

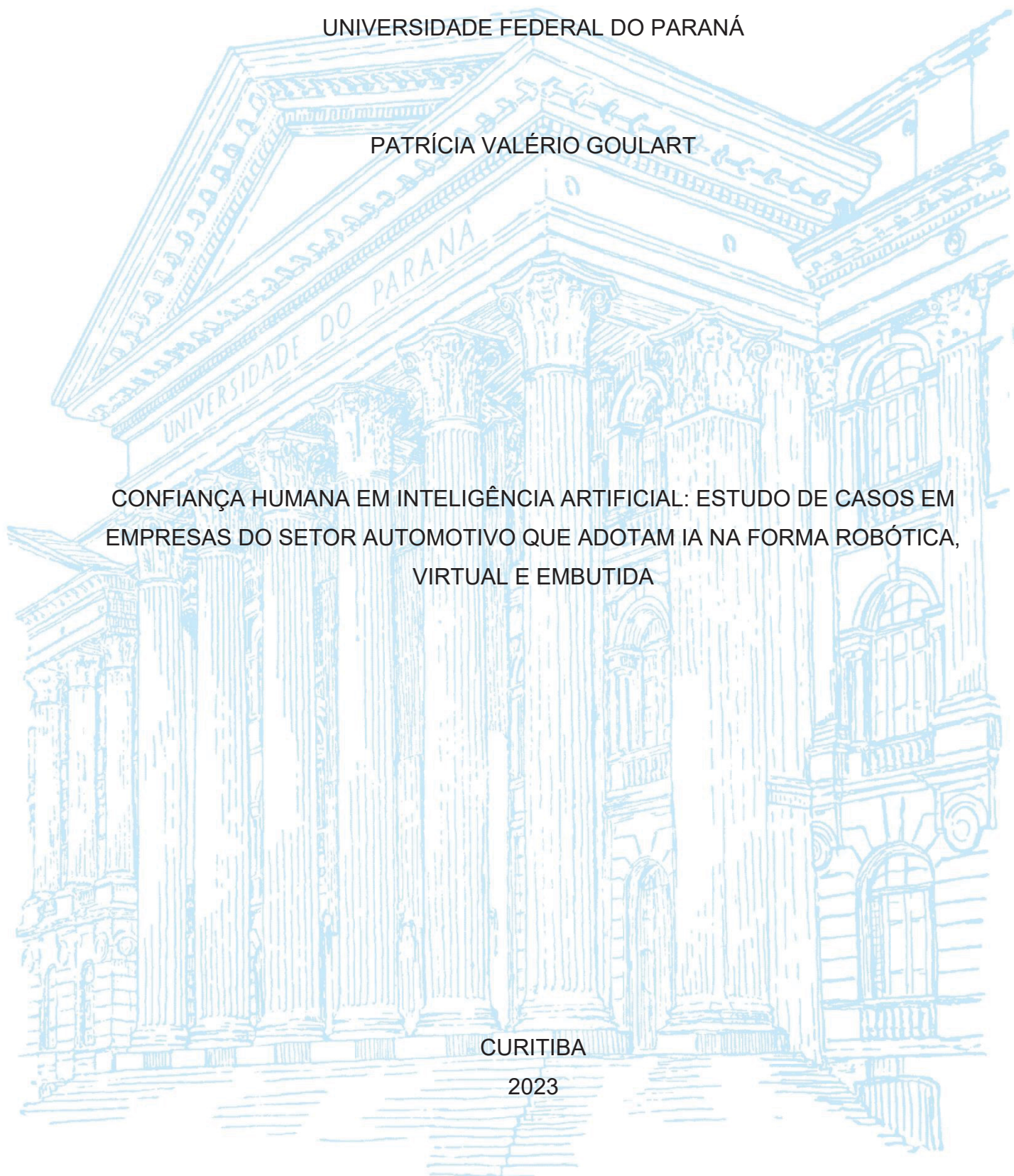
UNIVERSIDADE FEDERAL DO PARANÁ

PATRÍCIA VALÉRIO GOULART

CONFIANÇA HUMANA EM INTELIGÊNCIA ARTIFICIAL: ESTUDO DE CASOS EM
EMPRESAS DO SETOR AUTOMOTIVO QUE ADOTAM IA NA FORMA ROBÓTICA,
VIRTUAL E EMBUTIDA

CURITIBA

2023



PATRÍCIA VALÉRIO GOULART

CONFIANÇA HUMANA EM INTELIGÊNCIA ARTIFICIAL: ESTUDO DE CASOS EM
EMPRESAS DO SETOR AUTOMOTIVO QUE ADOTAM IA NA FORMA ROBÓTICA,
VIRTUAL E EMBUTIDA

Dissertação apresentada ao curso de Pós-Graduação em Gestão de Organizações, Liderança e Decisão, Setor de Ciências Sociais Aplicadas, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Mestre em Gestão de Organizações, Liderança e Decisão.

Orientador: Prof. Dr. Farley Simon Mendes Nobre

CURITIBA

2023

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
UNIVERSIDADE FEDERAL DO PARANÁ
SISTEMA DE BIBLIOTECAS – BIBLIOTECA DE CIÊNCIAS SOCIAIS APLICADAS

Goulart, Patrícia Valério

Confiança humana em inteligência artificial : estudo de casos em empresas do setor automotivo que adotam IA na forma robótica, virtual e embutida / Patrícia Valério Goulart. – Curitiba, 2023.

1 recurso on-line : PDF.

Dissertação (Mestrado) – Universidade Federal do Paraná, Setor de Ciências Sociais Aplicadas, Programa de Pós-Graduação em Gestão de Organizações, Liderança e Decisão. Orientador: Prof. Dr. Farley Simon Mendes Nobre.

1. Inteligência artificial. 2. Indústria automobilística.
3. Processo decisório. I. Nobre, Farley Simon Mendes.
II. Universidade Federal do Paraná. Programa de Pós-Graduação em Gestão de Organizações, Liderança e Decisão. III. Título.

Bibliotecária: Maria Lidiane Herculano Graciosa CRB-9/2008



MINISTÉRIO DA EDUCAÇÃO
SETOR DE CIÊNCIAS SOCIAIS E APLICADAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO GESTÃO DE
ORGANIZAÇÕES, LIDERANÇA E DECISÃO - 40001016173P9

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação GESTÃO DE ORGANIZAÇÕES, LIDERANÇA E DECISÃO da Universidade Federal do Paraná foram convocados para realizar a arguição da dissertação de Mestrado de **PATRICIA VALERIO GOULART** intitulada: **Confiança humana em inteligência artificial: Estudo de casos em empresas do setor automotivo que adotam IA na forma robótica, virtual e embutida, sob orientação do Prof. Dr. FARLEY SIMON MENDES NOBRE**, que após terem inquirido a aluna e realizada a avaliação do trabalho, são de parecer pela sua **APROVAÇÃO** no rito de defesa.

A outorga do título de mestra está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 17 de Abril de 2023.

Assinatura Eletrônica

03/05/2023 15:52:36.0

FARLEY SIMON MENDES NOBRE

Presidente da Banca Examinadora

Assinatura Eletrônica

03/05/2023 08:15:25.0

GUILHERME FRANCISCO FREDERICO

Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica

03/05/2023 09:18:40.0

EGON WALTER WILDAUER

Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Dedico este trabalho ao grande Arquiteto do Universo, o Leão de Judá, Ábba Pai. Àquele que deu seu filho único, Jesus, para que todo aquele que nele crê não pereça, mas tenha vida eterna.

AGRADECIMENTOS

À Deus, por me sustentar até quando não vejo.

À minha família, Rose, Joaquim, Tiago, Marly, Lara e Samuel, por todo suporte e amor durante esta jornada.

Ao meu amor, Rodrigo Lenz, pela compreensão nos períodos de ausência e por todo carinho e dedicação.

Às minhas melhores amigas Aline Carla e Berenice Chaves que durante este processo foram essenciais, agradeço pelas mãos amigas que me ajudaram a levantar da crise gerada pelos problemas de saúde que enfrentei durante todo o processo.

Ao professor e orientador Dr. Farley Simon Mendes Nobre, com quem tenho a honra de compartilhar esta caminhada, agradeço imensamente pela orientação e dedicação, por ser minha referência e inspiração. Que sorte a minha!

Aos professores, amigos e companheiros do curso Pós-Graduação em Gestão de Organizações, Liderança e Decisão Mestrado Profissional – PPGOLDD, em especial ao Roque, Alexandre, Rodolfo, Andressa, Wagner e Anderson por toda ajuda.

Aos meus gestores e parceiros na T-Systems do Brasil, especialmente ao Rodrigo Alves, Marcelo Torres, Marcio Miranda e Angelica Vitalli, que contribuíram para viabilizar esta jornada.

Aos meus mestres que me acompanham desde a universidade (UEMG), Gualberto Rabay, Alessandro Borges e Anderson Matos por acreditarem no meu potencial e por todo incentivo.

Aos meus mentores, Thiago Ayres, Armando Morais, Gustavo Prieto, Alessandro Andreotti e Thomas Walenta, pelo investimento da variável mais cara (o tempo), compartilhando comigo rico conhecimento.

Me surpreendi ao me deparar com um grande desafio. E ele se surpreendeu com a minha coragem de prosseguir, mesmo quando tive medo (Elaborado pela autora).

