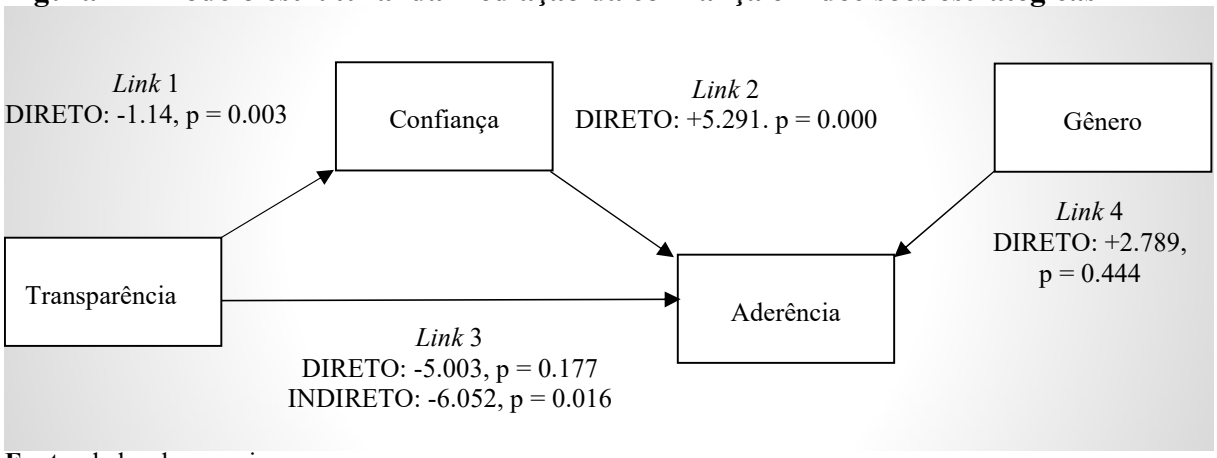
Para ratificar este entendimento e verificar se o fenômeno da mediação ocorre em ambos os cenários decisórios (isto é: tanto em decisões estratégicas quanto em decisões operacionais), procedeu-se também à análise por meio de equações estruturais. Esta análise demonstra as relações entre as principais variáveis abrangidas pelo estudo, ilustrando com simplicidade relações subjacentes geralmente complexas. A relação de mediação da Confiança é, no caso em

tela, a estrutura que conecta a variável independente (Transparência) à dependente (Aderência) em uma realidade específica.

Estatisticamente, a base dos modelos de equação estrutural também são regressões lineares. O foco da sua análise, contudo, está voltado para o exame dos efeitos diretos e indiretos provocados a partir da interação entre as variáveis. Nesse sentido, em termos de interpretação, é importante perceber dois elementos: o sinal da relação e a significância (p-valor) dela. O sinal indica a forma como um componente impactará o outro, e a significância demonstra a força estatística da ligação. Covariáveis também podem ter seu papel evidenciado neste modelo.

Assim, o modelo de equação estrutural foi primeiramente testado nos cenários em que o tipo de decisão era de natureza estratégica. De forma resumida (vide Figura 7), pode-se dizer que o efeito mediador da variável Confiança foi observado, pois: **i)** a Transparência teve associação negativa mas estatisticamente significante em relação à Confiança (p = 0.003); **ii)** a Confiança afeta de modo positivo e também significativo a Aderência (p = 0.000); **iii)** a Transparência afeta negativamente – de forma indireta, mas estatisticamente significante (p = 0.016) – a Aderência. Vale ressaltar quanto a este último ponto que a significância do efeito indireto é exatamente o que se espera: ela atesta que o efeito final da variável independente sobre a dependente não repercute de forma direta, mas sim mediada. Por fim, nesta modelagem, o Gênero – como já explicado anteriormente – foi mantido como covariável mas não demonstrou significância estatística na análise de mediação aplicada a decisões estratégicas.

Figura 7 – Modelo estrutural da mediação da confiança em decisões estratégicas

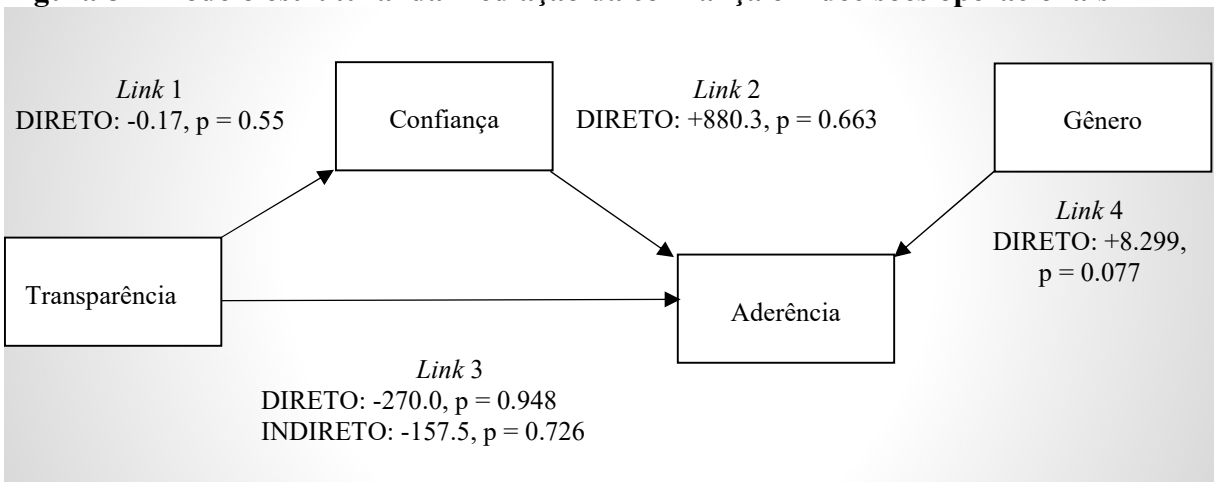


Fonte: dados da pesquisa.

No que concerne a decisões operacionais, contudo, aplicando este mesmo método de análise a partir de equações estruturais, a mediação da Confiança não ficou comprovada: tanto a sua conexão com a Transparência quanto o seu impacto sobre a Aderência não apresentaram significância estatística. Apenas, com relação ao Gênero, pode-se destacar que em decisões

operacionais esta variável parece produzir um efeito direto significativo (p = 0.077) sobre a Aderência. Complementando o que já foi comentado sobre esta variável no capítulo anterior, o resultado sugere que em decisões operacionais as mulheres manifestam Aderência significativamente maior que os homens. A Figura 8, a seguir, apresenta o modelo estrutural da mediação da Confiança em decisões operacionais.

Figura 8 – Modelo estrutural da mediação da confiança em decisões operacionais



Fonte: dados da pesquisa

Os dados obtidos, portanto, indicam que a Confiança satisfaz as condições necessárias e suficientes para que uma variável seja considerada mediadora. Tanto a metodologia de Baron e Kenny (1986) quanto a análise via equações estruturais validam esta afirmação. Reitere-se, contudo, que o papel mediador que a Confiança desempenha entre a Transparência e Aderência, e que restou demonstrado na presente tese, está circunscrito ao âmbito do uso de IA como subsídio para decisões estratégicas. Ressalve-se também que, uma vez que o impacto da Transparência sobre a Aderência em decisões estratégicas não foi exatamente igual a 0, pode-se inferir que – além da Confiança – deve haver outros fatores de mediação não capturados pelo modelo ora apresentado (BARON; KENNY, 1986).

Os principais achados relacionados às hipóteses H1 e H3, quando concatenados ao argumento da mediação, sugerem que, em decisões estratégicas, a alta transparência impacta negativamente a Confiança (maior transparência, menor confiança), ao passo que a Confiança mantém uma associação direta e positiva com a Aderência (maior confiança, maior aderência). É preciso destacar a relevância deste achado tendo em vista que, pelo menos com relação à sua primeira parte, ele contradiz frontalmente uma crença aparentemente razoável e aportada pela literatura acadêmica: a de que quanto mais transparente for o modelo de IA mais confiante estará seu usuário para lançar mão dela e aderir às suas recomendações. Esta crença é um dos

fatores que amparam e estimulam o estudo da *Artificial Explainable Intelligence* (XAI) e, portanto, ao ser contestada, ficam fragilizados alguns argumentos dela decorrentes.

Este resultado significa, em termos práticos, que a Aderência dos gestores às recomendações de um modelo de IA está diretamente associada ao grau de Confiança que eles têm na tecnologia em geral; e que esta Confiança – ao contrário do que a literatura acadêmica costuma sugerir – está inversamente atrelada ao nível Transparência da IA (mais transparência, portanto, resulta em menos confiança). Este mecanismo, ademais, é válido para decisões de natureza estratégica, mas não se verifica em decisões operacionais – o que indica que, diante de diferentes tipos de decisão, os gestores interpretam e avaliam os cenários de forma distinta.

5 DISCUSSÃO

A conexão da variável Transparência com a variável Aderência encontra sua principal justificativa em estudos de *Explainable Artificial Intelligence* (ADADI; BERRADA, 2018; MESKE *et al*., 2020; SCHEMMER *et al*., 2022; SHRESTHA; BEN-MENAHEM; VON KROGH, 2019; ZEDNIK, 2021). Porém, os dados obtidos no experimento ora aplicado atestam (com base nos testes aplicados à hipótese H1) que embora a Transparência tenha repercussão sobre a Aderência, o impacto da primeira sobre a última evidencia uma relação inversa, ou seja: mais transparência implica menos aderência.

Uma das possíveis explicações para esse impacto negativo da Transparência sobre a Aderência é exatamente a percepção de confiabilidade que perpassa a relação entre as duas variáveis. Isto porque **i)** a transparência em nível elevado nem sempre transmite ao usuário uma mensagem de confiabilidade (CHIEN *et al*. 2014; GLIKSON, 2020; HANCOCK *et al*., 2011; MILLER, 2015; SUTTON; HOLT; ARNOLD, 2016); e **ii)** a confiabilidade pode ser afetada e fragilizada por diversas situações, tais como: o desempenho da tecnologia, a cultura organizacional e experiência anteriores do usuário (CHIEN *et al.* 2014; HANCOCK *et al*., 2011) – o que, em última instância, diminui a disposição dos gestores em manifestar aderência às recomendações da IA.

A crença de que ao apresentar com clareza, detalhe, e profusão de explicações o funcionamento de um modelo de IA os seus usuários irão sentir-se mais confiantes para utilizá-lo é contestada por pesquisadores que estabelecem diversos outros requisitos, tão ou mais importantes que a Transparência, para que se chegue ao ponto de confiar em um algoritmo (GARNER *et al*., 2021; RIBEIRO; SINGH; GUESTRIN, 2016). Além disso, por meio da Transparência, o usuário compreende o funcionamento da IA e torna-se capaz de julgar suas recomendações, libertando-se da escravidão ao algoritmo – "o algoritmo me fez fazer isso" (NILSEN, 2020, p. 13) –, e rompendo com o consequencialismo "mais transparência » mais aderência". Sob este ponto de vista, portanto, seria possível e justificável imaginar um cenário de alta transparência em que houvesse, por parte dos gestores, menor Aderência.

Ao tratar sobre Transparência e IA alguns estudos destacam a presença de uma espécie de "vergonha" nas decisões de usuários que, sem compreender direito o funcionamento do modelo de IA, devido à carência de Transparência, não se sentem aptos a discordar deste modelo e acabam aderindo irrefletidamente ao que ele sugere (BONAVIA; BROX-PONCE, 2018; PANTANO; SCARPI, 2022). Este tipo de aderência manifesta um excesso de Confiança, também chamado "*overtrust*" ou "*benevolency*" (CHIEN *et al.,* 2014; GLIKSON, 2020;

HANCOCK *et al.*, 2011; MILLER, 2015; SUTTON; HOLT; ARNOLD, 2016) e pode, entre outras coisas, levar o usuário a supor e aceitar que o modelo "sabe o que faz" (ainda que o usuário não compreenda), cabendo-lhe apenas aderir à recomendação feita. Este comportamento é visto como preocupante e pode comprometer a qualidade da decisão e do processo decisório, levando a consequência danosas. Importante notar que a taxativa previsão de Nam e Lyons (2020, p. 5), segundo quem "na ausência de confiança em um sistema, usuários optarão por não utilizá-lo" encontra respaldo nos achados da presente tese, embora o caminho para a construção desta Confiança não seja tão intuitivo quanto os autores supuseram.

Não se pode desconsiderar também que, ao fornecer mais Transparência ao usuário do modelo de IA, abrem-se – para o usuário – mais possibilidades de interpretação acerca do modelo – algumas das quais podem não favorecer a Aderência (PREECE, 2018; SHRESTHA; BEN-MENAHEM; VON KROGH, 2019). Em cenário de alta Transparência, eventual desconfiança (*distrust*) em relação ao volume de informações que foram comunicados também não pode ser descartada: os detalhes, quando julgados excessivos, podem – de acordo com a leitura de alguns usuários – colocar o modelo de IA sob suspeição. Assim, por paradoxal que seja, mais Transparência pode, sim, significar menos Aderência.

Embora a *Explainable Artificial Intelligence* sustente que modelos de IA menos transparentes (*black box*) são mais precisos e eficazes que modelos mais transparentes (*white box*), ela não pontifica que isto é causa de maior ou menor aderência dos usuários às recomendações feitas por estes modelos. Assim, atenta à observação de Alufaisan *et al*. (2021), segundo quem não há um entendimento conclusivo sobre como a Transparência repercute nos modelos de IA, a presente tese amplia esta discussão ao sugerir (vide item 4.3) que a Confiança que o usuário tem na tecnologia é que permeia e explica a relação entre a Transparência da IA e a Aderência dos usuários.

Ao deparar-se com o resultado referente à hipótese H1 também convém admitir que o movimento de migração dos modelos de IA *black box* em direção a modelos *white box* (LOYOLA-GONZALEZ, 2019) perde parte de seu propósito, uma vez que não há evidências de que tal migração favoreceria o aprimoramento do processo decisório mediante a aceitação das recomendações fornecidas pela IA. Esta consideração vai ao encontro das conclusões de Lang (2018) que mostrou que, mesmo manipulando os *inputs* de um modelo de IA, os diferentes níveis de Transparência gerados não influenciam de maneira significativa a decisão dos gestores frente às sugestões propostas por tal modelo.

Outrossim, a Decisão parece não ser capaz, por si só, de alterar a percepção dos gestores com relação à IA. É preciso reconhecer que a Decisão, por mais que leve a ponderações e

comportamentos diversos, em vista do grau de risco e da amplitude do impacto a que está associada, não é *de per si* causa suficiente para impactar a adesão dos gestores ao que a IA sugere.

A explicação mais comum para a associação entre Decisão e Aderência proposta por *AI recommender systems* está ligada a questões subjetivas sobre como os decisores percebem e interpretam riscos e incertezas (HARL *et al*., 2020; KIM; SONG, 2022; MAIDA *et al.*, 2012; ZHOU, 2022): diante de decisões operacionais, por vezes classificadas como menos arriscadas, gestores parecem aderir mais facilmente às recomendações da IA; ao passo que em decisões estratégicas, normalmente associadas a riscos elevados, gestores sentem-se menos confortáveis para seguir uma recomendação de IA (de certo modo, aderir ao que a IA sugere, neste último caso, é visto como uma adição de risco a uma decisão que, por natureza, já é arriscada).

A insignificância (estatística) do impacto da Decisão sobre a Aderência pode ser teoricamente respaldada de várias maneiras: pela necessidade de compreender e avaliar com mais profundidade os sistemas frente às decisões específicas que eles supostamente auxiliam (LU *et al.*, 2022; HARL, 2020; ZHOU, 2022); por dificuldades de adequação entre o modelo decisório e a interface do *recommendation system* (PU; CHEN, 2006, 2007); por questões relacionadas à confiança dos usuários em situações de incerteza (KIM; SONG, 2022; MAIDA *et al.*, 2012) etc.