RESUMO

Contexto: A adoção de Inteligência Artificial (IA) se tornou cada vez mais presente nas organizações. Contudo, nem todas elas estão preparadas para as transformações causadas pelo uso dessa tecnologia que, de certa forma, pode executar funções de um ser humano. Publicações teóricas recentes apontam que um dos grandes desafios nesse campo se refere à compreensão de como a adoção de IA nas organizações afeta a confiança das pessoas que a utilizam e abrem uma lacuna de pesquisa no que se refere aos estudos empíricos sobre esse tema. **Objetivo:** Por conseguinte, esta dissertação tem como objetivo geral ***compreender empiricamente os fatores (cognitivos e emocionais) que afetam a confiança humana em IA nas organizações.*** **Métodos/resultados:** A metodologia de pesquisa se caracteriza como exploratória e qualitativa e envolve duas etapas distintas. Por se tratar de um tema emergente, primeiramente, o presente estudo buscou **(i) identificar na literatura quais são os fatores que afetam a confiança humana em IA nas organizações.** Para responder a esta pergunta, conduziu-se Revisão Integrativa (RI) para identificar e definir os diferentes fatores considerados pelas organizações no que tange a confiança em IA. Resultados e análises da RI permitiram a identificação de fatores que afetam a adoção de IA nas organizações, a exemplo de similaridade com comportamentos humanos, presença física, complexidade da tarefa, além de relacionar estes fatores aos diferentes tipos de IA (Robô, Agentes Virtuais e Bot/Embutida). Em seguida, conduziu-se estudo de casos múltiplos para **(ii) compreender quais são os fatores que influenciam a confiança humana em IA em empresas do setor automotivo e como estes fatores se relacionam aos diferentes tipos de IA (Robô, Agentes Virtuais e Bot/Embutida).** O setor automotivo foi selecionado por se apresentar como pioneiro e influenciador na adoção de IA, tanto em seus processos produtivos quanto nos seus produtos. Nesta segunda etapa, buscou-se compreender em que grau os fatores identificados na RI foram verificados no estudo de casos. **Contribuições:** A partir da RI, foram elaborados dois quadros. O primeiro quadro apresenta uma síntese da evolução dos conceitos de confiança humana, IA e fatores cognitivos e emocionais. O segundo quadro integra os principais fatores cognitivos e emocionais identificados na literatura e que influenciam confiança humana em IA nas organizações. A partir do estudo de casos, elaborou-se um terceiro quadro que contextualiza vantagens e desafios na adoção de IA em

organizações do setor automotivo. Esta dissertação contribui para o avanço de pesquisas sobre confiança humana em IA abordando uma importante questão social e tecnológica sob perspectivas teórica e empírica. Seus resultados auxiliam gestores com informações sobre os fatores que afetam a adoção de IA nas organizações e abre novas frentes de pesquisa futura sobre a adoção de IA em outros setores industriais. O presente estudo se limita ao setor automotivo de um país emergente e pode também ser estendido em pesquisas comparativas que envolvem países de regiões globais diversas.

Palavras-chave: Confiança humana. Inteligência artificial. Fatores cognitivos e emocionais. Estudo de casos. Setor automotivo. Decisão.

ABSTRACT

Context: The adoption of Artificial Intelligence (AI) has become increasingly present in organizations. However, not all of them are prepared for the transformations caused by the use of this technology that, in a way, can perform functions of a human being. Recent theoretical publications point out that one of the great challenges in this field refers to understanding how the adoption of AI in organizations affects the trust of the people who use it and open a research gap regarding empirical studies on this topic.

Objective: Therefore, **this dissertation aims to empirically understand the factors (cognitive and emotional) that affect human trust in AI in organizations.**

Methods/results: The research methodology is characterized as exploratory and qualitative and involves two distinct stages. As this is an emerging topic, first, this study sought to **(i) identify in the literature which factors affect human trust in AI in organizations.** To answer this question, an Integrative Review (IR) was conducted to identify and define the different factors considered by organizations regarding trust in AI. IR results and analyzes allowed the identification of factors that affect the adoption of AI in organizations, such as similarity with human behavior, physical presence, task complexity, in addition to relating these factors to the different types of AI (Robot, Virtual Agents and Bot/Embedded). Then, a multiple case study was conducted to **(ii) understand what are the factors that influence human trust in AI in companies in the automotive sector and how these factors are related to different types of AI (Robot, Virtual Agents and Bot/ built-in).** The automotive sector was selected for presenting itself as a pioneer and influencer in the adoption of AI, both in its production processes and in its products. In this second stage, we sought to understand to what degree the factors identified in the IR were verified in the case studies. **Contributions:** Based on the IR, two tables were created. The first table presents a synthesis of the evolution of the concepts of human trust, AI and cognitive and emotional factors. The second table integrates the main cognitive and emotional factors identified in the literature and that influence human trust in AI in organizations. Based on the case studies, a third table was created that contextualizes the advantages and challenges of adopting AI in organizations in the automotive sector. This dissertation contributes to the advancement of research on human trust in AI by addressing an important social and technological issue from theoretical and empirical perspectives. Its results help managers with information about the factors that affect the adoption of AI in

organizations and open new fronts for future research on the adoption of AI in other industrial sectors. The present study is limited to the automotive sector of an emerging country and can also be extended in comparative research involving countries from different global regions.

Keywords: Human trust. Artificial intelligence. Cognitive and emotional factors. Case Study. Automotive sector. Decision.

LISTA DE FIGURAS

FIGURA 1 - NÍVEL DE MATURIDADE DAS INDÚSTRIAS EM RELAÇÃO A IA.....	23
FIGURA 2 - PENETRAÇÃO DE IA NO SETOR AUTOMOTIVO NOS PRÓXIMOS 10 ANOS	24
FIGURA 3 - ESTRUTURA DA DISSERTAÇÃO	25
FIGURA 4 - MODELO DE CONFIANÇA EM IA	30
FIGURA 5 - DESENHO DE PESQUISA.....	34
FIGURA 6 - CLASSIFICAÇÃO DA PESQUISA.....	37
FIGURA 7 - FLUXO PRISMA	46
FIGURA 8 - DISTRIBUIÇÃO DE ARTIGOS POR DIMENSÃO	16
FIGURA 9 - FATORES PARA ADOÇÃO DE IA	20
FIGURA 10 - DESAFIOS NA ADOÇÃO DE IA.....	22
FIGURA 11 - AREAS MAIS AFETADAS POR IA.....	23
FIGURA 12 - BENEFÍCIOS DO USO E IA	24
FIGURA 13 - DESAFIOS DURANTE A IMPLEMENTAÇÃO DE IA.....	25
FIGURA 14 - RESULTADOS ESPERADOS COM ADOÇÃO DE IA.....	25
FIGURA 15 - PRINCIPAIS DESAFIOS DURANTE A IMPLEMENTAÇÃO DE IA	27
FIGURA 16 - AÇÕES PARA GERAR CONFIABILIDADE	30
FIGURA 17 - CARACTERÍSTICAS DA TAREFA MAIS COMUNS	30

LISTA DE QUADROS

QUADRO 1 - ETAPAS DA PESQUISA.....	40
QUADRO 2 - LISTA DE ENTREVISTADOS	42
QUADRO 3 - SÍNTESE DOS CONCEITOS E CONSTRUTOS IDENTIFICADOS NA RI	16
QUADRO 4 - EVOLUÇÃO DOS CONCEITOS.....	16
QUADRO 5 - SÍNTESE DOS FATORES COGNITIVOS E EMOCIONAIS IDENTIFICADOS NA RI.....	16
QUADRO 6 - MODELO DE GLIKSON E WOOLLEY	16
QUADRO 7 - INFORMAÇÕES SOBRE AS EMPRESAS ESTUDADAS	19
QUADRO 8 - TIPOS DE IA POR EMPRESA	32
QUADRO 9 - SÍNTESE DOS FATORES IDENTIFICADOS NO ESTUDO DE CASOS	32
QUADRO 10 – BENEFÍCIOS NA IMPLEMENTAÇÃO DE IA.....	36
QUADRO 11 - DEDAFIOS NA IMPLEMENTAÇÃO DE IA.....	37

LISTA DE TABELAS

TABELA 1 - QUANTIDADE DE PUBLICAÇÕES POR ANO	46
--	----

LISTA DE ABREVIATURAS OU SIGLAS

IA	Inteligência Artificial
RI	Revisão Integrativa

SUMÁRIO

1 INTRODUÇÃO	16
1.1 PROBLEMA DE PESQUISA	19
1.2 DEFINIÇÕES DOS OBJETIVOS DE PESQUISA	20
1.3 JUSTIFICATIVA TEÓRICA E PRÁTICA.....	21
1.4 ESTRUTURA DA DISSERTAÇÃO	25
2 FUNDAMENTAÇÃO TEÓRICO-EMPIRICA	26
2.1 ORGANIZAÇÕES	26
2.2 CONFIANÇA HUMANA.....	27
2.3 INTELIGÊNCIA ARTIFICIAL	28
2.4 FATORES COGNITIVOS E EMOCIONAIS.....	29
3 METODOLOGIA	32
3.1 ESPECIFICAÇÃO DO PROBLEMA	32
3.2 DEFINIÇÃO DOS TERMOS.....	34
3.3 CLASSIFICAÇÃO DA PESQUISA	36
3.4 DELINEAMENTO DA PESQUISA.....	38
3.5 ESCOLHA DOS CASOS	40
3.6 TÉCNICA DE COLETA E ANÁLISE DOS CASOS.....	41
3.7 VALIDADE E CONFIABILIDADE	43
4 REVISÃO INTEGRATIVA	45
4.1 TABULAÇÃO DOS DADOS	47
4.2 RESULTADOS E ANÁLISES DA RI.....	47
5 ESTUDO DE CASOS MÚLTIPLOS: ALFA, BETA E GAMA	18
5.1 RESULTADOS	19
5.2 ANÁLISES.....	31
5.3 DISCUSSÕES E PROPOSIÇÕES	38
6 CONSIDERAÇÕES FINAIS	41
6.1 CONTRIBUIÇÕES	42
6.2 IMPLICAÇÕES PARA GESTÃO E DECISÃO.....	42
6.3 RECOMENDAÇÕES PARA ESTUDOS FUTUROS.....	43
6.4 LIMITAÇÕES DA PESQUISA	44
REFERÊNCIAS	45

1 INTRODUÇÃO

Novas tecnologias associadas à Inteligência Artificial (IA) passaram a ser mais conhecidas, difundidas e aplicadas principalmente com a crescente utilização de mecanismos digitais pelas organizações e seus participantes a partir de transformações sociotécnicas que favoreceram a Indústria 4.0 (GHOBAKHLOO et al., 2021; NOBRE 2005, 2011; NOBRE et al., 2008; WANG et al., 2022; XU et al., 2021).

Com o avanço da Indústria 4.0 - que envolve transformações tecnológicas e organizacionais e busca alcançar maior eficiência e produtividade como uma estratégia de competitividade (NOBRE, 2011; NOBRE ET AL., 2008; XU et al., 2021) - as organizações passaram a ser cada vez mais desafiadas pela velocidade com que as mudanças ocorrem e requerem sua adaptação em ambientes complexos, dinâmicos e incertos. Raisch e Krakowski (2021, p.3) proferem que “avanços computacionais recentes e aumento exponencial de dados e novas técnicas de aprendizado de máquina, agora favorecem que as organizações usem soluções baseadas em IA” e outras tecnologias relacionadas.

À medida que empresas se aproximaram de modelos de Indústria 4.0, alguns estudos teóricos e empíricos puderam constatar as suas vantagens relacionadas à produtividade e deficiências associadas à sua falta de integração aos aspectos humanos e ergonômicos no trabalho além de questões mais amplas de cunho social e ecológico (BEIER, et al., 2020; BEIER; NIEHOFF; HOFFMANN, 2021; BEIER; KIEFER; KNOFF, 2022).

Por conseguinte, alguns autores anteciparam conceitos voltados para uma Indústria 5.0 que procura suprir essas deficiências à medida que procura reduzir disparidades e incompatibilidades entre o ser humano e a máquina a fim de utilizar essa combinação para melhoria de criatividade e aprendizado organizacional (NAHAVANDI, 2019; WANG et al., 2022; XU et al., 2021). Os proponentes da Indústria 5.0 reconhecem que os sistemas industriais podem contribuir para o alcance de objetivos sociais e ecológicos que vão além de geração de empregos e bem-estar de seus participantes. Ela busca se tornar um provedor de resiliência e prosperidade à medida que reconhece os limites do nosso planeta (XU et al., 2021). Neste sentido, enquanto a Indústria 4.0 tem foco na tecnologia e automação, a Indústria 5.0 tem foco no processo de *augmentation*, ou seja, na cooperação e sinergia entre homem-

máquina (Raisch e Krakowski, 2021), e em valores entregues nos ambientes interno e externo da organização (XU et al., 2021).

Referindo-se à IA, uma “nova geração de tecnologias capazes de interagir com o meio ambiente e que simula a inteligência humana” Glikson e Woolley (2020, p.628) passa a introduzir novos desafios no que se refere à sua integração nas organizações. **O desafio que interessa o estudo nesta dissertação é relativo à confiança das pessoas que são usuários de tecnologias de IA nas organizações.** Com a evolução das organizações industriais e das tecnologias de informação, as máquinas passaram a realizar funções mecânicas e cognitivas mais complexas nos ambientes produtivos. Dessa maneira, elevaram-se suas participações antes restritas ao chão de fábrica para níveis que envolvem dados gerenciais e organizacionais mais complexos (Nobre *et al.*, 2008). Por conseguinte, alguns estudos propuseram que máquina de IA, assim como seres humanos, podem ser classificadas como participantes nas organizações à medida que elas executam decisões em nome das organizações (NOBRE, 2005; NOBRE; TOBIAS; WALKER, 2009). Ao promover uma evolução na relação homem-máquina, espera-se potencializar capacidades humanas e organizacionais por meio de cooperação ao invés de automação (RAISCH; KRAKOWSKI, 2021).

Haverá uma interrupção massiva. A tecnologia está se desenvolvendo muito rapidamente, possivelmente de forma exponencial. Mas as pessoas são lineares. Quando as pessoas lineares se deparam com uma mudança exponencial, elas não serão capazes de se adaptar a isso facilmente. Então, claramente, algo está vindo ... E claramente a IA vai vencer [contra a inteligência humana]. Não está nem perto. Como as pessoas vão se ajustar a isso é um problema fascinante - mas para meus filhos e netos, não para mim, declara Daniel Kahneman em entrevista ao The Guardian (KAHNEMAN, 2021).

Tecnologias como IA podem ser percebidas pelas pessoas como um certo risco uma vez que estas tecnologias podem assumir funções similares às que as desempenham nas organizações. Porém, estas tecnologias tendem a criar ainda mais oportunidades mudando a forma e o tipo de tarefas que são executadas pelas pessoas ao invés de tirar seus empregos. (GLIKSON; WOOLLEY, 2020; CURIONI, 2018).

Conforme Asan, Bayrak e Choudury (2020), as interações dos seres humanos com IA são altamente influenciadas pela confiança. Promover confiança humana na utilização de IA é um fator chave para o sucesso da implantação desta tecnologia, principalmente porque os humanos temem serem substituídos pelas máquinas e este

temor gera resistência criando barreiras no avanço desta tecnologia nas organizações (CURIONI, 2018).

Nobre et al. (2008, 2009) explicam que máquinas cognitivas são necessárias para estender os limites humanos de capacidade computacional e gerenciamento de incertezas para modelos mais avançados de cognição e processamento de informações. Ao participar das organizações, as máquinas cognitivas são agentes da cognição organizacional e contribuem para melhorar os graus de cognição, inteligência e autonomia da organização (NOBRE, 2005). Humanos e máquinas são muitas vezes tidos como adversários em função dos efeitos da automação sobre o mercado de trabalho. Porém, suas fortalezas combinadas podem elevar o potencial de decisão e aprendizagem, aumentando-se suas capacidades e não o contrário (CURIONI, 2018; GLIKSON E WOOLLEY, 2020; NOBRE et al., 2009). Glikson e Woolley (2020) destacam que “focar em questões de confiança nos permite abordar não apenas o desuso (uma rejeição) da tecnologia, mas também seu uso indevido (um excesso de confiança inadequada na tecnologia) ou seu abuso (uso prejudicial para obter um ganho individual”.

Confiança possui aspectos cognitivos (baseado em percepções) e emocionais (baseado em relacionamentos e conexões) além de ter diferentes níveis na relação do usuário com a tecnologia (baixa confiança, que pode levar ao desuso, perda de tempo e desperdício de recursos; alta confiança, que poderia levar ao uso indevido ou ocasionar violação de segurança). Para tanto, focar em questões de como confiança pode ser influenciada, permite abordar o uso ou rejeição da tecnologia IA e por isto compreender em detalhes estes fatores é crítico para que as organizações possam potencializar o uso da IA. Neste contexto, esta dissertação tem como objetivo: **Compreender os fatores (cognitivos e emocionais) que influenciam a confiança humana em IA nas organizações.** Para responder à esta questão este estudo utilizou 2 estratégias em sua fundamentação.

Na primeira etapa, realizou-se uma **revisão integrativa da literatura** com objetivo de identificar quais são os fatores que afetam a confiança humana em IA nas organizações. Elaborou-se um quadro de análise teórico contendo conceitos de IA, confiança humana em IA e fatores cognitivos e emocionais que afetam a confiança humana em IA nas organizações. A revisão integrativa possibilitou identificar e estruturar estes fatores em um modelo que suportou o desenvolvimento da segunda etapa.

A segunda etapa constituiu-se de um **estudo de casos múltiplos** em empresas do setor automotivo em que a coleta de dados foi realizada através de técnicas de entrevista semiestruturadas. Neste estudo adotou-se o modelo teórico proposto em Glikson e Woolley (2020) para realização de estudo teórico e empírico sobre fatores (cognitivos e emocionais) que afetam a confiança humana em IA nas organizações. Optou-se por este modelo uma vez que abrange as principais formas de representação de IA (Robô Físico, Agente Virtual e Embutida) que influenciam elementos chave que moldam os fatores cognitivos e emocionais de confiança humana em IA que podem ser classificados nas seguintes dimensões: tangibilidade, transparência, confiabilidade, características das tarefas e comportamentos imediatos. Adicionalmente, o modelo de Glikson e Woolley (2020) mostrou abranger grande parte dos tipos de IA e de fatores cognitivos e emocionais encontrados na Revisão Integrativa (RI) conduzida nesta dissertação. Como contribuição, esta dissertação aborda uma importante questão social e tecnológica sob perspectivas teórica e empírica sobre confiança humana em IA. Seus resultados auxiliam gestores com informações sobre os fatores e dimensões que afetam a adoção de IA nas organizações.

1.1 PROBLEMA DE PESQUISA

O dinamismo e velocidade com que os ambientes de fabricação têm que se adaptar à transformação digital traz complexidades e incertezas, principalmente em relação ao volume exponencial de dados que são gerados. Para lidar com estes desafios, o uso de tecnologias como IA mostram alto potencial para suportar os fabricantes através de análises preditivas e mecanismos de tomada de decisão. Porém, ainda é relativamente baixa a adoção de tecnologias como IA na indústria devido à falta de preparação destas organizações, principalmente porque no ambiente real existem situações singulares e complexas (PERES et al., 2020).

À medida que o mercado gera possibilidades para que sistemas industriais produzam de forma cada vez mais personalizada, as empresas precisam desenvolver novas competências tecnológicas, gerenciais e organizacionais para acompanhar ambientes cada vez mais incertos e complexos (NOBRE et al, 2008; XU et al., 2021).

Conforme Lapinska et al. (2021), existem riscos e benefícios relacionados a implementação de IA. Benefícios como, aumento de eficiência e redução de custo com mão de obra, pode se tornar ao mesmo tempo um risco se notado pelos empregados como uma ameaça por medo de ficar sem emprego. Por isto, para reduzir essas preocupações as organizações devem promover confiança em IA durante a implementação. Nesse contexto, os principais fatores que determinam o nível de aceitação do uso de IA pelos empregados está diretamente relacionado ao grau de confiança em IA e, para tanto, conhecer quais os fatores (cognitivos e emocionais) que influenciam confiança em IA é determinante para aceitação, implementação e utilização de soluções com IA nas empresas.

Dessa forma, formulou-se a seguinte pergunta de pesquisa:

Quais são os fatores que afetam a confiança (cognitiva e emocional) humana em IA em empresas do setor automotivo?

1.2 DEFINIÇÕES DOS OBJETIVOS DE PESQUISA

A seguir, apresentam-se objetivo geral e objetivos específicos que norteiam esta dissertação.

1.2.1 Objetivo Geral

Compreender os fatores (cognitivos e emocionais) que afetam a confiança humana em IA nas organizações.

1.2.2 Objetivos Específicos

Por meio de Revisão Integrativa (RI), buscou-se i) identificar na literatura quais são os fatores que afetam a confiança humana em IA nas organizações. Elaborou-se um quadro de análise teórico contendo conceitos de IA, confiança humana em IA e fatores cognitivos e emocionais que afetam a confiança humana em IA nas organizações. Em seguida, deduziu-se proposições relacionando os fatores aos diferentes tipos de IA (Robô, Agentes Virtuais e Bot/Embutida).

- Por meio de estudo de casos, buscou-se **(ii) compreender quais são os fatores que influenciam a confiança humana em IA em empresas do setor automotivo e como estes fatores se relacionam aos diferentes tipos de IA (Robô, Agentes Virtuais e Bot/Embutida)**. Conduziu-se entrevistas com participantes de empresas do setor automotivo para compreender em que grau o quadro de análise teórico e as proposições deduzidas da RI podem ser verificadas empiricamente.
- Por final, por meio de triangulação dos resultados da RI e do estudo de casos, elaborou-se um quadro de análise teórico-empírico contendo benefícios e desafios encontrados na adoção de IA em organizações do setor automotivo.

1.3 JUSTIFICATIVA TEÓRICA E PRÁTICA

1.3.1 Teórica

Através de uma revisão de literatura é possível analisar e sintetizar o conhecimento científico já produzido sobre um determinado tema SNYDER (2019). Com isto, buscou-se identificar e analisar na literatura como está a evolução dos conceitos sobre construção de confiança humana em IA nas organizações a fim de compreender o estado da arte do tema.