Assim, para muitos talvez seja menos importante entender e identificar a tipologia de uma decisão e mais importante entender a lógica da ferramenta que recomenda esta ou aquela atitude diante da referida decisão. A Aderência, neste contexto, parece se dar mais pela percepção e compreensão de quão adequado é o modelo para orientar a decisão, do que por considerações sobre a natureza da decisão. Assim, em uma tomada de decisão com suporte de IA, outras ponderações parecem sobressair à avaliação dos gestores, de modo que avaliações sobre o tipo de decisão em questão não são tão significativas quanto se supõe.

Ademais, como já comentado, a Aderência está permeada pela Confiança que o gestor deposita na IA. Nesta perspectiva, a (Des)Confiança costuma estar relacionada a um dos polos da *Human-Computer Interaction,* isto é: ou o usuário tem em si aspectos que favorecem/desfavorecem a Confiança na máquina, ou a máquina traz em si elementos que estimulam/inibem a Confiança do usuário. A Decisão, porém, é um fator completamente externo, que não se vincula diretamente nem ao usuário nem à máquina. Talvez por isso não desempenhe papel relevante na Aderência.

Destaque-se ainda que a Decisão não está relacionada aos pilares que costumam amparar a introdução de inteligência artificial nas organizações. Nos termos de Lee e See

(2004), a Decisão não está diretamente vinculada à IA em termos de propósito, não ajuda a entender o processo, e não agrega qualquer valor ou significado em matéria de *performance*. Daí que o seu efeito sobre a Aderência acaba sendo inconsistente/irrelevante.

Diante disto, embora a tendência (em termos de média) continue aparentando que em decisões operacionais a Aderência é maior, a insignificância estatística demonstrada no teste da hipótese H2 desconstrói a argumentação teórica que gira em torno do trinômio Decisão – Percepção de Risco – Aderência, e sugere que há outros aspectos e variáveis, mais importantes que a Decisão, que sensibilizam a Aderência dos gestores de forma mais significativa e impactante.

Diante destas conclusões, extraídas ao testar H1 e H2, a constatação de H3 era presumível e lógica. Se **i)** em cenários de alta transparência a Aderência mostra-se reduzida (em relação a situações de baixa transparência); e se **ii)** em decisões estratégicas a Aderência é menor que em decisões operacionais; então, é compreensível que a combinação de alta transparência com decisão estratégica implique no menor nível de Aderência entre as quatro condições experimentais propostas no presente experimento.

A análise dos efeitos simples da ANCOVA que relacionou a Aderência a cada uma das condições experimentais esclarece dois pontos importantes para o entendimento de H3. Em primeiro lugar, constatou-se que a Decisão impacta significativamente a alta transparência. Em outros termos: quando o modelo de IA apresenta um nível de Transparência elevado, a média de Aderência difere significativamente entre decisões operacionais e decisões estratégicas. Em segundo lugar, verificou-se que a Transparência exerce influência significativa quando aplicada a decisões estratégicas, ou seja: diante de decisões estratégicas, a média de Aderência difere significativamente a depender do nível de Transparência (baixo/alto).

Alguns *insights* podem ser extraídos a partir dessas conclusões. De início, nota-se que, mesmo apropriando-se do funcionamento do modelo de IA por meio de informações claras e suficientes (= alta transparência), os gestores ponderam sobre a Decisão envolvida antes de manifestar sua Aderência: em decisões operacionais, eles aderem mais; e em decisões estratégicas, aderem menos.

Este achado é consistente com a argumentação clássica relativa aos efeitos da *accountability* sobre o decisor (BROWN, 1999; LERNER; TETLOCK, 1999; SIEGEL-JACOBS; YATES, 1996; SIMONSON, 1992), qual seja: nas decisões em que a exigência por *accountability* é mais acentuada (estratégicas), gestores preferem ancorar-se a modelos tradicionais. Arrematar uma decisão de amplo alcance e vital importância para a continuidade da organização com base na recomendação de um modelo de IA (mesmo de alta transparência)

atemoriza os gestores e, por isso, em decisões estratégicas eles optam por acatar as fórmulas prontas e sugestões já conhecidas que os modelos tradicionais oferecem. Doutro lado, em decisões operacionais, sendo menor a tensão e a pressão por *accountability*, a alta transparência parece encorajar os gestores a manifestar maior Aderência. O risco pessoal da decisão operacional é menor de modo que os gestores estão dispostos, nesse caso, a se afastar do modelo tradicional para seguir o modelo de IA.

Além disso, ao analisar a Aderência dos gestores nas decisões estratégicas vê-se que o nível de transparência é fator preponderante para a tomada de decisão. Isto é: em condições de alta transparência, gestores submetidos a decisões estratégicas manifestam significativamente menos Aderência às recomendações da IA que em condições de baixa transparência.

Mais uma vez, uma possível explicação para este fenômeno está ligada às consequências da *accountability* sobre a percepção dos gestores. Se por um lado, diante de uma IA de alta transparência a subsidiar decisões estratégicas, os gestores optam por ancorar-se à segurança dos modelos tradicionais; por outro lado, se estas mesmas decisões estratégicas são amparadas por modelos de IA de baixa transparência, os gestores se veem privados de importantes parâmetros decisórios (como simplicidade e clareza) e acabam "comprando o risco" de confiar na IA para "eximir-se" de tomar uma decisão difícil. Assim, frente à reduzida transparência da inteligência artificial, gestores preferem seguir a recomendação do algoritmo para, de certa forma, delegar uma atribuição sua e, assim, minimizar sua responsabilidade pessoal pela decisão tomada.

Os *insights* que partem de H3, portanto, dão conta de que o papel da *accountability* na interação entre as variáveis Decisão e Transparência apresenta duas possibilidades paradoxais: ou leva ao temor de responsabilização, e acaba ancorando os gestores aos modelos tradicionais de suporte à decisão; ou conduz a uma ousadia consciente, que os faz acatar as recomendações da IA na expectativa de eximir-se de suas responsabilidades.

6 CONCLUSÃO

A presente tese visou examinar se a aderência dos gestores às recomendações de um modelo de IA é afetada pela transparência do referido modelo e pelo tipo de decisão envolvida. Neste sentido, foram propostas três hipóteses: uma para avaliar o impacto individual da Transparência sobre a Aderência (H1); outra para avaliar o impacto individual da Decisão sobre a Aderência (H2); e uma última para examinar uma eventual influência conjunta de Transparência e Decisão sobre a Aderência (H3).

Assim, ao testar a hipótese H1 verificou-se que a Transparência afeta de forma significativa a Aderência ($F = 3.96$; $p = 0.0496$). Porém, contrariando a predição de H1, a relação entre Transparência e Aderência é de natureza inversa – e não direta – isto é: maior Transparência implica menor Aderência. Assim, o efeito principal da Transparência sobre a Aderência, embora existente, aponta para uma direção oposta à da previsão teórica formulada. A hipótese H1, então, foi rejeitada.

Como possibilidade de explicação deste resultado, pode-se aventar a ocorrência de dois fenômenos já mapeados pela literatura acadêmica: *overtrust* (excesso de confiança), para justificar a alta aderência em cenários de baixa transparência; e, paralelamente, *distrust* (desconfiança) para justificar a baixa aderência em cenário de alta transparência.

Doutro lado, o teste de hipótese aplicado a H2, sugere que a Decisão não é capaz, por si só, de alterar de forma estatisticamente significativa ($F = 2.52$; $p = 0.1158$) a percepção dos gestores – e, consequentemente, não provoca uma modificação substantiva na sua Aderência.

Três possibilidades de explicação emergem para esclarecer os achados relativos a H2: **i)** à luz da lógica proposta por Lee e See (2004), a Decisão não está diretamente vinculada à IA em termos de propósito, não ajuda a entender o processo, e não agrega qualquer valor ou significado em matéria de *performance*. Daí que o seu efeito sobre a Aderência acaba sendo inconsistente / irrelevante; **ii)** de acordo com a *Human-Computer Interaction,* tanto o usuário de um modelo de IA quanto a própria tecnologia de IA apresentam fatores intrínsecos que podem afetar a Aderência. Já a Decisão, como fator extrínseco à relação homem-máquina, não tem o condão de produzir esta mesma afetação; **iii)** a acentuada disseminação da *Explainable Artificial Intelligence*, focada mais em entender a lógica da ferramenta de suporte à decisão do que em identificar a tipologia da referida decisão, pode levar os gestores a minimizar a importância de considerar a Decisão nas suas escolhas e avaliações – e por isso a Aderência deles não tem vinculação com esta variável.

Por fim, o teste da hipótese H3, cujo objetivo era demonstrar que a interação entre as variáveis Transparência e Decisão tinha uma repercussão na variável Aderência, evidenciou que esta interação, de fato, afeta de forma significativa ($F = 3.24$; $p = 0.0752$) a Aderência dos gestores às recomendações do sistema de inteligência artificial. Entretanto, assim como no caso de H1, o sentido do resultado obtido foi contrário ao da predição teórica. Assim, não é possível afirmar que o efeito positivo da alta transparência sobre a Aderência é maior em decisões operacionais que em decisões estratégicas.

Na realidade, o que se apurou com relação aos achados de H3 foi o exato oposto: o efeito da alta transparência é, na verdade, negativo; e é significativamente mais intenso em decisões estratégicas que em decisões operacionais. Possíveis explicações para este fenômeno, podem seguir duas linhas de raciocínio (opostas, mas complementares): **i)** uma espécie de viés de ancoragem, faz com que os gestores, quando diante de decisões estratégicas, mesmo contando com uma IA de alta transparência, prefiram lastrear suas decisões nas recomendações dos modelos tradicionais de suporte à decisão (daí a pouca Aderência neste caso); **ii)** a pressão por *accountability* faz com que gestores, temendo os riscos de uma eventual responsabilização pessoal por decisões estratégicas tomadas, acatem as recomendações da IA – mesmo que esta apresente baixos níveis de Transparência – na expectativa de se eximir ou minimizar suas responsabilidades (daí a grande Aderência neste caso).

Esta tese contribui com algumas áreas da pesquisa acadêmica: i) a literatura relativa à *Explainable Artificial Intelligence* recebe um aporte – um tanto quanto disruptivo – com as conclusões ora apresentadas sobre a Transparência e sua interação com diferentes tipos de Decisão; ii) os estudos sobre *Decision-Making*, sobretudo os relativos a *decision-aids*, também foram incrementados com considerações sobre como características presentes nos modelos de IA e nos contextos decisórios interferem na formação da percepção dos gestores decisores; iii) e a pesquisa concernente a *Human-Computer Interaction*, notadamente aquela que versa sobre *recommender systems*, foi enriquecida com a explanação sobre a mediação exercida pelo elemento Confiança e sobre como este contribui para uma interação efetiva entre homem e máquina.

A análise de mediação, aliás, demonstra que a construção de uma relação de confiança na IA (e nas recomendações feitas por ela) está, pelo que aqui se expôs, mais relacionada a uma característica intrínseca do dispositivo de IA (transparência), do que à circunstância que conecta a máquina ao usuário (a tomada de uma decisão). As disposições interiores do próprio usuário com relação à tecnologia em geral também são relevantes e reverberam na sua relação com a

IA. Tudo isto representa mais uma importante contribuição oferecida pela presente tese àqueles que estudam a fenomenologia da Confiança na IA.

Na esteira da transformação digital em curso, tanto o planejamento estratégico quanto o operacional das organizações podem ser feitos com suporte de inteligência artificial. Ao abordar de forma específica cada tipo de decisão, a presente tese oferece *insights* – para gestores e para desenvolvedores – sobre o que considerar na elaboração de ferramentas de suporte ao planejamento para que contem com maior Aderência por parte de seus usuários.

Além de poder subsidiar o *design* de modelos de IA a partir dos novos conhecimentos ora agregados, as contribuições aqui oferecidas também trazem à tona pontos importantes a serem considerados em processos de implantação de modelos de IA nas organizações. Isto porque a aceitação, adaptação e integração das ferramentas de IA no dia a dia de uma organização sofre, por vezes, certa resistência por parte do corpo funcional. A tese ora proposta apresenta e discute possibilidades de explicação para este fenômeno (e outros a ele relacionados) viabilizando, assim, que se possa melhor entender e (re)definir a relação entre os humanos e a IA, no intuito de aumentar a aderência deles às propostas dos *recommender systems*.

A discussão sobre o papel dos gestores nesta era de tomada de decisões com suporte de IA também aufere um contributo significativo a partir dos achados da presente pesquisa. Entender as percepções, as preferências e a dinâmica que rege o raciocínio decisório quando diante de sistemas de inteligência distintos, e frente a diferentes tipos de decisão, enriquece esta discussão e lança luzes sobre pontos nebulosos que ainda a permeiam.

Como limitação ao presente estudo, destaque-se que, embora se requeira dos participantes que respondam à pesquisa de forma individual e isolada, é possível que os temas sob investigação já tenham sido comentados entre colegas e, portanto, tenha-se organicamente formado uma compreensão comum a alguns. Isto representa uma limitação do estudo, já que tal situação, se real, não pode ser claramente identificada, manipulada, e excluída da amostra. Avalia-se, porém, que o risco potencial de um eventual "entendimento compartilhado" é mínimo já que o setor onde atuam os participantes é segregado em várias divisões (pequenos núcleos de tarefas), cujas estações de trabalho situam-se em planos distintos. Assim, ainda que a divisão *x* tivesse a mesma opinião sobre o assunto α, sendo a sua composição diminuta em relação ao total do *staff*, não haveria enviesamento significativo do estudo. Embora seja um ponto que inspira cuidado e, de certa maneira, limita a generalização de inferências, entende-se que um tal risco não representa óbice à realização da pesquisa nem invalida suas conclusões.

Além disso, os cenários de cada caso, como já dito, foram hipotéticos. Embora isto consista a princípio em uma limitação – posto que circunscreve a investigação ao nível do imaginário e não do plano concreto –, o caráter preditivo do estudo torna justificável, e até necessário, que se trabalhe primeiro com esta perspectiva de teorização e exemplificação, para que – só depois – sejam realizadas pesquisas com o fito de testar modelos e ferramentas de inteligência artificial reais.

Como primeira sugestão para futuras pesquisas correlacionadas ao tema aqui desenvolvido, seria oportuno avaliar os impactos da transparência da IA sobre as decisões gerenciais a partir da implementação de um sistema real de inteligência artificial (quiçá em um experimento de campo). Além disso, o desenvolvimento de um modelo experimental capaz de capturar as nuances da influência da transparência sobre decisões de natureza tática (que não foi abordada no presente estudo) também complementaria o arcabouço de conhecimento sobre as interações entre os níveis de transparência e os tipos de decisão. Por fim, convém também realizar estudos para aprofundar o papel do gênero (aqui tratado apenas marginalmente, como covariável) na tomada de decisões.

Enfim, dinamizar o processo de transformação digital, dialogar com novos sistemas de inteligência, inovar na tomada de decisões, refinar os algoritmos de IA, repensar o papel dos gestores perante os avanços tecnológicos, estreitar e aprofundar a relação entre usuários e máquinas: todos estes componentes fazem parte da vida cotidiana das organizações (e da sociedade em geral) e requerem um esteio teórico sólido sobre o qual possam se firmar, desenvolver e ampliar. A presente tese, com as conclusões que derivam do experimento ora aplicado, contribui – na medida do recorte que se propôs a examinar – com a formação deste lastro e com o incremento do conhecimento científico nesta seara.