A estratégia empregada na busca de publicações considerou aquelas indexadas nas bases de dados *Web of Science* e *Scopus* em virtude da relevância e multidisciplinaridade, além de amplitude, por serem bases internacionalmente utilizadas. Estabeleceu-se como critérios de busca os termos “confiança humana” “OU” “pessoas confiam” “E” “inteligência artificial”, em inglês, cujos foram empregados na busca de títulos e resumos. A utilização do “E” se deu com objetivo de obter os dois termos juntos no retorno das buscas. Em relação ao tempo, optou-se por não delimitar uma série temporal a fim de identificar tudo que já foi publicado até o momento, independente do ano de publicação. A busca considerou ainda a delimitação das publicações do tipo “artigo” e idioma “inglês”. A pesquisa foi realizada em fevereiro do ano de 2022 e resultou em 25 artigos encontrados, já excluídos os repetidos.

Glikson; Woolley (2020) apontam uma necessidade crescente de pesquisas em organizações que já usam IA em seus sistemas de gestão ou tomada de decisão

que explorem o ambiente da vida real. Łapińska et al. (2021) consideram que “uma pesquisa entre funcionários com competências mais diversas na área de usuários de soluções de IA em seus locais de trabalho pode trazer uma nova luz sobre as questões examinadas”. Por ser um tema emergente, a justificativa teórica deste estudo se dá pela atualidade do tema proposto e pela lacuna em relação a pesquisas relacionadas a confiança humana em inteligência artificial. Tal lacuna pode ser constatada pelo baixo volume de trabalhos encontrados sobre este tema na busca realizada para realização deste estudo.

Por meio da RI, a proposta deste estudo é também buscar enriquecer a literatura disponível já que “para tópicos emergentes, o objetivo é antes criar conceituações iniciais ou preliminares e modelos teóricos, em vez de revisar modelos antigos” conforme destacado por Snyder (2019). E ajudar as organizações a construir confiança em IA e potencializar o uso desta tecnologia.

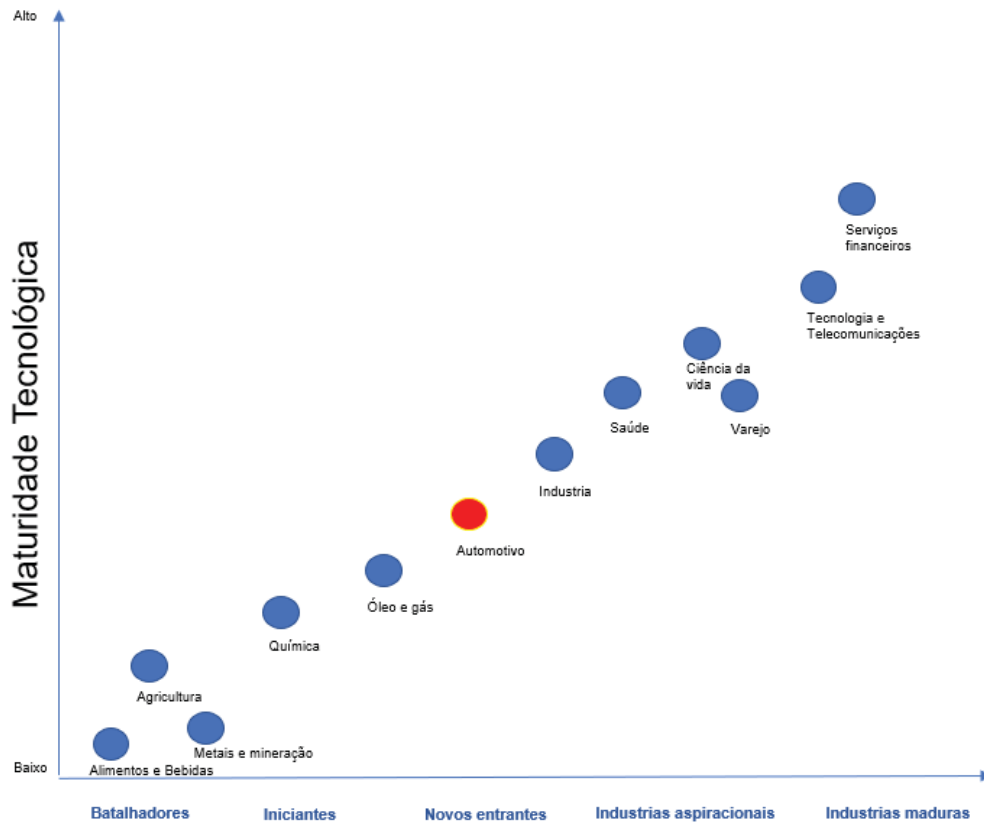
1.3.2 Prática

O setor automotivo é dos mais promissores em termos de crescimento e investimento em relação ao uso de IA tanto no processo produtivo como nos produtos. A realizada pelo CAPGEMINI INSTITUTE RESEARCH aponta que “a proporção de organizações automotivas que implantam IA em escala cresceu 10% em 2019 em comparação com 7% em 2017”, porém o progresso é ainda considerado modesto em relação a outros setores.

Atualmente, IA está sendo implementada no setor automotivo, incluindo design, cadeia de suprimentos, produção e pós-produção. Além disso, IA está sendo implementada em sistemas de 'assistência ao motorista' e 'avaliação de risco do motorista', o que está transformando o setor de transporte. Os serviços pós-venda, como manutenção preditiva e seguro, também estão se transformando com o uso de IA, FUTURE BRIDGE (2020).

Na FIGURA 1 é possível verificar a posição da indústria automotiva no rank de indústrias pesquisados pela FUTURE BRIDGE em 2020 cujos dados apontam que indústria automotiva embora “novata” em relação à adoção de IA, é um setor promissor.

FIGURA 1 - NÍVEL DE MATURIDADE DAS INDÚSTRIAS EM RELAÇÃO A IA



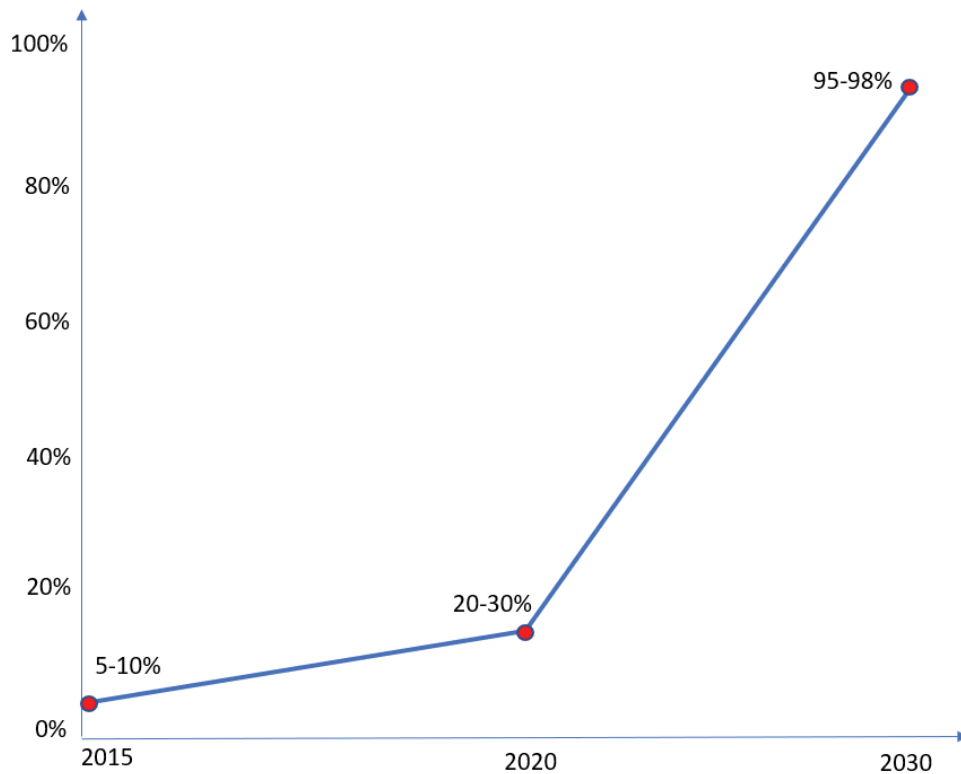
FONTE: Adaptado de FUTURE BRIDGE (2020).

Embora ainda com baixo nível de maturidade, há muita expectativa em relação a IA no setor automotivo para os próximos 10 anos, conforme apontado pela FUTURE BRIDGE (2020):

No setor automotivo, a implementação da inteligência artificial está em um ritmo muito lento. Em 2015, 5–10% dos veículos tinham alguma ou outra forma de sistema de IA instalado. Nos últimos cinco anos, a IA está sendo implantada para assistência a estacionamento, controle de cruzeiro e veículos autônomos de nível 1 e 2, o que resultou em maior adoção de tecnologia. Nos próximos 8 a 10 anos, espera-se que a tecnologia de IA faça parte de muitas aplicações automotivas. Portanto, até o final de 2030, pode-se esperar que 95–98% dos novos veículos tenham tecnologia de IA FUTURE BRIDGE (2020).

Na FIGURA 2 abaixo é possível verificar a projeção da penetração de IA no setor automotivo de 2019 a 2030.

FIGURA 2 - PENETRAÇÃO DE IA NO SETOR AUTOMOTIVO NOS PRÓXIMOS 10 ANOS



FONTE: Adaptado de FUTURE BRIDGE (2020).

Glikson e Woolley (2020, P.640) sugerem que “pesquisas futuras devem considerar o papel de confiança na IA dentro das organizações para entender melhor as dificuldades específicas que precisam ser abordadas para facilitar a sua utilização. Neste contexto, justifica-se estudar os fatores cognitivos e emocionais que afetam a confiança humana em IA em empresas do setor automotivo para suportar as pessoas e as organização nesta transição. Toda mudança requer preparação e passa por um período de adaptação, e neste sentido Glikson e Woolley (2020, P.633) destacam que “pesquisas existentes abordando a trajetória do ser humano e confiança em IA sugerem que a confiança inicial começa em um nível baixo e se desenvolve ao longo do tempo”.

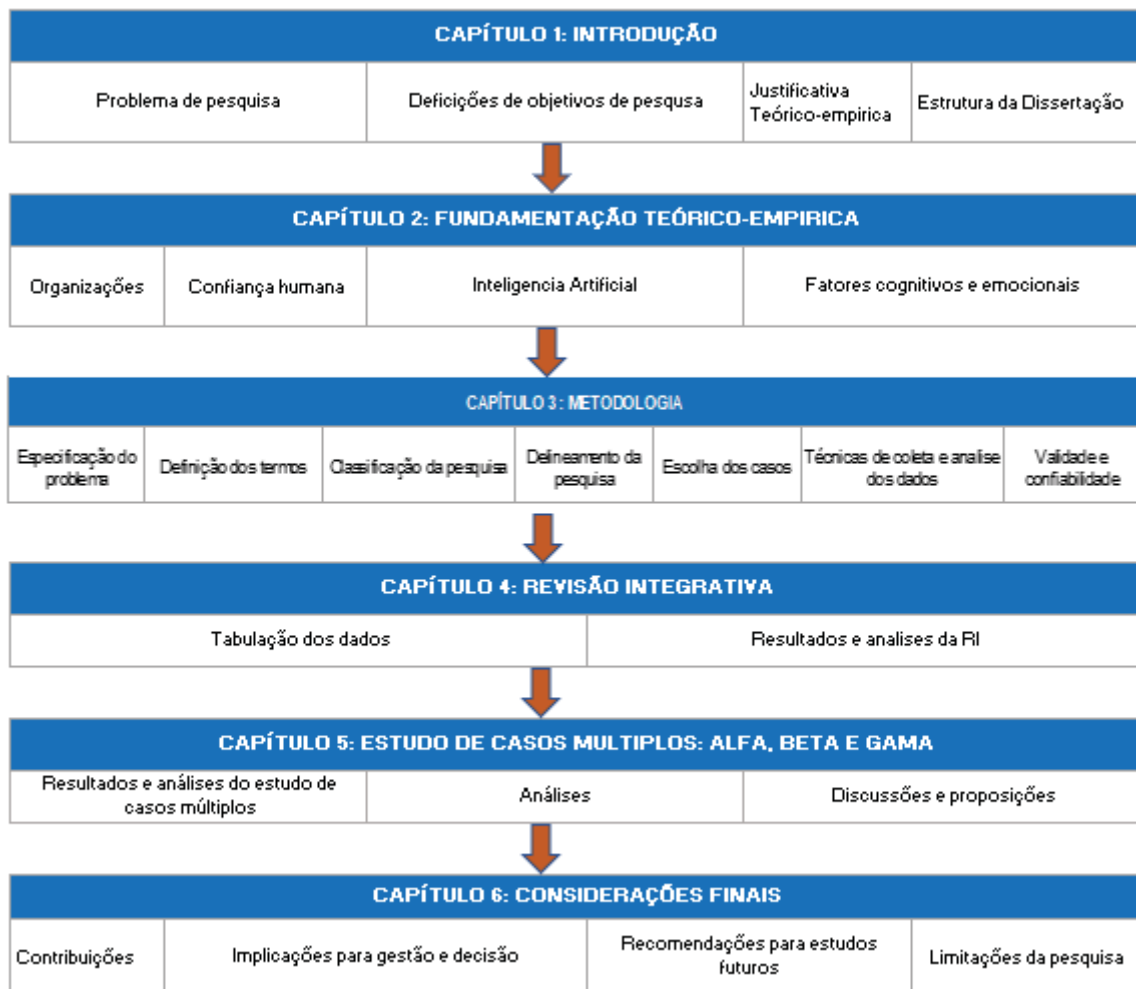
Para que sejam difundidos, sistemas artificialmente inteligentes precisam ser aceitos pela sociedade, principalmente em áreas em que as pessoas têm dificuldades de aceitar a interação com assistentes virtuais ou robôs físicos ao invés de humanos e isto requer confiança das pessoas no uso destas tecnologias o que implica na utilização da forma apropriada e que considere as nuances emocionais, empatia e subjetividade intrínsecos (HAGENDORFF; WEZEL, 2020). Desta forma, este trabalho

busca compreender a influência dos fatores cognitivos e emocionais que afetam confiança humana em IA em empresas do setor automotivo através de uma pesquisa organizada à partir do modelo proposto por GLIKSON; WOOLLEY (2020).

1.4 ESTRUTURA DA DISSERTAÇÃO

O presente trabalho está organizado em uma estrutura com 6 capítulos conforme apresentado na FIGURA 3, abaixo.

FIGURA 3 - ESTRUTURA DA DISSERTAÇÃO



FONTE: A autora (2022).

2 FUNDAMENTAÇÃO TEÓRICO-EMPIRICA

Este capítulo tem como objetivo promover reflexões acerca de conceitos e teorias que servirão de base para construção desta dissertação.

2.1 ORGANIZAÇÕES

Uma organização é um sistema com estruturas interativas complexas que envolve comportamentos sociais inter-relacionados entre pessoas e tecnologia através de poder, políticas e cultura organizacional. A função de uma organização é garantir a coerência entre os processos, padrões e normas estabelecidos pelas autoridades e tem duas tarefas fundamentais: tomar decisões e fazer as coisas acontecerem (SIMON, 1997), uma tentativa de colocar em ordem o fluxo de ações humanas necessárias para certos fins através da generalização e institucionalização das regras, padrões e significados particulares (NUTT; WILSON, 2010).

As pessoas passam pelo menos um terço do dia nas organizações, por isto elas são tão importantes. Este tempo faz com que as pessoas influenciem e sejam influenciadas pela organização (MARCH; SIMON, 1993).

O termo organização se refere ao padrão de comunicações e relações entre um grupo de seres humanos, incluindo os processos de tomada e implementação de decisões. Este padrão fornece aos membros da organização muitas das informações e muitas das suposições, objetivos e atitudes que entram em suas decisões, e fornece também um conjunto de expectativas estáveis e compreensíveis sobre o que os outros membros do grupo estão fazendo e como eles vai reagir ao que alguém diz e faz. O sociólogo chama esse padrão de "sistema de papéis"; estamos preocupados com a forma de sistema de funções conhecido como "organização" (SIMON, 1997, P.18-19).

“A organização administrativa é caracterizada pela especialização - tarefas específicas são delegadas a partes específicas da organização”, destaca Simon (1997). Por ser uma rede complexa os comportamentos de uma organização são orientados e influenciados pelos comportamentos daqueles que a operam executando as tarefas. Essas tarefas são distribuídas a partir da tomada de decisão e necessidade de execução, de acordo com as competências e habilidades dos indivíduos que por sua vez tem distintos objetivos e motivações pessoais cuja compreensão é imprescindível pois orientam as atividades da organização SIMON (1997).

Nas organizações uma das atividades mais importantes “é o desenvolvimento de programas para novas atividades que precisam entrar na rotina e isto implica na necessidade de tomar decisões”, conforme destacado por March e Simon (1993). O processo de tomada de decisão nas organizações envolve a coleta inteligente de fatos e valores para confecção de um planejamento e é limitado pelos seguintes motivos no que tange ao indivíduo: habilidades, hábitos e reflexos que não estão mais no reino do consciente, pelos valores e por aquelas concepções de propósito que o influenciam na tomada de decisão e pela extensão de seu conhecimento das coisas relevantes para seu trabalho (SIMON, 1997).

2.2 CONFIANÇA HUMANA

O termo confiança humana está relacionado a tendência de um indivíduo em assumir o risco de acreditar em algo que não se pode confirmar ou validar o resultado de forma que este fica vulnerável às ações de outra parte com base na expectativa de que esta irá realizar uma ação esperada ou prometida (SHARAN; ROMANO, 2020).

Nygard et al. (2017, p.182) destacam que “em um dado contexto, confiança pode ser usada como verbo transitivo, o que implica que existe um objeto (pessoa, máquina ou sistema) que é confiável. Isso poderia ser escrito na forma $A \rightarrow B$, onde A confia em B para cumprir algum propósito”. Confiança é contextual, específica para algum domínio com objetivos e propósito pretendidos cujo objetivo pode ser acessar recursos e dados, fornecer algum serviço, tomar uma decisão ou ainda controlar ou monitorar um processo. Nesse sentido, Nygard et al. (2017, p. 182) complementam ainda que “se uma pessoa confia em uma entidade que é outra pessoa ou em uma máquina ou sistema específico, há confiança de que a entidade entregará o desempenho conforme o esperado. A confiança, então, pode ser um substantivo que captura uma crença”.

“A confiança é um elemento vital tanto do lado humano quanto da máquina”, afirma Nygard et al. (2017, p.178). Neste sentido, Zhang et al. (2021, p.3) complementa afirmando que “A confiança humana é crítica para a adoção de sistemas de IA, especialmente em contextos de alta criticidade”. As pessoas usam a confiança como forma de orientar seu comportamento para tomar uma decisão racional quando elas não conseguem confiar nas normas sociais e em seus recursos cognitivos para tomar uma decisão. Esta decisão estará sempre baseada na crença de que a outra

parte, seja ela outra pessoa, uma empresa ou um sistema, fará o que se espera que seja feito (SHARAN; ROMANO, 2020).

2.3 INTELIGÊNCIA ARTIFICIAL

IA é uma forma de inteligência humana simulada por tecnologias complexas. Uma nova geração tecnológica capaz de interagir com o meio ambiente através da coleta e processamento de dados e interpretação de informações capazes de prever eventos, avaliar cenários, gera resultados, dar respostas ou formular instruções e se melhorar com base no conhecimento já adquirido (GLIKSON; WOOLLEY, 2020).

Lapinska et al. (2021) apontam que IA é um conceito extremamente amplo por abranger inúmeras tecnologias e soluções modernas que contam com algoritmos complexos e, por isto, é difícil encontrar uma definição comumente aceita de IA.

Asan et al., (2020) se referem a IA como “um processo de computador que toma decisões ótimas de maneira algorítmica com base em vários critérios, usando um ou mais modelos baseados em aprendizado de máquina”.

“A literatura futurística assume que a IA é um conjunto de algoritmos capazes de realizar todas as tarefas tão bem ou até melhor do que os humanos”, afirmam Glikson; Woolley (2020, p.628).

IA está diretamente relacionada à capacidade dos sistemas de tomar decisões ou realizar tarefas específicas com pelo menos uma representação parcial da inteligência humana, bem como a capacidade de aprender e melhorar com base nas informações recolhidas. É a contraparte à robótica em empresas de manufatura que apostam em robôs nos processos produtivos para facilitar a execução física dos operários e para apoiar a área executiva no processo de tomada de decisão cognitiva assumindo cada vez mais tarefas cognitivas complexas SASSMANNSHAUSEN et al. (2021).

Por ter a capacidade de traduzir a incerteza e complexidade em dados e decisões clínicas ou sugestões acionáveis - embora imperfeitas, IA pode ser encontrada em processos produtivos e produtos representada de várias formas, sendo as mais comuns:

- **Robô físico:** representações mecânicas físicas com similaridade humana contendo uma gama de funções e capacidades, que podem estar remotamente ou fisicamente presentes executando tarefas para

as quais foram programados (GLIKSON; WOOLLEY, 2020). Exemplo: Robô babá.