REFERÊNCIAS

ADADI, A.; BERRADA, M. Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). **IEEE Access**, v. 6, p. 52138-52160, 2018.

ALBUQUERQUE, P. H. M.; SAAVEDRA, C. A. P. B.; MORAIS, R. L.; ALVES, P. F.; YAOHAO, P. **Na era das máquinas, o emprego é de quem?** Estimação da probabilidade de automação de ocupações no Brasil. Rio de Janeiro: IPEA, 2019. (Texto para Discussão, 2457).

ALLES, M. G. Drivers of the use and facilitators and obstacles of the evolution of big data by the audit profession. **Accounting Horizons,** v. 29, n. 2, p. 439-449, 2015.

ALUFAISAN, Y.; MARUSICH, L. R.; BAKDASH, J. Z.; ZHOU, Y.; KANTARCIOGLU, M. Does explainable artificial intelligence improve human decision-making? **The Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21)**, v. 35, n. 8, p. 6618-6626, 2021.

AMASON, A. C.; MOONEY, A. C. The effects of past performance on top management team conflict in strategic decision making. **International Journal of Conflict Management**, v. 10, n. 4, p. 340-359, 1999.

AMERSHI, S.; WELD, D.; VORVOREANU, M.; FOURNEY, A.; ... HORVITZ, E. Guidelines for human-AI interaction. *In*: THE CHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, 2019, Glasgow. **Proceedings [...]**. New York: Association for Computing Machinery, 2019.

ANGELOV, P. P.; SOARES, E. A.; JIANG, R.; ARNOLD, N. I.; ATKINSON, P. M. Explainable artificial intelligence: an analytical review. **Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery,** v. 11, n. 5, e1424, 2021.

ANTHONY, R. **Planning and control systems:** a framework for analysis. Boston: Harvard University, 1965.

APPELBAUM, D.; KOGAN, A.; VASARHELYI, M.; YAN, Z. Impact of business analytics and enterprise systems on managerial accounting. **International Journal of Accounting Information Systems**, v. 25, p. 29-44, 2017.

ARAUJO, T.; HELBERGER, N.; KRUIKEMEIER, S.; DE VREESE, C. H. In AI we trust? Perceptions about automated decision-making by artificial intelligence. **AI and Society,** v. 35, n. 3, p. 611-623, 2020.

ARRIETA, A. B.; DÍAZ-RODRÍGUEZ, N.; DEL SER, J.; BENNETOT, A.; ... HERRERA, F. Explainable Artificial Intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. **Information Fusion,** v. 58, p. 82-115, 2020.

ASCARZA, E.; ISRAELI, A. Eliminating unintended bias in personalized policies using bias-eliminating adapted trees (BEAT). **Proceedings of the National Academy of Sciences of the United States of America,** v. 119, n. 11, p. 1-9, 2022.

ASHMOS, D. P.; DUCHON, D.; MCDANIEL JR, R. R. Participation in strategic decision making: The role of organizational predisposition and issue interpretation. **Decision Sciences,** v. 29, n. 1, p. 25-51, 1998.

ATHEY, B. S. C.; BRYAN, K. A.; GANS, J. S. The allocation of decision authority to human. **AEA Papers and Proceedings**, v. 110, p. 80-84, 2020.

BALDVINSDOTTIR, G.; BURNS, J.; NØRREKLIT, H.; SCAPENS, R. W. The image of accountants: from bean counters to extreme accountants. **Accounting, Auditing & Accountability Journal**, v. 22, n. 6, p. 858-8820 2009.

BANERJEE, S. Improving online rent-or-buy algorithms with sequential decision making and ML predictions. **Advances in Neural Information Processing Systems,** v. 33, p. 21072-21080, 2020.

BANSAL, G.; NUSHI, B.; KAMAR, E.; LASECKI, W. S.; … HORVITZ, E. Beyond accuracy: the role of mental models in human-AI team performance. **Proceedings of the AAAI Conference on Human Computation and Crowdsourcing,** v. 7, n. 1, p. 2-11, 2019.

BANSAL, H.; KHAN, R. A review paper on human computer interaction. **International Journals of Advanced Research in Computer Science and Software Engineering,** v. 8, p. 53-56, 2018.

BĂRBUȚĂ-MIȘU, N.; MADALENO, M.; ILIE, V. Analysis of risk factors affecting firms' financial performance – support for managerial decision-making. **Sustainability,** v. 11, n. 18, p. 4838, 2019.

BARNARD, C. I. **The functions of the executive**. Cambridge: Mass, 1938.

BARRAT, J. **Our final invention: artificial intelligence and the end of the human era**. Macmillan: St. Martin's Press, 2013.

BASI, R. S. Administrative decision making: a contextual analysis. **Management Decision**, v. 36, n. 4, p. 232-240, 1998.

BAUGUESS, S. W. **The role of big data, machine learning, and AI in assessing risks:** a regulatory perspective. New York: OpRisk North America, 2017.

BIGMAN, Y. E.; GRAY, K. People are averse to machines making moral decisions. **Cognition**, v. 181, p. 21-34, 2018.

BLUM, A. Machine learning theory. **Carnegie Melon University, School of Computer Science,** v. 26, 2007.

BOHLIN, T. P. **Practical grey-box process identification:** theory and applications. Berlin: Springer Science & Business Media, 2006.

BONAVIA, T; BROX-PONCE, J. Shame in decision making under risk conditions: Understanding the effect of transparency. **Plos One**, v. 13, n. 2, e0191990, 2018.

BONNER, S. E. Judgment and decision-making research in accounting. **Accounting Horizons,** v. 13, n. 4, p. 385, 1999.

BONOLLO, B.; ZHANG, Z. An experiment on group effects and Bayesian rationality. **UNSW Business School Research Paper**, 2018. Disponível em: http://dx.doi.org/10.2139/ssrn.3178550. Acesso em: 11 fev. 2023.

BRADY, M. Artificial intelligence and robotics. *In*: BRADY, M.; GERHARDT, L. A.; DAVIDSON, H. F. **Robotics and artificial intelligence**. Berlin: Springer, 1984. p. 47-63. (NATO ASI F, 21)

BRANDS, K.; HOLTZBLATT, M. Business analytics: transforming the role of management accountants. **Management Accounting Quarterly,** v. 16, n. 3, 2015.

BROUSSARD, M. **Artificial unintelligence:** How computers misunderstand the world. Cambridge: The MIT Press, 2018.

BROWN, C. L. "Do the right thing:" diverging effects of accountability in a managerial context. **Marketing Science**, v. 18, n. 3, p. 230-246, 1999.

BUSENITZ, L. W.; BARNEY, J. B. Differences between entrepreneurs and managers in large organizations: Biases and heuristics in strategic decision-making. **Journal of Business Venturing,** v. 12, n. 1, 9-30, 1997.

CARD, S. K.; MORAN, T. P.; NEWELL, A. **The psychology of human-computer interaction.** Hillside: L. Erlbaum Associates Inc., 1983.

CARROLL, J. M. Human-computer interaction: psychology as a science of design. **International Journal of Human-Computer Studies,** v. 46, n. 4, p. 501-522, 1997.

CASTELVECCHI, D. Can we open the black box of AI? **Nature News,** v. 538, n. 7623, p. 20, 2016.

CAVANAGH, K. Turn on, tune in and [don't] drop out: engagement, adherence, attrition, and alliance with internet-based interventions. *In*: BENNETT-LEVY, J. *et al*. (Eds.). **Oxford guide to low intensity CBT interventions. Oxford guides in cognitive behavioural therapy.** New York: Oxford University Press, 2010. p. 227-233.

CEPNI, E. The science and art of decision making. *In*: IEEE INTERNATIONAL CONFERENCE ON COGNITIVE INFORMATICS & COGNITIVE COMPUTING, 18.; 2019, Milan. **Proceedings [...]** Milan: IEEE, 2019. p. 272-277.

CGMA. **Re-inventing finance for a digital world:** the future of finance. London: CGMA, 2019 Disponível em: https://www.cimaglobal.com/Documents/Future%20of%20Finance/future-re-inventing-finance-for-a-digital-world.pdf. Acesso em: 11 fev. 2023.

CHEN, C. L. P.; ZHANG, C.-Y. Data-intensive applications, challenges, techniques and technologies: a survey on Big Data. **Information sciences,** v. 275, p. 314-347, 2014.

CHEN, J. Y.; PROCCI, K.; BOYCE, M.; WRIGHT, J.; … BARNES, M. **Situation awareness-based agent transparency.** Aberdeen: Army Research Laboratory, 2014. Disponível em: https://apps.dtic.mil/sti/pdfs/ADA600351.pdf. Acesso em: 11 fev. 2023.

CHENG, M. M.; COYTE, R. The effects of incentive subjectivity and strategy communication on knowledge-sharing and extra-role behaviours. **Management Accounting Research,** v. 25, n. 2, p. 119-130, 2014.

CHENGALUR-SMITH, I. N.; BALLOU, D. P.; PAZER, H. L. The impact of data quality information on decision making: an exploratory analysis. **IEEE Transactions on Knowledge and Data Engineering,** v. 11, n. 6, p. 853-864, 1999.

CHIEN, S. Y.; LEWIS, M.; SEMNANI-AZAD, Z.; SYCARA, K. An empirical model of cultural factors on trust in automation. **Proceedings of the Human Factors and Ergonomics Society**, v. 58, n. 1, p. 859-863, 2014.

CHIEN, S.-Y.; LEWIS, M.; SYCARA, K.; LIU, J.-S.; KUMRU, A. Influence of cultural factors in dynamic trust in automation. *In*: IEEE INTERNATIONAL CONFERENCE ON SYSTEMS, MAN, AND CYBERNETICS (SMC), 2016, Budapest. **Proceedings […]** Budapest: IEEE, 2016. p. 2884-2889.

CHONG, V. K.; EGGLETON, I. R. C. The decision-facilitating role of management accounting systems on managerial performance: the influence of locus of control and task uncertainty. **Advances in Accounting,** v. 20, p. 165-197, 2003.

CHOWDHURY, M.; SADEK, A. W. Advantages and limitations of artificial intelligence. **Artificial Intelligence Applications to Critical Transportation Issues,** v. 6, n. 3, p. 360-375, 2012.

COECKELBERGH, M. Artificial intelligence, responsibility attribution, and a relational justification of explainability. **Science and Engineering Ethics**, v. 26, n. 4, p. 2051-2068, 2020.

COKINS, G. Top 7 trends in management accounting. **Strategic Finance,** v. 95, n. 6, p. 21-30, 2013.

CONTI, M.; DEHGHANTANHA, A.; FRANKE, K.; WATSON, S. Internet of Things security and forensics: Challenges and opportunities. **Future Generation Computer Systems,** v. 78, p. 544-546, 2018.

COOPER, S.; CROWTHER, D.; CARTER, C. Challenging the predictive ability of accounting techniques in modelling organizational futures. **Management Decision**, v. 39, n. 2, p. 137-146, 2001.

CORAM, P. J.; WANG, L. The effect of disclosing key audit matters and accounting standard precision on the audit expectation gap. **International Journal of Auditing,** v. 25, n. 2, p. 270-282, 2021.

CORDESCHI, R. Automatic decision-making and reliability in robotic systems: some implications in the case of robot weapons. **AI & Society,** v. 28, n. 4, p. 431-441, 2013.

CRAMER, H.; EVERS, V.; RAMLAL, S.; VAN SOMEREN, M.; … WIELINGA, B. The effects of transparency on trust in and acceptance of a content-based art recommender. **User Modeling and User-Adapted Interaction,** v. 18, n. 5, p. 455-496, 2008.

CUI, L.; YANG, S.; CHEN, F.; MING, Z.; … QIN, J. A survey on application of machine learning for Internet of Things. **International Journal of Machine Learning and Cybernetics,** v. 9, n. 8, p. 1399-1417, 2018.

DARGNIES, M.-P.; HAKIMOV, R.; KÜBLER, D. F. Aversion to hiring algorithms: transparency, gender profiling, and self-confidence. **CESifo Working Paper**, n. 9968, 1011. Disponível em: https://doi.org/10.2139/ssrn.4238275. Acesso em: 11 fev. 2023.

DATHEIN, R. Inovação e revoluções industriais: uma apresentação das mudanças tecnológicas determinantes nos séculos XVIII e XIX. **Publicações DECON Textos Didáticos,** v. 2, 2003.

DAVENPORT, T. H. From analytics to artificial intelligence. **Journal of Business Analytics,** v. 1, n. 2, p. 73-80.

DELEN, D.; RAM, S. Research challenges and opportunities in business analytics. **Journal of Business Analytics,** v. 1, n. 1, p. 2-12, 2018.

DENWATTANA, N.; GETTA, J. R. A parameterised algorithm for mining association rules. *In*: AUSTRALASIAN DATABASE CONFERENCE, 12.; 2001. **Proceedings […]** Australia: IEEE, 2001. p. 45-51.

DHAMIJA, P.; BAG, S. Role of artificial intelligence in operations environment: a review and bibliometric analysis. **The TQM Journal,** v. 32, n. 4, p. 869-896, 2020.

DIEDERICH, J. Neo-luddism. *In*: COGNITIVE SYSTEMS MONOGRAPHS. **The psychology of artificial superintelligence.** Berlin: Springer, Cham, 2021. v. 42, p. 73-93.

DIEDERICH, S.; BRENDEL, A. B.; MORANA, S.; KOLBE, L. On the design of and interaction with conversational agents: an organizing and assessing review of human-computer interaction research. **Journal of the Association for Information Systems,** v. 23, n. 1, p. 96-138, 2022.

DING, K.; LEV, B.; PENG, X.; SUN, T.; VASARHELYI, M. A. Machine learning improves accounting estimates: evidence from insurance payments. **Review of Accounting Studies,** v. 25, n. 3, p. 1098-1134, 2020.

DIRICAN, C. The impacts of robotics, artificial intelligence on business and economics. **Procedia-Social and Behavioral Sciences**, v. 195, p. 564-573, 2015.

DORAN, D.; SCHULZ, S.; BESOLD, T. R. What does explainable AI really mean? A new conceptualization of perspectives. **arXiv preprint,** n. 1710.00794, 2017. Disponível em: https://doi.org/10.48550/arXiv.1710.00794. Acesso em: 11 fev. 2023.

DOŠILOVIĆ, F. K.; BRČIĆ, M.; HLUPIĆ, N. Explainable artificial intelligence: a survey. *In*: INTERNATIONAL CONVENTION ON INFORMATION AND COMMUNICATION TECHNOLOGY, ELECTRONICS AND MICROELECTRONICS, 41.; 2018. **Proceedings […]** Opatija, Croatia: IEEE, 2018. p. 210-215.

DREYFUS, H. L. **What computers still can't do:** a critique of artificial reason. Cambridge: The MIT Press, 1992.

DUNEGAN, K. J.; DUCHON, D.; BARTON, S. L. Affect, risk, and decision criticality: replication and extension in a business setting. **Organizational Behavior and Human Decision Processes,** v. 53, n. 3, p. 335-351, 1992.

DWIPUTRA, F.; MUSTIKASARI, E. Literary review on the antecedent of ethical dilemma in management accounting profession. **Jurnal Riset Akuntansi Dan Bisnis Airlangga,** v. 6, n. 1, 2021.

ERNEST, N.; CARROLL, D.; SCHUMACHER, C.; CLARK, M.; … LEE, G. Genetic fuzzy based artificial intelligence for unmanned combat aerial vehicle control in simulated air combat missions. **Journal of Defense Management,** v. 6, n. 1, p. 2167-0374, 2016.