- **Agentes virtuais:** representação inteligente com característica humanas sem presença física, uma identidade distinta como *chat bot* ou avatar (GLIKSON; WOOLLEY (2020). Exemplo: Lu, o *chat bot* premiado do magazine Luiza.
- **Embutida:** é uma forma de IA invisível para os humanos, não possui representação visual e por ser totalmente incorporada a sistemas e pode ser incorporada a produtos e serviços tais como mecanismos de busca, programas e aplicativos. Exemplo: GPS. (GLIKSON; WOOLLEY (2020).

2.4 FATORES COGNITIVOS E EMOCIONAIS

Confiança é uma construção e pode ser influenciada por fatores cognitivos (avaliação racional e características situacionais) ou emocionais (baseado no afeto, conexões sociais) (GLIKSON; WOOLLEY, 2020). Neste sentido, as autoras apresentam como tais fatores influenciam o grau de confiança em cada tipo de IA (Robô físico, Agentes virtuais e embutida) propõem em seu framework as dimensões de confiança cognitiva (tangibilidade, transparência, confiabilidade, características da tarefa e comportamentos imediatos) e emocional (tangibilidade comportamentos imediatos e Antropomorfismo).

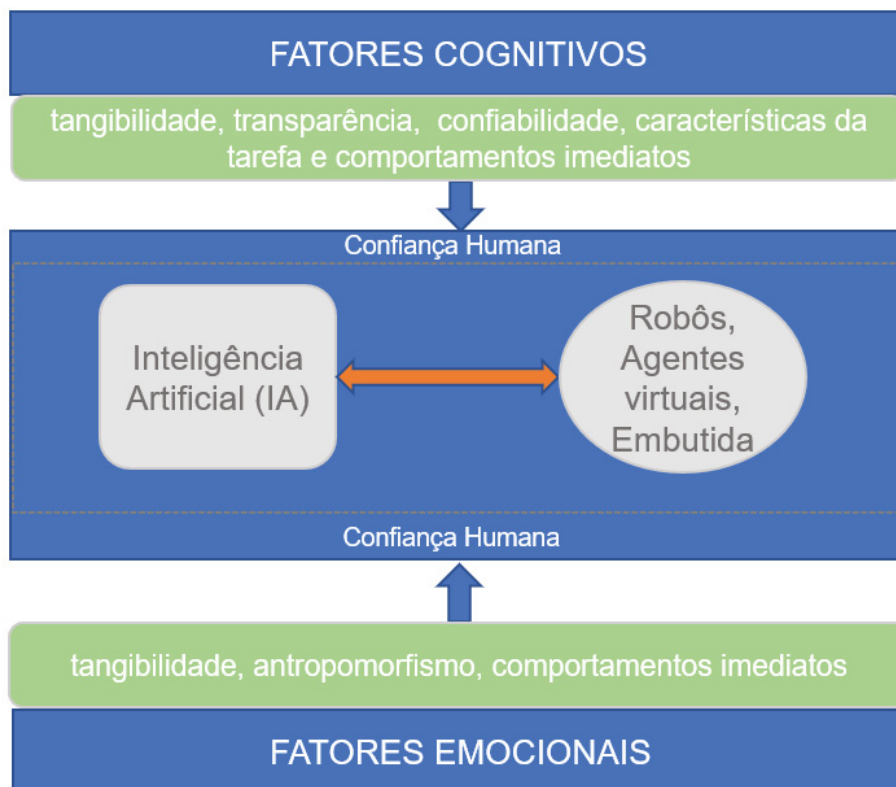
- **Tangibilidade:** Se refere a capacidade de ser percebido ou tocado durante o desenvolvimento de confiança (GLIKSON; WOOLLEY, 2020);

- **Transparência:** Está relacionada ao quanto a lógica interna da tecnologia é aparente para os usuários, reflete o nível em que as regras operacionais subjacentes estão sendo consideradas e apresentadas para desenvolver confiança na nova tecnologia (GLIKSON; WOOLLEY, 2020);

- **Confiabilidade:** Capacidade de apresentar o mesmo comportamento esperado ao longo do tempo, também é fundamental para a confiabilidade da tecnologia (GLIKSON; WOOLLEY, 2020);

- Características da tarefa: São tarefas relacionadas ao trabalho que a “tecnologia” ou IA está realizando. Como IA lida com julgamentos amplamente técnicos ou interpessoais pode ser um importante antecedente o papel das características da tarefa no desenvolvimento da confiança cognitiva (GLIKSON; WOOLLEY, 2020);
- Comportamentos imediatos: É o grau de percepção física e/ou psicológica relacionado à proximidade entre as pessoas. Referem-se a gestos socialmente orientados destinados a aumentar a proximidade interpessoal, como proatividade, escuta ativa e responsividade (GLIKSON; WOOLLEY, 2020); e
- Antropomorfismo: Semelhança humana, refere-se à percepção da tecnologia ou de um objeto como tendo qualidades humanas, como sentimentos. Percepções podem ser impulsionadas por recursos de interface, como a forma humana do robô, por características comportamentais, como olhar, e por enquadramento intencional, como dar a um robô ou bot um nome humano (GLIKSON; WOOLLEY, 2020).

FIGURA 4 - MODELO DE CONFIANÇA EM IA



FONTE: A autora (2022); Adaptado de Glikson e Woolley (2020)

Asan et al., (2020, p. 3) afirmam que “IA pode ser influenciado por vários fatores humanos”, tais como:

- Educação;
- Experiências anteriores;
- Preconceitos do usuário;
- Percepção em relação a automação;
- Propriedades do sistema de IA: controlabilidade, transparência e complexidade do modelo, riscos associados e outros.

Edmonds et al. (2019, p. 2) infere que “fatores cognitivos são uma crença ou uma avaliação com boas razões racionais”. Corroborando com esta afirmação Nygard et al. (2017, p.184) acrescenta que “em uma abordagem cognitiva, os estados de crença são a base da confiança”. Algumas características podem ajudar a fornecer base quantitativa para medir a confiança, tais como: disponibilidade, confiabilidade, segurança, usabilidade, acessibilidade, desempenho computacional e confiabilidade (NYGARD et al., 2017). Complementarmente, Sharan e Romano (2020, p.3) destacam que estes fatores podem estar relacionados “ao meio ambiente (colaboração da equipe, tarefa / contexto), relacionado ao parceiro (recursos e capacidades do parceiro), relacionados a humanos (características, estados, fatores cognitivos e emocionais).

Abaixo são destacados alguns fatores emocionais e cognitivos:

- **Fatores emocionais:** competência, habilidades e experiência e afinidade digital (GLIKSON; WOOLLEY, 2020);
- **Fatores cognitivos:** percepção de qualidade, plausibilidade, compulsividade das explicações (SASSMANNSHAUSEN et al., 2021).

O antropomorfismo “aumenta a resiliência da confiança em agentes cognitivos”, afirmam Sharan e Romano (2020, p.2). Aumentar a confiança humana em IA requer aumento da percepção de “cuidado”, características que incluem demonstrações de sentimentos e emoções ajudam neste sentido pois acredita-se que os robôs não têm estas características que ajudam a compreender o outro, como por exemplo, empatia YOKOI et al. (2021).

3 METODOLOGIA

Este capítulo descreve os procedimentos metodológicos empreendidos para alcançar o objetivo deste trabalho. A metodologia de pesquisa é a ciência que estuda como a pesquisa é feita cientificamente e tem como propósito resolver sistematicamente um problema de pesquisa através dos passos e critérios que são adotados por um pesquisador, de forma lógica, na busca de possíveis respostas para o problema de pesquisa apontado através da utilização de métodos, técnicas e ferramentas de pesquisa (KOTHARI, 2004).

Método e teoria são interdependentes e sua correlação permite que ambos possam alcançar os objetivos da pesquisa, sendo o método é o processo pelo qual os objetivos poderão ser alcançados (VERGARA, 2016).

Com isto, o presente capítulo foi estruturado da seguinte forma: a) especificação do problema; b) abordagem metodológica c) delimitação e desenho de pesquisa d) definição dos termos e) classificação da pesquisa; f) delineamento da pesquisa g) escolha dos casos e; h) técnicas de coleta e análise dos dados.

3.1 ESPECIFICAÇÃO DO PROBLEMA

Esta pesquisa tem por finalidade responder ao seguinte problema:

Quais são os fatores (cognitivos e emocionais) que afetam a confiança humana em IA nas organizações?

A partir da questão central apresentada no primeiro capítulo, e considerando os problemas e objetivos delineados neste estudo, foram elaboradas as seguintes perguntas de pesquisa.

3.1.1 Perguntas de pesquisa

- Identificar na literatura quais são os fatores que afetam a confiança humana em IA nas organizações
- Compreender quais são os fatores que influenciam a confiança humana em IA em empresas do setor automotivo e como estes fatores se relacionam aos diferentes tipos de IA (Robô, Agentes Virtuais e Bot/Embutida).

3.1.2 Desenho de pesquisa

Tendo em vista os objetivos apresentados na Introdução e os conceitos destacados na Fundamentação Teórica, criou-se o desenho de pesquisa ilustrado pela FIGURA 5 abaixo, destacando as duas etapas do presente estudo, RI e estudo de casos múltiplos.

Na etapa de RI, buscou-se **identificar na literatura quais são os fatores que afetam a confiança humana em IA nas organizações** com o objetivo de criar um quadro teórico de análise contendo conceitos de IA, confiança humana em IA e fatores cognitivos e emocionais que afetam a confiança humana em IA nas organizações.

Na segunda etapa, composta pelo **Estudo de Caso**, será elaborado um segundo quadro de análise com base nos resultados da RI para contextualizar os benefícios e dificuldades na adoção de IA em organizações do setor automotivo a fim de compreender **quais são os fatores que influenciam a confiança humana em IA em empresas do setor automotivo e como estes fatores se relacionam aos diferentes tipos de IA (Robô, Agentes Virtuais e Bot/Embutida)**. Nesta etapa, selecionou-se 3 empresas do setor automotivo que adotam IA em seus processos e produtos com a finalidade de verificar em que grau os resultados teóricos da RI são atendidos empiricamente.

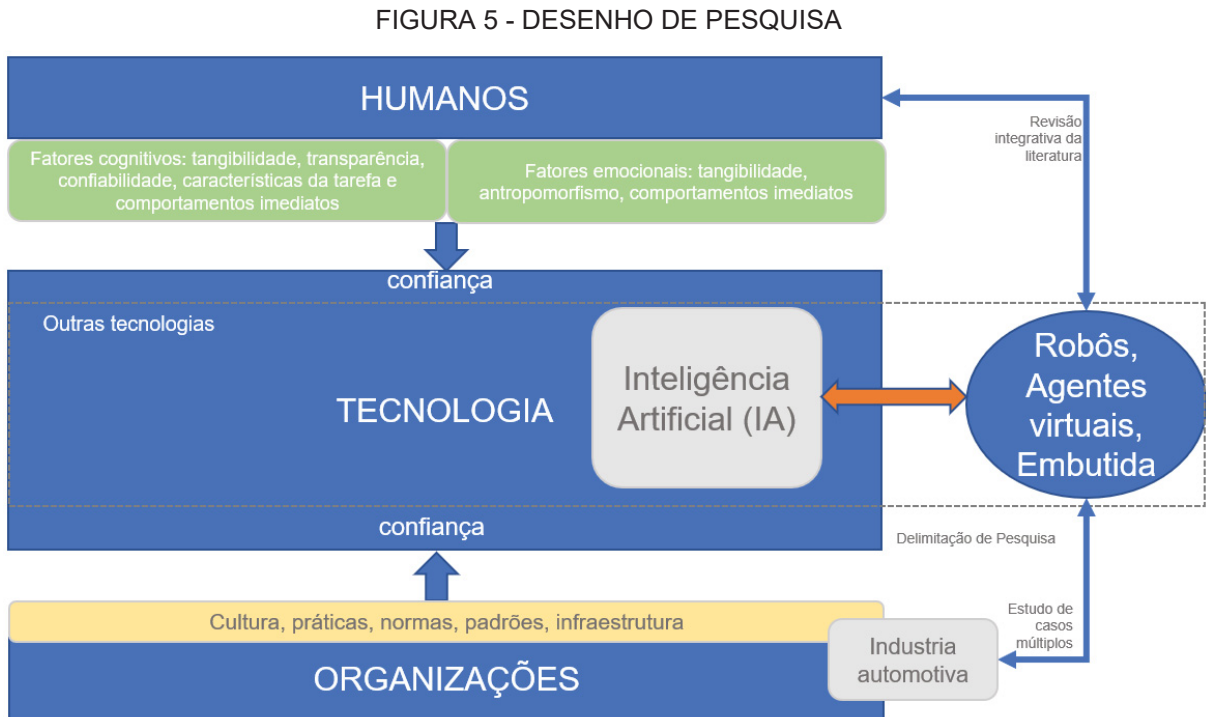
O estudo é feito considerando o impacto da relação entre confiança humana e os fatores cognitivos e emocionais que afetam a relação humano-IA no contexto de diferentes formas de representação de IA (robôs físicos, agentes virtuais e bots) em empresas do setor automotivo.

O nível de confiança humana em tecnologia, principalmente em IA deve ser calibrado pois influencia diretamente o resultado do uso. Glikson e Woolley (2020, p.629), ressaltam que “baixa confiança em tecnologia altamente capaz leva ao desuso e altos custos em termos de desperdício de tempo e baixa eficiência de trabalho”.

Desta forma, este estudo tem interesse em identificar quais fatores cognitivos e emocionais influenciam confiança humana em IA nos processos produtivos e produtos de empresas do setor automotivo e contribuir no processo de aprendizagem e promoção da construção de confiança no uso de IA colaborando com o avanço do setor ao “examinar os recursos de IA que facilitam o desenvolvimento da confiança, como tangibilidade, transparência, confiabilidade e comportamentos imediatos no

contexto da tarefa que está sendo executada” de acordo com o exposto por Glikson e Woolley (2020, p. 630).

A FIGURA 5, abaixo, consolida e sintetiza os conceitos estudados neste trabalho.



FONTE: A autora (2020); Adaptado de Glikson e Woolley (2020)

3.2 DEFINIÇÃO DOS TERMOS

Nesta sessão são apresentadas as definições conceituais (DC) e definições operacionais (DO) das principais categorias analíticas que embasam este trabalho com a finalidade de elucidar o significado dos principais conceitos, termos, definições e construtos utilizados de forma transparente a fim de não comprometer a confiabilidade dos resultados em relação ao fenômeno estudado (MARTINS; THEÓPHILO, 2007)

É indispensável a conceituação e termos e variáveis, a definição dos principais fim de que o investigador e interessados nos resultados compartilhem os mesmos entendimentos sobre os conceitos, definições, possíveis construtos e variáveis incluídas no estudo, compreendendo de maneira semelhante os resultados, conclusões e limitações da investigação (MARTINS; THEÓPHILO, 2007, p.30).

3.2.1 Confiança Humana

DC: Confiança humana é um aspecto essencial das relações interpessoais e se refere a um estado psicológico, uma crença, um mecanismo para lidar com a incerteza diante do que é conhecido e desconhecido. Representa uma tendência comportamental gerada a partir de perspectivas diversas cujo objeto da confiança não são apenas pessoas, mas também, instituições, sistemas computacionais e outros objetos (NYGARD et al., 2017); GLIKSON; WOOLLEY, 2020; ASAN et al., 2020; YOKOI et al., 2021; SHARAN; ROMANO, 2020; LAPINSKA et al., 2021).

DO: O conceito dos fatores de confiança humana propostos no modelo teórico de GLIKSON; WOOLLEY (2020) foi operacionalizado por meio dos roteiros utilizados para guiar as entrevistas semiestruturadas gravadas que ocorreram de forma virtual utilizando a ferramenta Microsoft Teams, cujas forneceram informações construção do QUADRO 9 - SÍNTESE DOS FATORES IDENTIFICADOS NO ESTUDO DE CASOS.

3.2.2 Inteligência artificial

DC: IA é uma tecnologia moderna altamente de alta capacidade de processamento de dados baseada em conceitos complexos e abrangentes com a finalidade de melhorar a tomada de decisão em situações envolvendo incerteza. Baseada em sistemas computacionais, IA tende a imitar a inteligência humana na forma de robôs físicos (representações mecânicas físicas com similaridade humana), virtuais (representação inteligente com característica humanas sem presença física) ou embutidos (forma de IA invisível para os humanos por ser totalmente incorporada a sistemas e produtos) e com capacidade de aprendizagem interativa das funções cognitivas humanas para processar informações gerando resultados (NYGARD et al., 2017); CURIONI, 2018); EDMONDS et al., 2019; GLIKSON; WOOLLEY, 2020; ASAN et al., 2020); ZHANG et al., 2021); LAPINSKA et al., 2021)).

DO: Para operacionalizar a pesquisa e análises foram preparados roteiros para entrevistas semiestruturadas gravadas, considerados os 3 tipos de representação de IA baseados no modelo proposto por (GLIKSON; WOOLLEY (2020):

robótica, virtual e embutida, onde os entrevistados identificaram quais tipos de IA foram implementados em cada organização e responderam as questões correlacionadas.

3.2.3 Fatores cognitivos e emocionais

DC: Fatores cognitivos e emocionais são fatores característicos do comportamento humano, relacionados a avaliações cognitivas ou racionais/situacionais ou emocionais/afetivas. Os fatores cognitivos e emocionais são diretamente afetados pelas crenças, culturas e ambiente em que as pessoas estão inseridas. Os fatores cognitivos envolvem avaliação racional e características situacionais, já os fatores emocionais são baseados no afeto e conexões sociais (GLIKSON; WOOLLEY (2020). No modelo de Glikson; Woolley (2020a) os fatores cognitivos são classificados nas dimensões de tangibilidade, transparência, confiabilidade, características da tarefa e comportamentos imediatos, enquanto os fatores emocionais são classificados nas dimensões de tangibilidade, comportamentos imediatos e antropomorfismo (GLIKSON; WOOLLEY (2020).

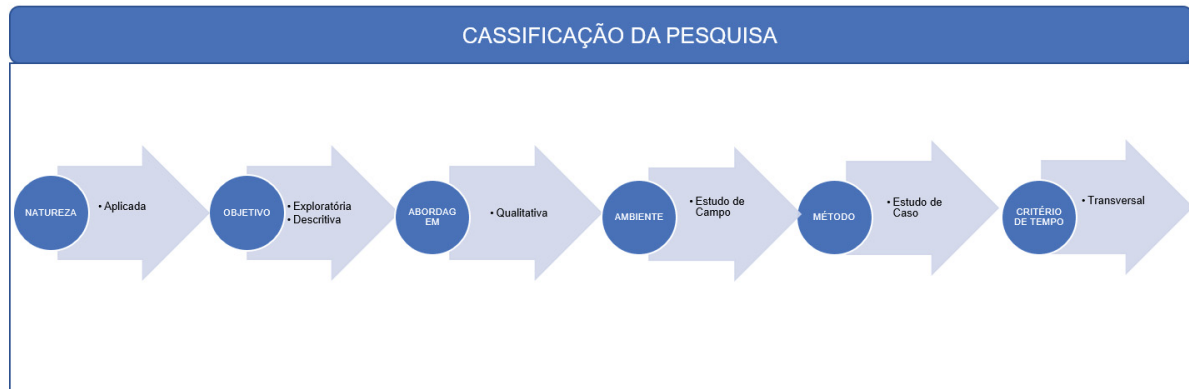
DO: Os conceitos de fatores cognitivos e emocionais propostos no modelo de Glikson e Woolley (2020) foram operacionalizados por meio da utilização de roteiros para entrevistas semiestruturadas e gravadas, realizadas através da ferramenta Microsoft Teams, onde as perguntas foram direcionadas considerando cada uma das dimensões apontadas acima, para cada tipo de IA, conforme Anexo I. A utilização de roteiros para as entrevistas semiestruturadas, conforme apresentado no capítulo 5, forneceu insumos para análise da classificação dos achados sobre as empresas estudadas.

3.3 CLASSIFICAÇÃO DA PESQUISA

Na

FIGURA 6, abaixo, é apresentada a classificação da pesquisa em relação à natureza, objetivo, abordagem, ambiente e critério de tempo.

FIGURA 6 - CLASSIFICAÇÃO DA PESQUISA



FONTE: A autora (2022).

Em relação à natureza, este trabalho se classifica como pesquisa aplicada pois, conforme definido por Pereira (2010, p.20), neste tipo de pesquisa “os conhecimentos adquiridos são utilizados para aplicação prática e voltados para a solução de problemas concretos da vida moderna”.

Os objetivos empregados nesta pesquisa são de caráter exploratório-descritivo por se tratar de uma pesquisa que busca explorar conceitos e fatos com pouca bibliografia.