FEDERAÇÃO BRASILEIRA DE BANCOS. **Pesquisa Febraban de Tecnologia Bancária 2019**. São Paulo, 2019. Disponível em: https://cmsarquivos.febraban.org.br/Arquivos/documentos/PDF/Pesquisa-FEBRABAN-Tecnologia-Bancaria-2019.pdf. Acesso em: 11 fev. 2023.

FEDERAÇÃO BRASILEIRA DE BANCOS. **Pesquisa Febraban de Tecnologia Bancária 2022:** investimentos em tecnologia. São Paulo, 2022. Disponível em: https://cmsarquivos.febraban.org.br/Arquivos/documentos/PDF/pesquisa-febraban-2022-vol-2.pdf. Acesso em: 11 fev. 2023.

FELZMANN, H.; FOSCH-VILLARONGA, E.; LUTZ, C.; TAMÒ-LARRIEUX, A. Towards transparency by design for artificial intelligence. **Science and Engineering Ethics**, v. 26, n. 6, p. 3333-3361, 2020.

FERNANDES, A. M.; SCHNORRENBERGER, D.; RENGEL, R. Influência das características do decisor sobre os vieses da heurística da representatividade. **Revista Ambiente Contábil**, v. 12, n. 2, p. 298-317, 2020.

FERREIRA, J. J.; MONTEIRO, M. The human-AI relationship in decision-making: AI explanation to support people on justifying their decisions. **arXiv preprint,** n. 2102.05460, 2021. Disponível em: https://arxiv.org/abs/2102.05460. Acesso em: 11 fev. 2023.

FETZER, J. H. What is Artificial Intelligence? *In:* STUDIES IN COGNITIVE SYSTEMS. **Artificial intelligence:** its scope and limits. Springer, Dordrecht, 1990. v. 4, p. 3-27.

FISHBEIN, M.; AJZEN, I. Belief, attitude, intention, and behavior: An introduction to theory and research. **Philosophy and Rhetoric**, v. 10, n. 2, 1977.

FORBES INSIGHTS. **The Internet of Things:** from theory to reality-how companies are leveraging the IoT to move their businesses forward. New Jersey, 2017. Disponível em: https://events.pentaho.com/rs/680-ONC-130/images/internet-of-things-forbes-insights.pdf. Acesso em: 11 fev. 2023.

FREY, C. B.; OSBORNE, M. A. The future of employment: how susceptible are jobs to computerisation? **Technological Forecasting and Social Change,** v. 114, p. 254-280, 2017.

FRIEDMAN, A. L.; LYNE, S. R. Activity-based techniques and the death of the beancounter. **European Accounting Review,** v. 6, n. 1, p. 19-44, 1997.

GARCIA, R.; OLAK, P. A.; CLEMENTE, A.; FADEL, B. Teoria dos prospectos: vieses de percepção do usuário da informação no processo decisório. **Revista de Ensino, Educação e Ciências Humanas,** v. 13, n. 1, 2012.

GARDNER, H. **Frames of mind theory of multiple intelligences**: developing talent in young people. New York: Basic Books, 1983.

GARDNER, A.; SMITH, A. L.; STEVENTON, A.; COUGHLAN, E.; OLDFIELD, M. Ethical funding for trustworthy AI: proposals to address the responsibilities of funders to ensure that projects adhere to trustworthy AI practice. **AI and Ethics**, v. 2, p. 1-15, 2022.

GARRISON, R. H.; NOREEN, E. W.; BREWER, P. C. **Contabilidade gerencial**. Porto Alegre: AMGH, 2013.

GAVIRIA, C. I. G. **The unforeseen consequences of artificial intelligence (ai) on society:** a systematic review of regulatory gaps generated by AI in the U.S. Santa Monica, CA: RAND Corporation, 2020. Disponível em: https://www.rand.org/pubs/rgs_dissertations/RGSDA319-1.html. Acesso em: 11 fev. 2023.

GEETHA, R.; BHANU, S. R. D. Recruitment through artificial intelligence: a conceptual study. **International Journal of Mechanical Engineering and Technology,** v. 9, n. 7, p. 63-70, 2018.

GHANBARI, M.; VASELI, S. The role of management accounting in the organization. **International Research Journal of Applied and Basic Sciences,** v. 9, n. 11, p. 1913-1915, 2015.

GLIKSON, E.; WOOLLEY, A. W. Human trust in artificial intelligence: review of empirical research. **Academy of Management Annals,** v. 14, n. 2, p. 627-660, 2020.

GONZÁLEZ-MENDOZA, J. A.; CAÑIZARES-ARÉVALO, J. D. J.; CARDENAS-, M. **Decision-Making, Rationality, and Human Action,** v. 6, n. 1, p. 3977-3991, 2022.

GOTTHARDT, M.; KOIVULAAKSO, D.; PAKSOY, O.; SARAMO, C. … LEHNER, O. Current state and challenges in the implementation of smart robotic process automation in accounting and auditing. **ACRN Journal of Finance and Risk Perspectives**, v. 9, n. 1, p. 90-102, 2020.

GOW, I. D.; LARCKER, D. F.; REISS, P. C. Causal inference in accounting research. **Journal of Accounting Research,** v. 54, n. 2, p. 477-523, 2016.

GRACE, K.; SALVATIER, J.; DAFOE, A.; ZHANG, B.; EVANS, O. When will AI exceed human performance? Evidence from AI experts. **Journal of Artificial Intelligence Research,** v. 62, p. 729-754, 2018.

GRANLUND, M.; LUKKA, K. From bean-counters to change agents: the Finnish management accounting culture in transition. **LTA,** v. 3, n. 97, p. 213-255, 1997.

GRAY, G. L.; CHIU, V.; LIU, Q.; LI, P. The expert systems life cycle in AIS research: what does it mean for future AIS research? **International Journal of Accounting Information Systems,** v. 15, n. 4, p. 423-451, 2014.

GUJARATI, D. N.; PORTER, D. C. **Econometria básica 5.** Porto Alegre: AMGH, 2011

GÜNGÖR, H. Creating value with artificial intelligence: a multi-stakeholder perspective. **Journal of Creating Value,** v. 6, n. 1, p. 72-85, 2020.

GUNKEL, D. J. Mind the gap: responsible robotics and the problem of responsibility. **Ethics and Information Technology,** v. 22, p. 307-320, 2020.

GUNNING, D.; STEFIK, M.; CHOI, J.; MILLER, T.; … YANG, G. Z. XAI –explainable artificial intelligence. **Science Robotics,** v. 4, n. 37, eaay7120, 2019.

GUPTA, V. Why explainability should be the core of your ai application. **Forbes**, 23 jan. 2023. Disponível em: https://www.forbes.com/sites/forbestechcouncil/2023/01/23/why-explainability-should-be-the-core-of-your-ai-application/. Acesso em: 5 maio 2023.

HALEEM, A.; RAISAL, I. The study of the influence of information technology sophistication on the quality of accounting information system in bank branches at Amapara District, Sri Lanka. In: ANNUAL RESEARCH CONFERENCE, 5.; 2016, Oluvil. **Proceedings [...]** Oluvil, Sri Lanka, FMC, SEUSL, 2016. p. 114-124.

HALL, M. Accounting information and managerial work. **Accounting, Organizations and Society**, v. 35, n. 3, p. 301-315, 2010.

HANCOCK, P. A.; BILLINGS, D. R.; SCHAEFER, K. E.; CHEN, J. Y.; … PARASURAMAN, R. A meta-analysis of factors affecting trust in human-robot interaction. **Human Factors**, v. 53, n. 5, p. 517-527, 2011.

HANNA, R. C.; LEMON, K. N.; SMITH, G. E. Is transparency a good thing? How online price transparency and variability can benefit firms and influence consumer decision making. **Business Horizons,** v. 62, n. 2, p. 227-236, 2019.

HARL, M.; WEINZIERL, S.; STIERLE, M.; MATZNER, M. Explainable predictive business process monitoring using gated graph neural networks. **Journal of Decision Systems**, v. 29, Supplement 1, p. 312-327, 2020.

HARRISON, E. F.; PELLETIER, M. A. A paradigm for strategic decision success. **Management Decision**, v. 33, n. 7, 1995

HARRISON, F. E.; PELLETIER, M. A. Levels of strategic decision success. **Management Decision,** v. 38, n. 2, p. 107-118, 2000.

HARTMANN, F. G.; MAAS, V. S. Why business unit controllers create budget slack: involvement in management, social pressure, and machiavellianism. **Behavioral Research in Accounting**, v. 22, n. 2, p. 27-49, 2010.

HERM, L.-V.; WANNER, J.; SEUBERT, F.; JANIESCH, C. I don't get it, but it seems valid! The connection between explainability and comprehensibility in (X) AI research. **ECIS 2021 Research Papers**, v. 82, n. 1413, 2021.

HILL, R. K. What an algorithm is. **Philosophy & Technology,** v. 29, p. 35-59, 2016.

HIRST, D. E.; KOONCE, L.; VENKATARAMAN, S. How disaggregation enhances the credibility of management earnings forecasts. **Journal of Accounting Research,** v. 45, n. 4, p. 811-837, 2007.

HÖDDINGHAUS, M.; SONDERN, D.; HERTEL, G. The automation of leadership functions: would people trust decision algorithms? **Computers in Human Behavior,** v. 116, 106635, 2021.

HOFFMAN, C. **Accounting and auditing in the digital age.** Harvard: Digital Financial Reporting, 2017. Disponível em: http://xbrlsite.azurewebsites.net/2017/Library/AccountingAndAuditingInTheDigitalAge.pdf. Acesso em: 11 fev. 2023.

HOLSAPPLE, C. W.; JOSHI, K. D. A knowledge management ontology. *In*: HOLSAPPLE, C. W.; **Handbook on Knowledge Management 1**. Berlin: Springer, 2004. p. 89-124.

HORVITZ, E. J.; BREESE, J. S.; HENRION, M. Decision theory in expert systems and artificial intelligence. **International Journal of Approximate Reasoning,** v. 2, n. 3, p. 247-302, 1988.

IKEDA, E. K. **Knowledge management in IOT ecosystems:** how to generate intelligence and connectivity. 2020. Dissertação (Mestrado em Gestão de Projetos) – Universidade Nove de Julho, São Paulo, 2020.

IMANE, A.; DRISS, H. Strategic management for organizational performance: from which come the mistakes of strategic decision-making. **European Journal of Economics and Business Studies,** v. 3, n. 3, p. 291-300, 2017.

ISHIZAKA, A.; SIRAJ, S. Are multi-criteria decision-making tools useful? An experimental comparative study of three methods. **European Journal of Operational Research,** v. 264, n. 2, p. 462-471, 2018.

JARRAHI, M. H. Artificial intelligence and the future of work: human-AI symbiosis in organizational decision making. **Business Horizons,** v. 61, n. 4, p. 577-586, 2018.

JUSSUPOW, E.; SPOHRER, K.; HEINZL, A. Identity threats as a reason for resistance to artificial intelligence: survey study with medical students and professionals. **JMIR Formative Research,** v. 6, n. 3, e28750, 2022.

KACHELMEIER, S. J.; THORNOCK, T. A.; WILLIAMSON, M. G. Communicated values as informal controls: promoting quality while undermining productivity? **Contemporary Accounting Research,** v. 33, n. 4, p. 1411-1434, 2016.

KANE, G. C.; PALMER, D.; PHILLIPS, A. N.; KIRON, D.; BUCKLEY, N. **Achieving digital maturity:** adapting your company to a changing world. Cambridge: The MIT Press, 2017. Disponível em: https://sloanreview.mit.edu/projects/achieving-digital-maturity/. Acesso em: 11 fev. 2023.

KATZ, B. **The intelligence edge:** opportunities and challenges from emerging technologies for US intelligence. Washington, DC: Center for Strategic and International Studies, 2020.

KELLY, J. L.; SELFRIDGE, O. G. Sophistication in computers: a disagreement. **IRE Transactions on Information Theory,** v. 8, n. 2, p. 78-80, 1962.

KESUMA, S. A.; SAIDIN, S. Z.; AHMI, A. IT sophistication: implementation on state owned banks in Indonesia. **International Review of Management and Marketing,** v. 6, n. 8, p. 234-239, 2016.

KHANZODE, K. C. A.; SARODE, R. D. Advantages and disadvantages of artificial intelligence and machine learning: a literature review. **International Journal of Library & Information Science,** v. 9, n. 1, p. 3, 2020.

KHAVAS, Z. R. A review on trust in human-robot interaction. **arXiv preprint,** n. 2105.10045, 2021. Disponível em: http://arxiv.org/abs/2105.10045. Acesso em: 11 fev. 2023.

KIM, T.; SONG, H. Communicating the limitations of AI: the effect of message framing and ownership on trust in artificial intelligence. **International Journal of Human-Computer Interaction**, v. 39, n. 4, p. 790-800, 2022.

KOCHER, M. G.; POGREBNA, G.; SUTTER, M. The determinants of managerial decisions under risk. **Working Papers in Economics and Statistics,** Universität Innsbruck, n. 2008,4, 2008.

KOLASINSKA, A.; LAURIOLA, I.; QUADRIO, G. Do people believe in artificial intelligence? A cross-topic multicultural study. *In*: EAI INTERNATIONAL CONFERENCE ON SMART OBJECTS AND TECHNOLOGIES FOR SOCIAL GOOD, 5.; 2019, Valencia. **Proceedings […]** Valencia: EAI, 2019. p. 31-36.

KROLL, A. Grey-box models: concepts and application. **New Frontiers in Computational Intelligence and Its Applications,** v. 57, p. 42-51, 2000.

LAMOND, D. On the value of management history: absorbing the past to understand the present and inform the future. **Management Decision,** v. 43, n. 10, p. 1273-1281, 2005.

LANG, B. **Three studies examining the effects of business analytics on judgment and decision making in accounting.** 2018. Thesis (Ph.D in Philosophy) – University of Central Florida, Florida, 2018.

LANGER, M.; OSTER, D.; SPEITH, T.; HERMANNS, H.; ... BAUM, K. What do we want from Explainable Artificial Intelligence (XAI)? – a stakeholder perspective on XAI and a conceptual model guiding interdisciplinary XAI research. **Artificial Intelligence,** v. 296, 103473, 2021.

LARKIN, C.; DRUMMOND, O. C.; ÁRVAI, J. Will people accept advice from artificial intelligence for consequential risk management decisions? **Journal of Risk Research,** v. 25, n. 4, p. 407-422, 2022.

LARSSON, S.; HEINTZ, F. Transparency in artificial intelligence. **Internet Policy Review,** v. 9, n. 2, 2020.

LAUDON, K. C.; LAUDON, J. P. **Essentials of management information systems**. New Jersey: Pearson, 2011

LAWSON, R. **Management accounting competencies:** fit for purpose in a digital age? Montvale: Institute of Management Accountants, 2019. Disponível em:

https://cma.pace.edu.vn/en/resources/research/management-accounting-competencies-fit-for-purpose-in-a-digital-age. Acesso em: 11 fev. 2023.

LEARY, M. R. **Introduction to behavioral research methods.** 6th ed. New Jersey: Pearson, 2012.

LEAVY, S.; O'SULLIVAN, B.; SIAPERA, E. Data, power and bias in artificial intelligence. **arXiv preprint,** n. 2008.07341, 2020. Disponível em: https://arxiv.org/abs/2008.07341. Acesso em: 11 fev. 2023.

LEE, J. D.; SEE, K. A. Trust in automation: designing for appropriate reliance. **Human Factors,** v. 46, n. 1, p. 50-80, 2004.

LEE, J.; MORAY, N. Trust, control strategies and allocation of function in human-machine systems. **Ergonomics,** v. 35, n. 10, p. 1243-1270, 1992.

LEGG, S.; HUTTER, M. Universal intelligence: a definition of machine intelligence. **Minds and Machines,** v. 17, n. 4, p. 391-444, 2007.

LEHNER, O. M.; ITTONEN, K.; SILVOLA, H.; STRÖM, E.; WÜHRLEITNER, A. Artificial intelligence based decision-making in accounting and auditing: ethical challenges and normative thinking. **Accounting, Auditing & Accountability Journal,** v. 35, n. 9, p. 109-135, 2022.

LEITNER-HANETSEDER, S.; LEHNER, O. M.; EISL, C.; FORSTENLECHNER, C. A profession in transition: actors, tasks and roles in AI-based accounting. **Journal of Applied Accounting Research,** v. 22, n. 3, 2021.

LEOPOLD, T. A.; RATCHEVA, V.; ZAHIDI, S. **The future of jobs:** employment, skills, and workforce strategies for the Fourth Industrial Revolution. Geneva: World Economic Forum, 2016. Disponível em: https://www3.weforum.org/docs/WEF_Future_of_Jobs.pdf. Acesso em: 11 fev. 2023.

LERNER, J. S.; TETLOCK, P. E. Accounting for the effects of accountability. **Psychological Bulletin**, v. 125, n. 2, p. 255, 1999.

LEWIS, M.; SYCARA, K; WALKER, P. The role of trust in human-robot interaction. *In*: ABBASS, H.; SCHOLZ, J.; REID, D. (Eds.) **Foundations of trusted autonomy.** Berlin: Springer, Cham, 2018. v. 117, p. 135-159.

LEYER, M.; DOOTSON, P.; OBERLÄNDER, A. M.; KOWALKIEWICZ, M. Decision-making with artificial intelligence: towards a novel conceptualization of patterns. *In*: PACIFIC ASIA CONFERENCE ON INFORMATION SYSTEMS, 24., 2020, Dubai. **Proceedings [...]** Dubai: PACIS, 2020.

LI, C.; PAN, R.; XIN, H.; DENG, Z. Research on artificial intelligence customer service on consumer attitude and its impact during online shopping. **Journal of Physics: Conference Series,** v. 1575, n. 1, 012192, 2020.

LI, E. Y.; MCLEOD JR, R.; ROGERS, J. C. Marketing information systems in the Fortune 500 companies: past, present, and future. **Journal of Management Information Systems,** v. 10, n. 1, p. 165-192, 1993.

LI, L.; LASSITER, T.; OH, J.; LEE, M. K. Algorithmic hiring in practice: recruiter and HR professional's perspectives on AI use in hiring. *In*: AAAI/ACM CONFERENCE ON AI, ETHICS, AND SOCIETY, 21., 2021. New York. **Proceedings […]** New York: ACM, 2021. p. 166-176.

LI, Z.; ZHENG, L. The impact of artificial intelligence on accounting. **Advances in Social Science, Education and Humanities Research (ASSEHR)**, v. 181, p. 813-816, 2018.

LIBBY, R.; BLOOMFIELD, R.; NELSON, M. W. Experimental research in financial accounting. **Accounting, Organizations and Society,** v. 27, n. 8, p. 775-810, 2002.

LIBBY, T.; SALTERIO, S. E.; WEBB, A. The balanced scorecard: the effects of assurance and process accountability on managerial judgment. **The Accounting Review,** v. 79, n. 4, p. 1075-1094, 2004.

LICHT, K. F.; LICHT, J. F. Artificial intelligence, transparency, and public decision-making. **AI & Society,** v. 35, n. 4, p. 917-926, 2020.

LONGONI, C.; BONEZZI, A.; MOREWEDGE, C. K. Resistance to medical artificial intelligence. **Journal of Consumer Research,** v. 46, n. 4, p. 629-650, 2019.

LOSBICHLER, H.; LEHNER, O. M. Limits of artificial intelligence in controlling and the ways forward: a call for future accounting research. **Journal of Applied Accounting Research**, v. 2, n. 2, p. 365-382, 2021.

LOURENÇO, S. M. Field experiments in managerial accounting research. **Foundations and Trends® in Accounting,** v. 14, n. 1, p. 1-72, 2019.

LOYOLA-GONZALEZ, O. Black-box vs. white-box: understanding their advantages and weaknesses from a practical point of view. **IEEE Access,** v. 7, p. 154096-154113, 2019.

LU, H.; MA, W.; WANG, Y.; ZHANG, M.; WANG, X.; LIU, Y.; CHUA, T.; MA, S. User perception of recommendation explanation: are your explanations what users need? **ACM Transactions on Information Systems,** v. 41, n. 2, p. 1-31, 2022.

LU, J.; WU, D.; MAO, M.; WANG, W.; ZHANG, G. Recommender system application developments: a survey. **Decision Support Systems,** v. 74, p. 12-32, 2015.

LUFTMAN, J.; BEN-ZVI, T. Key issues for IT executives 2010: judicious IT investments continue post-recession. **MIS Quarterly Executive,** v. 9, n. 4, 2010.

MADSEN, M.; GREGOR, S. Measuring human-computer trust. *In:* AUSTRALASIAN CONFERENCE ON INFORMATION SYSTEMS, 11., 2000, Brisbane. **Proceedings […]** Brisbane, Australia: AAIS, 2000. v. 53, p. 6-8.

MAIDA, M.; MAIER, K.; OBWEGESER, N.; STIX, V. The effect of sensitivity analysis on the usage of recommender systems. *In:* **2nd Workshop on Human Decision Making in Recommender Systems in conjunction with the 6th ACM conference on Recommender Systems (RecSys 2012)**, p. 15-18, 2012.

MAKARIUS, E. E.; MUKHERJEE, D.; FOX, J. D.; FOX, K. A. Rising with the machines: a sociotechnical framework for bringing artificial intelligence into the organization. **Journal of Business Research,** v. 120, p. 262-273, 2020.

MANNES, A. Governance, risk, and artificial intelligence. **AI Magazine,** v. 41, n. 1, p. 61-69, 2020.

MARCH, J. G.; SIMON, H. A. **Organizations.** New York: Wiley, 1958.

MARCUS, G.; DAVIS, E. **Rebooting AI:** Building artificial intelligence we can trust. New York: Vintage, 2019.

MARION, J. C.; RIBEIRO, O. M. **Introdução à contabilidade gerencial.** São Paulo: Saraiva, 2017.

MARRONE, M.; HAZELTON, J. The disruptive and transformative potential of new technologies for accounting, accountants and accountability: a review of current literature and call for further research. **Meditari Accountancy Research,** v. 27, n. 5, 2019.

MASTILAK, C.; MATUSZEWSKI, L.; MILLER, F.; WOODS, A. Evaluating conflicting performance on driver and outcome measures: the effect of strategy maps. **Journal of Management Control**, v. 23, n. 2, p. 97-114, 2012.

MCKNIGHT, D. H.; CARTER, M.; THATCHER, J. B.; CLAY, P. F. Trust in a specific technology: an investigation of its components and measures. **ACM Transactions on Management Information Systems**, v. 2, n. 2, 2011.

MCKNIGHT, H.; CARTER, M.; CLAY, P. Trust in technology: development of a set of constructs and measures. **DIGIT 2009 Proceedings**, v. 10, 2009. Disponível em: https://aisel.aisnet.org/digit2009/10/. Acesso em: 11 fev. 2023.

MERCADO, J. E.; RUPP, M. A.; CHEN, J. Y. C.; BARNES, M. J.; ... PROCCI, K. Intelligent agent transparency in human-agent teaming for multi-UxV management. **Human Factors,** v. 58, n. 3, p. 401-415, 2016.

MESKE, C.; BUNDE, E.; SCHNEIDER, J.; GERSCH, M. Explainable artificial intelligence: objectives, stakeholders, and future research opportunities. **Information Systems Management,** v. 39, n. 1, p. 53-63, 2020.

MEYER, J. A. Information overload in marketing management. **Marketing Intelligence & Planning**, v. 6, n. 3, p. 200-209, 1998.

MICHEL, A. H. **The black box, unlocked:** predictability and understandability in military AI. Geneva: United Nations Institute for Disarmament Research, 2020. Disponível em: https://doi.org/10.37559/SecTec/20/AI1. Acesso em: 11 fev. 2023.

MILLER, E. E. P. **Trust in people and trust in technology:** expanding interpersonal trust to technology-mediated interactions. 2015. Thesis (Ph.D in Psychology) – University of South Florida, Florida, 2015.

MINSKY, M. Thoughts about artificial intelligence. *In*: KURZWEIL, R. **The age of intelligent machines.** Cambridge: The MIT Press, 1990.

MISNI, F.; LEE, L. S. A review on strategic, tactical and operational decision planning in reverse logistics of green supply chain network design. **Journal of Computer and Communications,** v. 5, n. 8, p. 83-104, 2017.

MODHA, D. S.; ANANTHANARAYANAN, R.; ESSER, S. K.; … SINGH, R. Cognitive computing. **Communications of the ACM,** v. 54, n. 8, p. 62-71, 2011.

MUIR, B. M. Trust in automation: part I: theoretical issues in the study of trust and human intervention in automated systems. **Ergonomics,** v. 37, n. 11, p. 1905-1922, 1994.

MUIR, B. M.; MORAY, N. Trust in automation: part II: experimental studies of trust and human intervention in a process control simulation. **Ergonomics,** 1996. 39, n. 3, p. 429-460, 1996.

MYERS, B.; HOLLAN, J.; CRUZ, I.; BRYSON, S.; ... IOANNIDIS, Y. Strategic directions in human-computer interaction. **ACM Computing Surveys**, v. 28, n. 4, p. 794-809, 1996.

MYERS, J. L. **Fundamentals of experimental design.** 3. ed. Boston: Allyn and Bacon, 1979.

NAM, C. S.; LYONS, J. B. **Trust in Human-Robot Interaction**. Cambridge: Academic Press, 2020.

NASTESKI, V. An overview of the supervised machine learning methods. **Horizons,** v. 4, p. 51-62, 2017.

NAZAR, M.; ALAM, M. M.; YAFI, E.; SU'UD, M. M. A systematic review of human–computer interaction and explainable artificial intelligence in healthcare with artificial intelligence techniques. **IEEE Access**, v. 9, p. 153316-153348, 2021.

NEIL, C. **Artificial Intelligence:** 4 books in 1: AI for beginners + AI for business + machine learning for beginners + machine learning and artificial intelligence (English edition). London: Alicex, 2020.

NEISSER, U.; BOODOO, G.; BOUCHARD JR, T. J.; BOYKIN, A. W.; ... URBINA, S. Intelligence: knowns and unknowns. **American Psychologist,** v. 51, n. 2, p. 77-101, 1996.

NIELSEN, S. Management accounting and the idea of machine learning. **Economics Working Papers**, Department of Economics and Business Economics, Aarhus University, n. 2020-09, 2020. Disponível em: https://ideas.repec.org/p/aah/aarhec/2020-09.html. Acesso em: 11 fev. 2023.

NOORAIE, M. Decision magnitude of impact and strategic decision-making process output: the mediating impact of rationality of the decision-making process. **Management Decision**, v. 46, n. 4, p. 640-655, 2008.

NORDLANDER, T. E. **AI surveying:** artificial intelligence in business. 2001. Dissertation (Full-Time MSc in Management Science) – Montfort University, Montfort, 2001.

NORRIS, W. R.; PATTERSON, A. E. **Automation, autonomy, and semi-autonomy:** a brief definition relative to robotics and machine systems. Urbana-Champaign: University of Illinois

at Urbana-Champaign, 2019. Disponível em: http://hdl.handle.net/2142/104214.2019. Acesso em: 11 fev. 2023.

NOVIANTY, I. Strategic management accounting: challenges in accounting practices. **Research Journal of Finance and Accounting,** v. 6, n. 9, p. 7-13, 2015.

NUNES, I.; JANNACH, D. A systematic review and taxonomy of explanations in decision support and recommender systems. **User Modeling and User-Adapted Interaction,** v. 27, n. 3, p. 393-444, 2017.

OESTERREICH, T. D.; TEUTEBERG, F.; BENSBERG, F.; BUSCHER, G. The controlling profession in the digital age: understanding the impact of digitisation on the controller's job roles, skills and competences. **International Journal of Accounting Information Systems**, v. 35, n. C, 2019.

OPPENHEIMER, D.; MEYVIS, T.; DAVIDENKO, N. Instructional manipulation checks: detecting satisficing to increase statistical power. **Journal of Experimental Social Psychology**, v. 45, n. 4, p. 867-872, 2009.

OUSSAR, Y.; DREYFUS, G. How to be a gray box: dynamic semi-physical modeling. **Neural Networks**, v. 14, n. 9, p. 1161-1172, 2001.

PANTANO, Eleonora; SCARPI, Daniele. I, robot, you, consumer: Measuring artificial intelligence types and their effect on consumers emotions in service. **Journal of Service Research**, v. 25, n. 4, p. 583-600, 2022.

PASSATH, T.; MERTENS, K. Decision making in lean smart maintenance: criticality analysis as a support tool. **IFAC-PapersOnLine,** v. 52, n. 10, p. 364-369, 2019.

PAULO, S. F. A terceira revolução industrial e a estagnação da acumulação capitalista. **Mundo Livre: Revista Multidisciplinar,** v. 5, n. 2, p. 54-77, 2019.

PERRAULT, R.; SHOHAM, Y.; BRYNJOLFSSON, E.; CLARK, J.; … NIEBLES, J. C. **The AI Index 2019 Annual Report.** Stanford, CA: AI Index Steering Committee, Human-Centered AI Institute, 2019.

PHILLIPS-WREN, G. AI tools in decision making support systems: a review. **International Journal on Artificial Intelligence Tools,** v. 21, n. 2, 1240005, 2012.

POMEROL, J. C. Artificial intelligence and human decision making. **European Journal of Operational Research**, v. 99, n. 1, p. 3-25, 1997.

PREECE, A. Asking "Why" in AI: explainability of intelligent systems–perspectives and challenges. **Intelligent Systems in Accounting, Finance and Management,** v. 25, n. 2, p. 63-72, 2018.

PU, P.; CHEN, L. Trust-inspiring explanation interfaces for recommender systems. **Knowledge-Based Systems**, v. 20, n. 6, p. 542-556, 2007.

PU, P.; CHEN, L. Trust building with explanation interfaces. *In*: INTERNATIONAL CONFERENCE ON INTELLIGENT USER INTERFACES, 11., 2006, Sidney. **Proceedings [...]** Sidney: IUI, 2006. p. 93-100.

PUSTEJOVSKY, J.; KRISHNASWAMY, N. Embodied human computer interaction. **KI-Künstliche Intelligenz**, v. 35, n. 3, p. 307-327, 2021.

RADERMACHER, F. J. Decision support systems: scope and potential. **Decision Support Systems,** v. 12, n. 4-5, p. 257-265, 1994.

RAYMOND, L.; PARÉ, G. Measurement of information technology sophistication in small manufacturing businesses. **Information Resources Management Journal (IRMJ),** v. 5, n. 2, p. 4-16, 1992.

REN, F.; BAO, Y. A review on human-computer interaction and intelligent robots. **International Journal of Information Technology & Decision Making,** v. 19, n. 1, p. 5-47, 2020.

RENN, O. The role of risk perception for risk management. **Reliability Engineering and System Safety**, v. 59, n. 1, p. 49-62, 1998.

RIABACKE, A. Managerial decision making under risk and uncertainty. **IAENG International Journal of Computer Science**, v. 32, n. 4, 2006.

RIBEIRO, M. T.; SINGH, S.; GUESTRIN, C. "Why should I trust you?" Explaining the predictions of any classifier. *In*: ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, 22., 2016, San Francisco. **Proceedings [...]** San Francisco: ACM SIGKDD, 2016. p. 1135-1144.

ROSE, J.M.; ROSE, A.M.; NORMAN, C. S.; MAZZA, C. R. Will disclosure of friendship ties between directors and CEOs yield perverse effects? **The Accounting Review,** v. 89, n. 4, p. 1545-1563, 2014.

ROSE, K.; ELDRIDGE, S.; CHAPIN, L. The internet of things: an overview. **The Internet Society (ISOC)**, v. 80, p. 1-50, 2015.

ROSENFELD, A.; RICHARDSON, A. Explainability in human–agent systems. **Autonomous Agents and Multi-Agent Systems,** v. 33, n. 6, p. 673-705, 2019.

RYAN, C.; BERGIN, M.; WELLS, J. S. G. Theoretical perspectives of adherence to web-based interventions: a scoping review. **International Journal of Behavioral Medicine,** v. 25, n. 1, p. 17-29, 2018.

SAGIROGLU, S.; SINANC, D. Big data: a review. *In*: INTERNATIONAL CONFERENCE ON COLLABORATION TECHNOLOGIES AND SYSTEMS (CTS), 2013, San Diego. **Proceedings [...]** San Diego: IEEE, 2013. p. 42-47.

SAMEK, W.; MÜLLER, K. R. Towards explainable artificial intelligence. *In:* LNCS. **Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)**. Berlim: Springer, Cham, 2019. v. 1170, p. 5-22

SARMAH, S. S. Concept of artificial intelligence, its impact and emerging trends. **International Research Journal of Engineering and Technology (IRJET),** v. 6, n. 11, p. 6, 2019.

SCHAEFER, K. E.; CHEN, J. Y. C.; SZALMA, J. L.; HANCOCK, P. A. A meta-analysis of factors influencing the development of trust in automation: implications for understanding autonomy in future systems. **Human Factors,** v. 58, n. 3, p. 377-400, 2016.

SCHEMMER, M.; HEMMER, P.; NITSCHE, M.; KUHL, N.; VOSSING, M. A meta-analysis of the utility of explainable artificial intelligence in human-AI decision-making. **Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society,** v. 1, n. 1, 2022.

SCHOLTEN, L.; VAN KNIPPENBERG, D.; NIJSTAD, B. A.; DE DREU, C. K. Motivated information processing and group decision-making: effects of process accountability on information processing and decision quality. **Journal of Experimental Social Psychology,** v. 43, n. 4, p. 539-552, 2007.

SCHWAB, K. **The fourth industrial revolution**. New York: Currency, 2017.

SEARLE, J. R. Minds, brains, and programs. **Behavioral and Brain Sciences,** v. 3, n. 3, p. 417-424, 1980.

SENA, T. H. R. **As instituições financeiras na era da tecnologia de informação e comunicação:** um novo modelo de relacionamento com a sociedade. Ponta Grossa, Atena, 2020.

SHAH, J. Algorithm for human-assisting robots. **Advanced Manufacturing Technology,** v. 33, n. 8, p. 3-4, 2012.

SHARMA, R.; SHISHODIA, A.; GUNASEKARAN, A.; MIN, H.; MUNIM, Z. H. The role of artificial intelligence in supply chain management: mapping the territory. **International Journal of Production Research,** v. 60, n. 24, p. 1-24, 2022.

SHATILO, O. The impact of external and internal factors on strategic management of innovation processes at company level. **Ekonomika**, v. 98, n. 2, p. 85-96, 2019.

SHI, Y. The impact of artificial intelligence on the accounting industry. *In:* AISC. **Advances in Intelligent Systems and Computing**. Berlim: Springer, Cham, 2019. v. 928, p. 971-978

SHIN, D. The effects of explainability and causability on perception, trust, and acceptance: implications for explainable AI. **International Journal of Human-Computer Studies,** v. 146, p. 102551, 2021.

SHINDE, P. P.; SHAH, S. A review of machine learning and deep learning applications. *In*: INTERNATIONAL CONFERENCE ON COMPUTING COMMUNICATION CONTROL AND AUTOMATION, 4., 2018, Pune, India. **Proceedings […].** Pune: IEEE, 2018. p. 1-6.

SHIVAKUMAR, R. How to tell which decisions are strategic. **California Management Review**, v. 56, n. 3, p. 78-97, 2014.

SHRESTHA, Y. R; BEN-MENAHEM, S. M.; VON KROGH, G. Organizational decision-making structures in the age of artificial intelligence. **California Management Review**, v. 61, n. 4, 000812561986225, 2019.

SIEGEL-JACOBS, K.; YATES, J. F. Effects of procedural and outcome accountability on judgment quality. **Organizational Behavior and Human Decision Processes**, v. 65, n. 1, p. 1-17, 1996.

SIMON, H. A. A behavioral model of rational choice. **Quarterly Journal of Economics**, v. 69, n. 1, p. 99-118, 1955.

SIMON, H. A. Making management decisions: the role of intuition and emotion. **Academy of Management Perspectives**, v. 1, n. 1, p. 57-64, 1987.

SIMON, H. A. The compensation of executive. **Sociometry,** v. 20, p. 32-35, 1957.

SIMON, H. A. Theories of bounded rationality. **Decision and Organization,** v. 1, n. 1, p. 161-176, 1972.

SIMONSON, I.; NYE, P. The effect of accountability on susceptibility to decision errors. **Organizational Behavior and Human Decision Processes**, v. 51, n. 3, p. 416-446, 1992.

SINGH, D.; TRIPATHI, G.; JARA, A. J. A survey of Internet-of-Things: future vision, architecture, challenges and services. *In*: IEEE WORLD FORUM ON INTERNET OF THINGS (WF-IOT), 2014, Seoul. **Proceedings [...]** Seoul: IEEE, 2014. p. 287-292.

SINGH, J. V. Performance, slack, and risk taking in organizational decision making. **Academy of Management Journal**, v. 29, n. 3, p. 562-585, 1986.

SKALFIST, P.; MIKELSTEN, D.; TEIGENS, V. **Inteligência artificial:** a quarta revolução industrial. Cambridge: Cambridge Stanford Books, 2019.

SMITH, E. R. Research design. *In*: REIS, H. T.; JUDD, C. M. (Eds.). **Handbook of research methods in social and personality psychology.** Cambridge: Cambridge University Press, 2000. p. 17-39.

STANCIU, V.; TINCA, A. Solid knowledge management: the ingredient companies need for performance: a Romanian insight. **Accounting and Management Information Systems,** v. 16, n. 1, p. 147-163, 2017.

STRUSANI, D.; HOUNGBONON, G. V. The role of artificial intelligence in supporting development in emerging markets, **EMCompass**, n. 69, 2019.

SUTTON, S. G.; HOLT, M.; ARNOLD, V. The reports of my death are greatly exaggerated: artificial intelligence research in accounting. **International Journal of Accounting Information Systems**, v. 22, p. 60-73, 2016.

SWELLER, J. Cognitive load during problem solving: effects on learning. **Cognitive Science,** v. 12, n. 2, p. 257–285, 1988.

TAMMINEN, P. **Operationalizing transparency and explainability in artificial intelligence through**. 2022. Dissertation (Master in Economy) – University of Turku, Turku, 2022.

TAYLOR, A. M. The future of the professions: how technology will transform the work of human experts (Book Review). **Social Work Education**, v. 35, p. 371-372, 2016.

TAYLOR, R. N. Age and experience as determinants of managerial information processing and decision-making performance. **Academy of Management Journal**, v. 18, n. 1, p. 74-81, 1975.

TENG, X. Discussion about artificial intelligence's advantages and disadvantages compete with natural intelligence. **Journal of Physics: Conference Series**, v. 1187, n. 3, 032083, 2019.

TROJANOWSKI, M.; KUŁAK, J. The impact of moderators and trust on consumer's intention to use a mobile phone for purchases. **Central European Management Journal**, v. 25, n. 2, p. 91-116, 2017.

TRONCO, P. B.; LÖBLER, M. L; SANTOS, L. G.; NISHI, J. M. Heurística da ancoragem na decisão de especialistas: resultados sob teste de manipulação. **Revista de Administração Contemporânea,** v. 23, p. 331-350, 2019.

TRONG, H. B.; KIM, U. B. T. Application of information and technology in supply chain management: case study of artificial intelligence–a mini review. **European Journal of Engineering and Technology Research**, v. 5, n. 12, p. 19-23, 2020.

TRUNK, A.; BIRKEL, H.; HARTMANN, E. On the current state of combining human and artificial intelligence for strategic organizational decision making. **Business Research,** v. 13, n. 3, p. 875-919, 2020.

TÜREGÜN, N. Impact of technology in financial reporting: the case of Amazon Go. **Journal of Corporate Accounting & Finance,** v. 30, n. 3, p. 90-95, 2019.

TURING, A. M. Computing machinery and intelligence. **Mind,** v. 59, n. 236, p. 433-460, 1950.

TURNER, S. E. **Analysis of operational decision making by nurse executives in hospital settings:** the viability of the Turner decision-making model. 2000. Thesis (Ph.D in Psycology) – The University of Alabama, Huntsville, 2000.

TURNER, S. E. The Turner decision-making model: a four-perspective framework. **Journal of the Alabama Academy of Science**, v. 74, n. 2, p. 116-117, 2003.

UNGSON, G. R.; BRAUNSTEIN, D. N.; HALL, P. D. Managerial information processing: a research review. **Administrative Science Quarterly,** v. 26, n. 1, p. 116-134, 1981.

VENKATESH, V.; BALA, H. Technology acceptance model 3 and a research agenda on interventions. **Decision Sciences**, v. 39, n. 2, p. 273-315, 2008.

VON ESCHENBACH, W. J. Transparency and the black box problem: why we do not trust AI. **Philosophy & Technology,** v. 34, n. 4, p. 1607-1622, 2021.

VULTUREANU-ALBIŞI, A.; BĂDICĂ, C. Recommender systems: an explainable AI perspective. *In*: INTERNATIONAL CONFERENCE ON INNOVATIONS IN INTELLIGENT SYSTEMS AND APPLICATIONS (INISTA), 2021, Kocaeli, Turkey. **Proceedings […]** Kocaeli: IEEE, 2021.

WALLACH, W. Implementing moral decision-making faculties in computers and robots. **AI and Society,** v. 22, n. 4, p. 463-475, 2008.

WAMBA-TAGUIMDJE, S.-L.; WAMBA, S. F.; KAMDJOUG, J. R. K.; WANKO, C. E. T. Influence of artificial intelligence (AI) on firm performance: the business value of AI-based transformation projects. **Business Process Management Journal**, v. 26, n. 7, p. 1893-1924, 2020.

WANG, L.; RAU, P.-L. P.; EVERS, V.; ROBINSON, B. K.; HINDS, P. When in Rome: the role of culture & context in adherence to robot recommendations. *In*: ACM/IEEE INTERNATIONAL CONFERENCE ON HUMAN-ROBOT INTERACTION (HRI), 5., 2010, Osaka. **Proceedings [...]**. Osaka: IEEE, 2010. p. 359-366.

WARD, P. T.; DURAY, R. Manufacturing strategy in context: environment, competitive strategy and manufacturing strategy. **J Oper Manag,** v. 18, n. 2, p. 123-138, 2000.

WATSON, H. J. Preparing for the cognitive generation of decision support. **MIS Quarterly Executive**, v. 16, n. 3, p. 153-169, 2017.

WEIMEI, C. On the transition from financial accounting to management accounting in the era of artificial intelligence. **Advances in Vocational and Technical Education**, v. 3, n. 4, 2021, p. 78-82, 2021.

WILSON, H. J.; DAUGHERTY, P. R. Collaborative intelligence: humans and AI are joining forces. **Harvard Business Review,** v. 96, n. 4, p. 114-123, 2018.

XU, W. Toward human-centered AI: a perspective from human-computer interaction. **Interactions,** v. 26, n. 4, p. 42-46, 2019.

YAGODA, R. E.; GILLAN, D. J. You want me to trust a ROBOT? The development of a human-robot interaction trust scale. **International Journal of Social Robotics**, v. 4, n. 3, p. 235–248, 2012.

YAMPOLSKIY, R. V. Unexplainability and incomprehensibility of AI. **Journal of Artificial Intelligence and Consciousness**, v. 7, n. 2, p. 277-291, 2020.

YANG, X.; HAUGEN, S. Classification of risk to support decision-making in hazardous processes. **Safety Science**, v. 80, p. 115-126, 2015.

YOUNG, B. S.; ARTHUR JR, W.; FINCH, J. Predictors of managerial performance: More than cognitive ability. **Journal of Business and Psychology**, v. 15, n. 1, p. 53-72, 2000.

ZEDNIK, C. Solving the black box problem: a normative framework for explainable artificial intelligence. **Philosophy & Technology**, v. 34, n. 2, p. 265-288, 2021.

ZERILLI, J.; KNOTT, A.; MACLAURIN, J.; GAVAGHAN, C. Transparency in algorithmic and human decision-making: is there a double standard? **Philosophy & Technology**, v. 32, n. 4, p. 661-683, 2019.

ZHANG, Y.; NAUMAN, U. Artificial unintelligence: anti-intelligence of intelligent algorithms. In: INTERNATIONAL CONFERENCE ON INTELLIGENCE SCIENCE, 2., 2018, Beijing. **Proceedings [...]** Beijing: ICIS, 2018. p. 333-339.

ZHANG, Chao et al. Ethical impact of artificial intelligence in managerial accounting. **International Journal of Accounting Information Systems**, v. 49, p. 100619, 2023.

ZHAO, X.; CAO, W.; ZHU, H.; MING, Z.; ASHFAQ, R. A. R. An initial study on the rank of input matrix for extreme learning machine. **International Journal of Machine Learning and Cybernetics**, v. 9, n. 1-3, p. 867-879, 2018.

ZHOU, M. **Understanding and improving recommender systems' performance in the presence of practical user-, item-, and marketing-oriented considerations**. 2022. Tese (Ph.D in Business Administration) – University of Minnesota, Minneapolis, 2022.

ZUBOFF, S. **In the age of the smart machine:** the future of work and power. New York, Basic Books, 1988.

APÊNDICE A – SUMÁRIO DE JUSTIFICATIVAS METODOLÓGICAS

Aspecto	Opção Teórico-Metodológica	Fundamento	Justificativa
Base Teórica	*Human-Computer Interaction*	Card, Moran e Newell (1983)	Fundamento teórico comum aos estudos que versam sobre as interações humanas com ferramentas tecnológicas. Frequentemente utilizado também para as pesquisam que tratam do design de interfaces com vistas ao aprimoramento da relação homem-máquina para obtenção de melhores resultados.
Estratégia de Pesquisa	Experimental	Gow, Larcker e Reiss (2016)	Tendo o estudo a proposta de identificar relações causais, a metodologia experimental - quando possível - é a forma de investigação costumeiramente recomendada pelos metodólogos
Tipo de Experimento	*Artefactual Field Experiment*	Lourenço (2019)	A utilização de um grupo não-padronizado de participantes, característica deste tipo experimental, proporciona uma validade externa maior que a dos experimentais puramente laboratoriais. Com uma já elevada validade interna, em razão do rígido controle de variáveis exógenas, esta opção figura como a mais adequada frente às demais possibilidades experimentais
Desenho Experimental	*Between Subjects Design*	Leary (2012)	Quando a finalidade é compreender o comportamento de diferentes grupos experimentais (como é o caso), estando cada um deles submetido a apenas uma condição experimental, este desenho é adequado e aplicável
Participantes	Gerentes, Consultores e Assessoras atuantes em Unidades Táticas	Myers (1979)	Escolha realizada conforme os propósitos específicos da pesquisa.
Tamanho da Amostra	Mínimo de 100 indivíduos	Smith (2000)	Segregada em quatro grupos experimentais (cada um com no mínimo 25 indivíduos), a quantidade mínima de participantes da amostra se alinha - com folga - à práxis atual da pesquisa em contabilidade (que é de estabelecer cerca de 20-25 participantes por condição experimental), e é equalizada mediante distribuição aleatória, mas coesa, dos indivíduos entre os grupos.
Periodicidade da Tarefa Experimental	Unifásica	-	A comparação ao longo do tempo não é propósito deste estudo. Focar na captura e na explicação de determinada reação humana (no caso presente, na reação diante de uma decisão orçamentária com suporte de inteligência artificial), em uma perspectiva pontual – isto é: transversal no tempo – previne o trabalho contra a obsolescência própria dos estudos que versam exclusivamente sobre tecnologia, e dispensa a necessidade de uma abordagem longitudinal.
Instrumento de Coleta de Dados	Survey Monkey®	-	Opção amplamente utilizada em estudos similares ao proposto. Permite desenvolver com facilidade o instrumento de pesquisa, distribuí-lo virtual e aleatoriamente aos respondentes, e analisar com profundidade e clareza as respostas obtidas. O preenchimento da tarefa por parte dos respondentes também é bastante intuitivo.
Análise de Resultados	ANOVA e ANCOVA	Gujarati e Porter (2011)	Os métodos de análise de variância e covariância tem sido tradicionalmente aplicados em estudos experimentais na área de contabilidade. A diferença entre as médias dos grupos é facilmente capturada por meio destes métodos – o que explica e justifica o uso deles dentro do presente escopo.

APÊNDICE B – MAPA CONCEITUAL BÁSICO

Termo	Conceito	Fundamento
Artificial Intelligence	*"Ability of machines to make decisions and learn in a similar manner as humans. It is commonly known as a branch of computer science that deals with the simulation of behaviors of computer intelligence. Commonly abbreviated as AI, artificial intelligence refers to a computer system that can complete tasks that normally require the intelligence of humans, including recognition of speech, visual perception, decision-making, and language translations".*	Neil (2020)
Decision	*"Refers to making up one's mind about the issue at hand and taking a course of action".*	Bonner (1999)
Operational Decision	*"Covers all the unforeseen events, particular situations that arise during the execution of the operations. The operational decision relates to the day-to-day operation with the aim of making the process of resource transformation as efficient as possible. They are very frequent, their impact is short-term".*	Imane e Driss (2017)
Strategic Decision	*"Concerns the main axes of development of the company, and the relations of the company with the external environment. This type of decision determines the future of the company by setting the fundamental orientations that commit it in the long term".*	Imane e Driss (2017)
Judgement	*"Refers to forming an idea, opinion, or estimate about an object, an event, a state, or another type of phenomenon".*	Bonner (1999)
Process	*"Describes how the automation operates"*	Lee e See (2004)
Transparency	*"Quality of an interface pertaining to its abilities to afford an operator's comprehension about an intelligent agent's intent, performance, future plans, and reasoning process".*	Chen *et al.* (2014)
Adherence	*"The active use of an intervention as prescribed by those delivering the programme"*	Cavanagh (2010)

APÊNDICE C – TRATAMENTO DE VIESES

Fase	Tipo de *Confounder*	Significado	Decisão Metodológica
Formulação	Viés do Participante	Quando a experiência pessoal (atual ou pretérita) do participante cria ou alimenta exagerado preconceito com relação às questões que lhe estão sendo formuladas ou com relação à sua própria participação no estudo	Compatibilizar o perfil dos potenciais participantes com o perfil desejado para a amostra
	Viés do Pesquisador	Quando a experiência pessoal (atual ou pretérita) do pesquisador cria ou alimenta exagerado preconceito na formulação das hipóteses, na elaboração dos instrumentos de coleta e análise de dados, e/ou na conclusão e interpretação dos resultados	Triangulação de todas as decisões metodológicas e construções textuais
Seleção	Viés de Autosseleção	Quando é dada aos próprios participantes da pesquisa a possibilidade de escolher a qual grupo de estudo irão pertencer	Aleatorizar a seleção de participantes
	Viés de Alocação	Quando o pesquisador escolhe uma forma não-aleatória de distribuição dos participantes da pesquisa entre os grupos de estudo	Aleatorizar a distribuição de participantes
	Viés de Participação	Quando qualquer fator (interno ou externo) influencia ou mesmo obriga o participante a fazer parte da amostra	Colher Termo de Consentimento Livre e Esclarecido
	Viés de Amostra	Quando a amostra é composta por uma fração não representativa da população estudada	Validar o perfil da amostra a partir de questões sociodemográficas
Coleta	Viés de Cognição Numérica	Quando, diante de um conjunto de dados numéricos, dadas as limitações do sistema humano de processamento de informações, o participante de uma pesquisa cede à tendência de responder rapidamente, sem analisar os dados que lhe estão sendo apresentados	Evitar apresentação de tabelas com números em excesso de forma a simplificar questões que demandem resposta numérica
	Viés de Perdas na Amostra	Quando uma amostra sofre alterações significativas em suas características como decorrência da exclusão/desistência de participantes	Uniformizar a amostra de modo que eventuais exclusões e/ou desistências não a descaracterizem
Mensuração	Viés de Mensuração	Uso de técnicas inadequadas de mensuração ou subjetividade na escala de medidas	Utilizar escalas consagradas pela literatura e realizar pré-testes
	Viés de Aferição	Quando há diferenças (ou mesmo mudança total) quanto aos instrumentos/técnicas de mensuração	Unificar e manter os instrumentos/técnicas de mensuração
Análise	Viés de Confirmação	Quando, antes de checar a validação de todos os pressupostos estabelecidos no *design* do método, o pesquisador enxerga nos achados a confirmação de seu ponto de vista pessoal e de suas hipóteses	Redigir a discussão dos resultados só após a completa análise dos achados
	Erro Tipo I	Rejeitar a hipótese nula quando ela é verdadeira	Aplicar testes de validação
	Erro Tipo II	Não rejeitar a hipótese nula quando ela é falsa	Aplicar testes de validação
	Causalidade Reversa	Atribuir a causa de um determinado fenômeno a uma razão diversa daquela que efetivamente o origina	Triangulação na revisão da análise
	Outlier(s)	Ter o resultado influenciado pela presença de um elemento discrepante em relação à média e cujo peso é relevante a ponto de distorcer a compreensão da amostra e dos achados	Antes de discutir/interpretar os achados, eliminar outlier(s) dos resultados, se houver
	Overfitting	Superestimar os achados	Triangulação na revisão da análise
	Underfitting	Subestimar os achados	Triangulação na revisão da análise

APÊNDICE D – INSTRUMENTO DE COLETA DE DADOS

UNIVERSIDADE FEDERAL DE PERNAMBUCO

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Prezado(a),

Você está sendo convidado(a) a participar voluntariamente de um estudo realizado pelo doutorando e funcionário bolsista do Banco do Brasil, Gustavo Souza. Este estudo faz parte de um projeto de pesquisa desenvolvido na Universidade Federal de Pernambuco, sob orientação do Prof. Cláudio de Araújo Wanderley, Ph.D. (UFPE) e coorientação do Prof. Dr. Andson Braga de Aguiar (USP).

O objetivo da pesquisa é examinar a percepção das pessoas quanto ao uso de uma ferramenta tecnológica para auxiliar o processo decisório. Não há respostas certas ou erradas, e você não será julgado(a) ou avaliado(a) com base no que você responde. O tempo médio estimado de preenchimento das respostas é de até 10 minutos.

Como parte deste experimento de pesquisa, você responderá também algumas perguntas de cunho sociodemográfico com a única intenção de entender os dados gerais dos(as) participantes, sem qualquer intuito de identificar sua participação. É garantido o anonimato, sigilo, privacidade e confidencialidade das respostas. As respostas serão analisadas em conjunto com as de outros(as) participantes, o que garante que seus dados não serão identificados.

Em qualquer fase do preenchimento, você terá o direito de abandonar o experimento. Sua decisão de não participar da pesquisa não afetará de forma alguma seu relacionamento com os pesquisadores responsáveis, tampouco com a universidade ou com a empresa onde você trabalha.

Pede-se ler com atenção todos os esclarecimentos e instruções contidas no instrumento de pesquisa. Se você tiver alguma dúvida, comentário, ou desejar receber os resultados com as conclusões deste estudo, entre em contato com o pesquisador responsável:

GUSTAVO HENRIQUE COSTA SOUZA
Fone: (81) 99885-4039
E-mail: gustavo.csouza@ufpe.br / g.souza@bb.com.br

Diante desses esclarecimentos, confira e assinale uma das duas alternativas a seguir:

- Aceito, voluntariamente, participar do estudo - tal como proposto - na qualidade de respondente
- Não concordo em participar (neste caso, encerre sua participação).

INSTRUÇÕES ESPECÍFICAS

CENÁRIO I - TRANSPARÊNCIA ALTA EM DECISÃO OPERACIONAL

Obrigado por aceitar participar deste estudo!

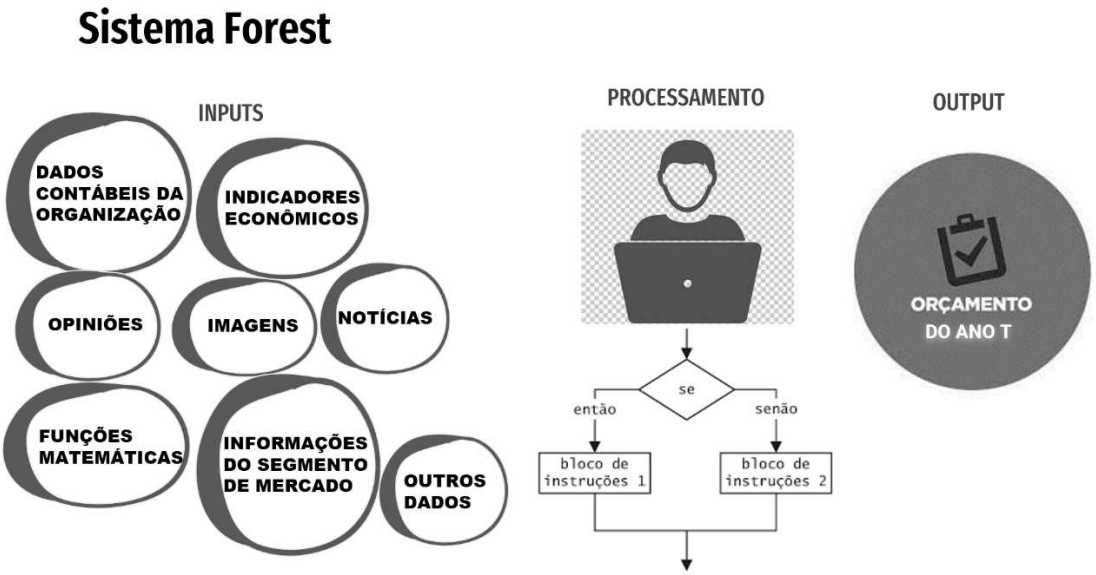
Considere que você é o(a) gestor(a) encarregado(a) da parte de suprimentos e estocagem do BANCO ART S/A. É sua responsabilidade tomar decisões de curto prazo sobre aquisições, reposições e alienações de bens móveis.

O BANCO ART S/A constatou que – devido ao uso intenso – a vida útil do teclado dos computadores tem se reduzido e, como consequência, o número de pedidos de substituição deste tipo de equipamento tem aumentado. Funcionários alegam que a autorização para adquirir um novo teclado demora muito a ser concedida – o que acaba impactando o fluxo das atividades cotidianas de atendimento ao público nas agências.

Para solucionar a questão e agilizar a reposição de teclados danificados, a empresa cogita fazer uma compra massificada de teclados e distribui-los entre as agências para que sejam mantidos em almoxarifado e possam estar à disposição em caso de necessidade. Você deve determinar o valor que será destinado a esta compra massificada de teclados.

A política tradicional da organização diz que deve-se apurar a média de gastos com aquisição de teclados nos últimos 3 anos e corrigir o valor a partir de um índice de inflação. Assim, de acordo com esta metodologia, deve ser realizado um gasto de R$ 485.000,00 com aquisição de teclados.

Para auxiliar este tipo de decisão, o BANCO ART S/A disponibiliza também uma ferramenta tecnológica desenvolvida a partir de modelos de inteligência artificial chamada Sistema Forest. Neste sistema, as regras de cálculo utilizadas para prever gastos são estabelecidas por programadores. O funcionamento do Sistema Forest, portanto, depende de escolhas humanas pré-definidas. De acordo com ele, deve ser realizado um gasto de R$ 440.000,00 com aquisição de teclados.



CENÁRIO II - TRANSPARÊNCIA BAIXA EM DECISÃO OPERACIONAL

Obrigado por aceitar participar deste estudo!

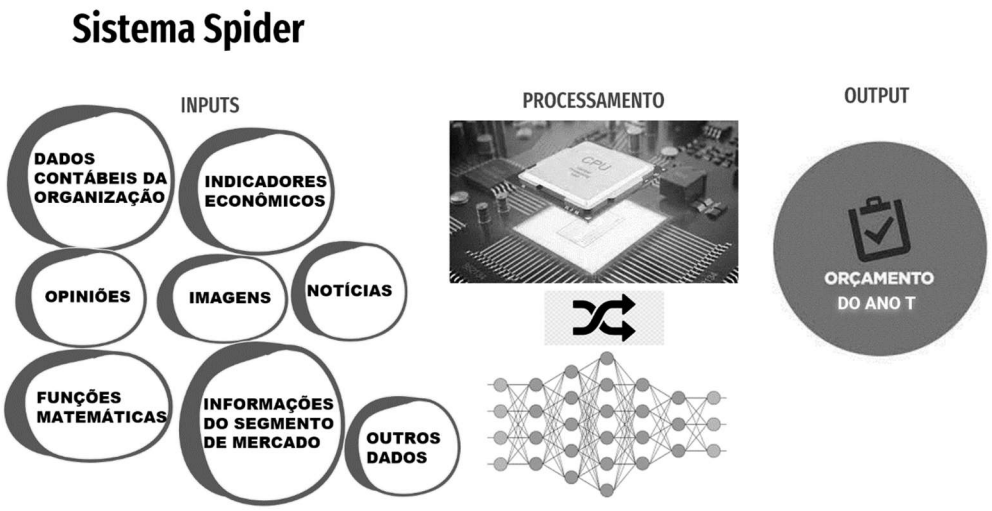
Considere que você é o(a) gestor(a) encarregado(a) da parte de suprimentos e estocagem do BANCO ART S/A. É sua responsabilidade tomar decisões de curto prazo sobre aquisições, reposições e alienações de bens móveis.

O BANCO ART S/A constatou que – devido ao uso intenso – a vida útil do teclado dos computadores tem se reduzido e, como consequência, o número de pedidos de substituição deste tipo de equipamento tem aumentado. Funcionários alegam que a autorização para adquirir um novo teclado demora muito a ser concedida – o que acaba impactando o fluxo das atividades cotidianas de atendimento ao público nas agências.

Para solucionar a questão e agilizar a reposição de teclados danificados, a empresa cogita fazer uma compra massificada de teclados e distribui-los entre as agências para que sejam mantidos em almoxarifado e possam estar à disposição em caso de necessidade. Você deve determinar o valor que será destinado a esta compra massificada de teclados.

A política tradicional da organização diz que deve-se apurar a média de gastos com aquisição de teclados nos últimos 3 anos e corrigir o valor a partir de um índice de inflação. Assim, de acordo com esta metodologia, deve ser realizado um gasto de R$ 485.000,00 com aquisição de teclados.

Para auxiliar este tipo de decisão, o BANCO ART S/A disponibiliza também uma ferramenta tecnológica desenvolvida a partir de modelos de inteligência artificial chamada Sistema Spider. Neste sistema, as regras de cálculo utilizadas para prever gastos são estabelecidas pelo próprio sistema. O funcionamento do Sistema Spider, portanto, depende de associações aleatórias entre os dados que o alimentam. De acordo com ele, deve ser realizado um gasto de R$ 440.000,00 com aquisição de teclados.



CENÁRIO III -TRANSPARÊNCIA ALTA EM DECISÃO ESTRATÉGICA

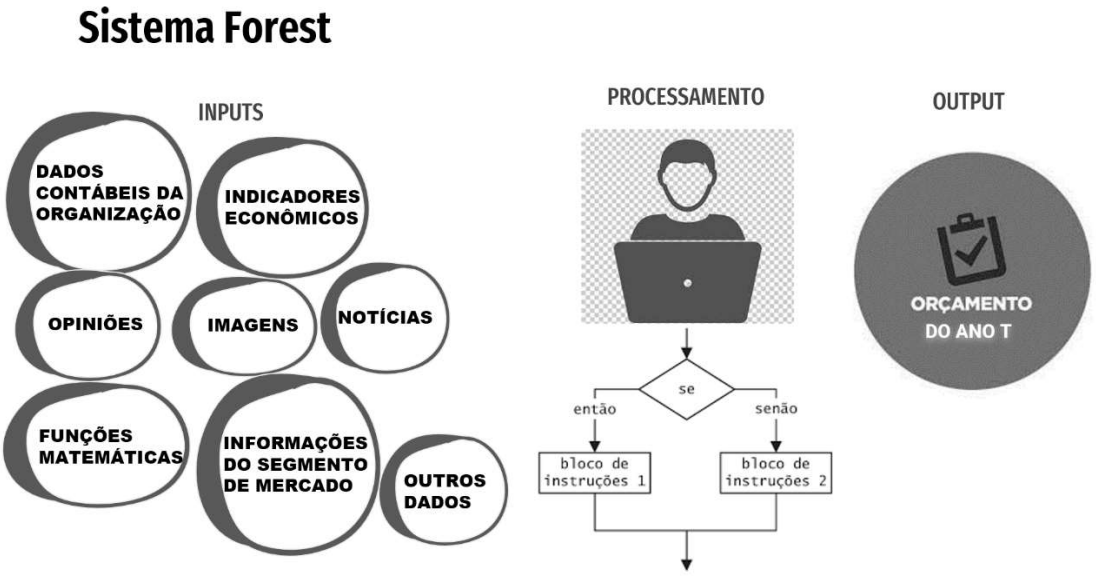
Obrigado por aceitar participar deste estudo!

Considere que você é o(a) Diretor(a) de Infraestrutura do BANCO ART S/A. É sua responsabilidade, como membro do Alto Escalão da organização, decidir sobre investimentos na estrutura física da organização, com vistas à expansão dos negócios no longo prazo.

O BANCO ART S/A constatou que os clientes têm demandado um serviço de orientação específico quanto ao uso de cartões de crédito. Bancos concorrentes, inclusive, já criaram plataformas de atendimento especializadas em assessorar os clientes com relação a isso. Para se alinhar ao mercado, e visando atingir seus objetivos de longo prazo em termos de captação e retenção de clientes, o BANCO ART S/A cogita fazer o mesmo, e você deve determinar o valor a ser investido na criação desta nova plataforma de atendimento.

A política tradicional do BANCO ART S/A diz que deve-se apurar a média de investimentos de capital feitos pela organização nos últimos 3 anos e corrigir o valor a partir de um índice de inflação. Assim, de acordo com esta metodologia, devem investidos R$ 485.000,00 na criação desta nova plataforma de atendimento.

Para auxiliar este tipo de decisão, o BANCO ART S/A disponibiliza também uma ferramenta tecnológica desenvolvida a partir de modelos de inteligência artificial chamada Sistema Forest. Neste sistema, as regras de cálculo utilizadas para prever gastos são estabelecidas por programadores. O funcionamento do Sistema Forest, portanto, depende de escolhas humanas pré-definidas. De acordo com ele, devem ser investidos R$ 440.000,00 na criação da nova plataforma de atendimento.



CENÁRIO IV- TRANSPARÊNCIA BAIXA EM DECISÃO ESTRATÉGICA

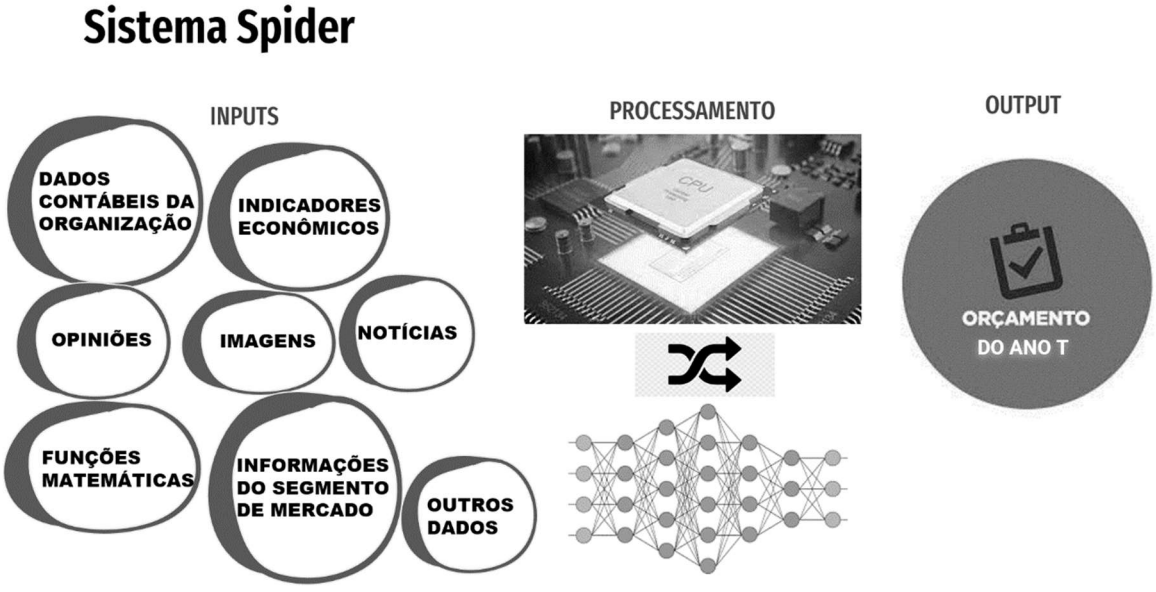
Obrigado por aceitar participar deste estudo!

Considere que você é o(a) Diretor(a) de Infraestrutura do BANCO ART S/A. É sua responsabilidade, como membro do Alto Escalão da organização, decidir sobre investimentos na estrutura física da organização, com vistas à expansão dos negócios no longo prazo.

O BANCO ART S/A constatou que os clientes têm demandado um serviço de orientação específico quanto ao uso de cartões de crédito. Bancos concorrentes, inclusive, já criaram plataformas de atendimento especializadas em assessorar os clientes com relação a isso. Para se alinhar ao mercado, e visando atingir seus objetivos de longo prazo em termos de captação e retenção de clientes, o BANCO ART S/A cogita fazer o mesmo, e você deve determinar o valor a ser investido na criação desta nova plataforma de atendimento.

A política tradicional do BANCO ART S/A diz que deve-se apurar a média de investimentos de capital feitos pela organização nos últimos 3 anos e corrigir o valor a partir de um índice de inflação. Assim, de acordo com esta metodologia, devem investidos R$ 485.000,00 na criação desta nova plataforma de atendimento.

Para auxiliar este tipo de decisão, o BANCO ART S/A disponibiliza também uma ferramenta tecnológica desenvolvida a partir de modelos de inteligência artificial chamada Sistema Spider. Neste sistema, as regras de cálculo utilizadas para prever gastos são estabelecidas pelo próprio sistema. O funcionamento do Sistema Spider, portanto, depende de associações aleatórias entre os dados que o alimentam. De acordo com ele, devem ser investidos R$ 440.000,00 na criação da nova plataforma de atendimento.



TAREFA EXPERIMENTAL[6]

Diante das informações apresentadas, por favor responda:

1. Em uma escala de 0 a 100, qual a probabilidade de você votar em concordância com a recomendação do Sistema Forest/Spider? _____
2. Na sua opinião, que valor deve ser destinado à compra massificada de teclados? _____
3. Na sua opinião, que valor deve ser investido na criação da nova plataforma de atendimento?

VERIFICAÇÃO DE MANIPULAÇÃO

As regras que a ferramenta de inteligência artificial utiliza para prever valores são:

☐ Estabelecidas por programadores e, portanto, dependem de escolhas humanas predefinidas
☐ Estabelecidas pelo próprio sistema e, portanto, dependem de associações aleatórias entre os dados

O tipo de decisão que me foi solicitada...

☐ Caracteriza-se como decisão de curto prazo e visa garantir a disponibilidade de bens móveis utilizados nas atividades cotidianas
☐ Caracteriza-se como decisão de longo prazo e visa garantir a expansão dos negócios em um cenário de concorrência

FAMILIARIDADE COM IA, HABILIDADE COM IA E CONFIANÇA NA TECNOLOGIA

Qual das opções melhor descreveria a sua familiaridade com a inteligência artificial?

☐ Nada familiar
☐ Pouco familiar
☐ Moderadamente familiar
☐ Muito familiar
☐ Extremamente familiar

Qual o seu grau de habilidade com ferramentas de inteligência artificial?

☐ Nenhum
☐ Básico
☐ Principiante
☐ Intermediário
☐ Avançado
☐ Expert

Indique o seu nível de concordância em relação às seguintes afirmações [Escala Likert: (1) discordo totalmente; (2) discordo parcialmente; (3) não concordo nem discordo; (4) concordo parcialmente; (5) concordo totalmente]:

Dimensão	Afirmação	1	2	3	4	5
Confiança na Tecnologia	Eu normalmente confio em uma tecnologia até que ela me dê uma razão para não mais confiar					
	Eu normalmente dou à tecnologia o benefício da dúvida quando a utilizo pela primeira vez					
	Minha postura típica é de confiar em novas tecnologias até que elas me provem que eu não devo confiar nelas					

[6] Por uma questão de coerência com o texto de cada cenário descrito, os grupos experimentais que receberam os cenários I e II não receberam a pergunta 3. E os grupos que receberam os cenários III e IV não receberiam a pergunta 2.

QUESTÕES PÓS-EXPERIMENTAIS

Questões Sociodemográficas

Qual a sua idade? _____

Qual o seu gênero?

- ☐ Masculino
- ☐ Feminino
- ☐ Não desejo informar

Qual o último nível (concluído) da sua formação escolar/acadêmica?

- ☐ Ensino Médio
- ☐ Graduação
- ☐ Especialização
- ☐ Mestrado
- ☐ Doutorado

Qual a sua área de formação escolar/acadêmica?

- ☐ Ciências Agrárias
- ☐ Ciências Biológicas
- ☐ Ciências da Saúde
- ☐ Ciências Exatas
- ☐ Ciências Humanas
- ☐ Ciências Sociais Aplicadas
- ☐ Engenharias
- ☐ Linguística, Letras e Artes
- ☐ Outras
- ☐ Não se aplica

Qual o seu cargo atual?

- ☐ Gerente de Administração
- ☐ Gerente de Negócios
- ☐ Gerente de Núcleo
- ☐ Superintendente
- ☐ Gerente de Cobrança e Reestruturação de Ativos Operacionais
- ☐ Gerente Geral de Unidade
- ☐ Assessor(a)
- ☐ Consultor(a)

Há quanto anos (completos) você está no seu cargo atual? _____
Há quantos anos (completos) você exerce cargos de natureza gerencial nesta organização? _____
Há quantos anos (completos) você trabalha nesta organização? _____

APÊNDICE E – CODIFICAÇÃO DO INSTRUMENTO DE COLETA DE DADOS

CÓDIGO	QUESTÃO	ESCALA
pconcord	Em uma escala de 0 a 100, qual a probabilidade de você votar em concordância com a recomendação...	0 - 100
valor	Na sua opinião, que valor deve ser...	–
resptrans	As regras que a ferramenta de inteligência artificial utiliza para prever valores são...	0 – Baixa 1 – Alta
respdec	O tipo de decisão que me foi solicitada...	0 – Operacional 1 – Estratégica
conftec1	Eu normalmente confio em uma tecnologia até que ela me dê uma razão para não mais confiar	1 – Discordo totalmente 2 – Discordo parcialmente 3 – Não concordo nem discordo 4 – Concordo parcialmente 5 – Concordo totalmente
conftec2	Eu normalmente dou à tecnologia o benefício da dúvida quando a utilizo pela primeira vez	
conftec3	Minha postura típica é de confiar em novas tecnologias até que elas me provem que eu não devo confiar nelas	
familiar	Qual das opções melhor descreveria a sua familiaridade com a inteligência artificial?	1 – Nada familiar 2 – Pouco familiar 3 – Moderadamente familiar 4 – Muito familiar 5 – Extremamente familiar
habil	Qual o seu grau de habilidade com ferramentas de inteligência artificial?	1 – Nenhum 2 – Básico 3 – Principiante 4 – Intermediário 5 – Avançado 6 – Expert
idade	Qual a sua idade?	–
genero	Qual o seu gênero?	1 – Masculino 2 – Feminino 3 – Não desejo informar
grauesc	Qual o último nível (concluído) da sua formação escolar/acadêmica?	1 – Ensino Médio 2 – Graduação 3 – Especialização 4 – Mestrado 5 – Doutorado
areaform	Qual a sua área de formação escolar/acadêmica?	1 – Ciências Agrárias 2 – Ciências Biológicas 3 – Ciências da Saúde 4 – Ciências Exatas 5 – Ciências Humanas 6 – Ciências Sociais Aplicadas 7 – Engenharias 8 – Linguística, Letras e Artes 9 – Outras 10 – Não se aplica
cargo	Qual o seu cargo atual?	1 – Gerente de Administração 2 – Gerente de Negócios 3 – Gerente de Núcleo 4 – Superintendente 5 – Gerente de Cobrança e Reestruturação de Ativos 6 – Gerente Geral de Unidade 7 – Assessor 8 – Consultor
tcargo	Há quantos anos (completos) você está no seu cargo atual?	–
tgerente	Há quantos anos (completos) você exerce cargos de natureza gerencial nesta organização?	–
tcasa	Há quantos anos (completos) você trabalha nesta organização?	–

APÊNDICE F – LEGENDA DOS COMPONENTES / VARIÁVEIS DE ANÁLISE

CÓDIGO	COMPONENTE / VARIÁVEL
pconcord	Probabilidade de concordância com a recomendação da IA
valor	Valor da decisão tomada pelo respondente
resptrans	Resposta à manipulação da transparência
respdec	Resposta à manipulação do tipo de decisão
conftec1	Confiança na tecnologia
conftec2	
conftec3	
familiar	Familiaridade com a IA
habil	Habilidade no manuseio de IA
idade	Idade
genero	Gênero
grauesc*	Grau de escolaridade
areaform*	Área de formação acadêmica
cargo	Cargo
tcargo **	Tempo no cargo atual
tgerente **	Tempo em cargos de gestão
tcasa **	Tempo de empresa

***** representam a variável indicada na *Libby Box* como "Formação"
****** representam a variável indicada na *Libby Box* como "Experiência".