Pesquisas exploratórias são aquelas que têm por objetivo explicitar e proporcionar maior entendimento de um determinado problema. A exploração é necessária a fim de desenvolver ideias e listar as questões de pesquisa valem a pena ser utilizadas (VOSS, 2010). Este tipo de pesquisa é realizado “em área na qual há pouco conhecimento acumulado e sistematizado. Por sua natureza de sondagem, não comporta hipóteses que, todavia, poderão surgir durante ou ao final da pesquisa”, conforme destacado por Vergara (2016, p.48). A finalidade da pesquisa exploratória é “desenvolver, esclarecer e modificar conceitos e ideias, tendo em vista a formulação de problemas mais precisos ou hipóteses a serem testadas em estudos posteriores”, conforme exposto por Gil (2019, p.26).

Pode ser considerada também como descritiva, pois, ainda que não tenha o compromisso de explicar os fenômenos que descreve, pode servir de explicação para

tal através da exposição de características de determinada população ou de determinado fenômeno, podendo ainda, estabelecer correlações entre variáveis e definir sua natureza (VERGARA, 2016). São pesquisas que tem por objetivo estudar características de um grupo tais como: opiniões, atitudes e crenças de uma determinada população (GIL, 2019).

A abordagem escolhida para esta investigação foi a pesquisa qualitativa, pois, emerge de paradigmas fenomenológicos e interpretativos. Dá ênfase em processos e significados pois tem como objetivo interpretar os significados das intenções dos atores. Os dados obtidos deste tipo de pesquisa são analisados de forma indutiva e representam os atos e expressões humanas (GODOI; BANDEIRA-DE-MELLO; SILVA, 2010) baseando-se “no pressuposto de que a realidade pode ser vista sob múltiplas perspectivas”, conforme apontado por Gil (2019, p.175). Creswell (2014, p.49-50) acrescenta que:

A pesquisa qualitativa começa com pressupostos e o uso de estruturas interpretativas/teóricas que informam o estudo dos problemas da pesquisa, abordando os significados que os indivíduos ou grupos atribuem a um problema social ou humano. Para estudar esse problema, os pesquisadores qualitativos usam uma abordagem qualitativa da investigação, a coleta de dados em um contexto natural sensível às pessoas e aos lugares em estudo e a análise dos dados que é tanto indutiva quanto dedutiva e estabelece padrões ou temas. O relatório final ou a apresentação incluem as vozes dos participantes, a reflexão do pesquisador, uma descrição complexa e interpretação do problema e a sua contribuição para a literatura ou um chamado à mudança (CRESWELL, 2014, p.49-50).

Vergara (2016, p.49) descreve a pesquisa de campo como: “investigação empírica realizada no local onde ocorre ou ocorreu um fenômeno ou que dispõe de elementos para explicá-lo. Pode incluir entrevistas, aplicação de questionários, testes e observação participante ou não”. Sendo assim, quanto ao ambiente, o presente estudo se enquadra em uma pesquisa de campo, uma vez que a coleta de dados foi feita por meio da estruturação de roteiros para entrevistas semiestruturadas e observações não participantes realizadas nas empresas selecionadas.

Quanto à dimensão de tempo, esta pesquisa se classifica como um estudo de corte transversal pois representa uma foto do momento específico em que o estudo foi realizado (COOPER, D. R.; SCHINDLER, 2016).

3.4 DELINEAMENTO DA PESQUISA

Foram adotadas duas estratégias de pesquisa para responder as questões de do presente trabalho: revisão de literatura com abordagem integrativa e estudo de casos.

Para **identificar na literatura quais são os fatores que afetam a confiança humana em IA nas organizações**, conduziu-se RI e elaborou-se um quadro de análise contendo conceitos de IA, confiança humana em IA e fatores cognitivos e emocionais que afetam a confiança humana em IA nas organizações. O método de RI foi escolhido por auxiliar a coleta, análise e construção de categorias a partir de pesquisas anteriores. Uma RI contribui integrando perspectivas de estudos teóricos e empíricos para responder a questões de pesquisa com um poder que nenhum estudo individual tem, conforme apontado por Snyder (2019).

Por se tratar de um tema emergente, a abordagem escolhida foi a RI, pois a “maioria das revisões integrativas da literatura se destina a abordar dois tipos de tópicos - tópicos maduros ou tópicos novos e emergentes”, destaca Torracco (2005, p. 357). Além disso, considerando o exposto por Snyder (2019, p. 334), “uma revisão integrativa é abordagem que pode ser útil quando o objetivo da revisão não é cobrir todos os artigos já publicados sobre o tema, mas sim para combinar perspectivas para criar novos modelos teóricos”.

Ainda, com o objetivo de compreender **quais são os fatores que influenciam a confiança humana em IA em empresas do setor automotivo e como estes fatores se relacionam aos diferentes tipos de IA (Robô, Agentes Virtuais e Bot/Embutida)**, conduziu-se estudo de casos.

Voss (2010, p. 197) define que “um estudo de caso é uma história de um fenômeno passado ou atual, extraída de múltiplas fontes de evidência. Pode incluir dados de observação direta e entrevistas sistemáticas, bem como de arquivos públicos e privados”. Seuring (2008, p. 130) infere que “um estudo de caso é uma investigação empírica que investiga um fenômeno contemporâneo dentro de seu contexto de vida real, especialmente quando os limites entre o fenômeno e o contexto não são claramente evidentes”. É possível pesquisar a mesma questão ou questões diferentes em distintos casos na mesma empresa, por exemplo (SEURING, 2008). A lógica de replicação é o ponto central para o uso do método de estudo de caso pois cada caso pode servir como um experimento distinto (EISENHARDT; GRAEBNER, 2014).

Uma das forças do método estudo de casos que é permite explorar variáveis desconhecidas inicialmente em seu cenário natural o que ajuda a responder as questões “por que” e “o quê” com base na natureza do fenômeno completo através de investigação e observação prática (VOSS, 2010).

Um dos principais motivos para a popularidade e relevância de construção de teoria a partir de estudos de caso é que é uma das melhores (se não a melhor) das pontes de rica evidência qualitativa para a pesquisa dedutiva convencional. Sua ênfase no desenvolvimento de construções, medidas e proposições teóricas testáveis torna a pesquisa de caso indutiva consistente com a ênfase na teoria testável dentro da pesquisa dedutiva convencional (EISENHARDT; GRAEBNER, 2014).

Voss (2010) afirma que “casos múltiplos têm maior validade externa do que casos únicos”, para tanto este trabalho considerou a abordagem de estudo de casos múltiplos por se tratar experimentos replicáveis para teoria emergente pois enfatizam o contexto em que os fenômenos ocorrem na realidade (EISENHARDT; GRAEBNER, 2014) levando em conta os elementos dos achados que possam contribuir para evidenciá-los de forma a elaborar mais perguntas em sobre cada caso (TRIVIÑOS, 1987).

Para realização deste estudo foram consideradas as etapas sugeridas por Eisenhardt (1989), descritas no **Erro! Fonte de referência não encontrada.** abaixo.

QUADRO 1 - ETAPAS DA PESQUISA

ETAPA	DESCRIÇÃO DA ATIVIDADE
Etapa 1 - Iniciar planejamento	Definir a pergunta de pesquisa e fornecer medidas de base para os construtos.
Etapa 2 - Selecionar casos	Especificar a população com base na teoria existente
Etapa 3 - Criar protocolos e instrumentos	Identificar instrumentos e definir os protocolos para coleta e análise de dados.
Etapa 4 - Analisar dados coletados	Realizar a análise dos dados utilizando diferentes métodos, incluir notas do campo estudado.
Etapa 5 – Relacionar à literatura	Utilizar a literatura para comparar e validar as evidências.
Etapa 6 - Encerrar pesquisa	Concluir o estudo.

FONTE: Adaptado de Eisenhardt (1989, p.533).

3.5 ESCOLHA DOS CASOS

A escolha adequada dos casos é imperativa na concepção da teoria de estudo de casos (EISENHARDT, 1989) e a resposta da pergunta de pesquisa e lacuna devem

estar diretamente relacionadas com as escolhas das unidades de análise de forma a gerar oportunidade de conectar aspectos empíricos à teoria (YIN, 2014).

Uma boa amostragem envolve pelo menos duas ações: conectar as questões de pesquisa de forma a estabelecer limites do que se pode estudar e criar uma estrutura de amostra para ajudar a descobrir, confirmar ou qualificar os processos básicos ou construtos que sustentam o estudo (VOSS, 2010). Com isto, o principal critério para escolha das empresas selecionadas é não probabilístico e intencional, principalmente considerando a localização, a similaridade na característica da implementação de IA nos processos, onde há uma expectativa de encontrar evidências nos casos selecionados em relação à pergunta de pesquisa, possibilitando o estudo de casos múltiplos.

3.6 TÉCNICA DE COLETA E ANÁLISE DOS CASOS

Para realização da primeira etapa, composta pela RI, foram aplicados os métodos destacados na sessão 4: Revisão Integrativa da Literatura.

Na segunda etapa, será conduzido o estudo de casos em 3 empresas multinacionais do setor automotivo. Em geral os estudos de caso combinam métodos como a coleta de dados métodos como arquivos, entrevistas, questionários e observações gerando evidencias que podem ser qualitativas (palavras), quantitativa (números), ou ambos (EISENHARDT, 1989). Para este estudo a coleta de dados será realizada por meio de dados primários: entrevista semiestruturada e gravada. “As entrevistas semiestruturadas utilizam um guia de entrevista com uma lista de tópicos a serem abordados, mas a entrevista pode seguir várias trajetórias decorrentes dos tópicos designados” (CHRISTENBERRY, 2017).

As entrevistas semiestruturadas foram escolhidas como ferramenta de coleta de dados pelos pontos positivos destacados por Yin (2015): são direcionadas e dão foco aos tópicos estudados e por serem perceptíveis fornecendo explicações, visões pessoais como percepção, atitudes e significados e permitem a obtenção de informações exclusivas pela interpretação do entrevistado (EISENHARDT, 1989).

Os entrevistados foram selecionados considerando os seguintes requisitos: envolvimento com projetos de implementação de IA e disponibilidade para as entrevistas e possíveis questionamentos de verificação. Foram realizadas 7 entrevistas semiestruturadas em profundidade e utilizando o Software Microsoft

Teams para gravação e transcrição a fim de que informações importantes não sejam descartadas (CHRISTENBERRY, 2017).

Seuring (2008, p. 128) aponta que “os estudos de caso permitem a observação direta do campo, o que foi particularmente adequado [...]” por fornecer benefícios como urgência, cobrindo os eventos em tempo real no contexto exato de cada caso (YIN, 2015) o que foi possível uma vez que a pesquisadora trabalha em uma empresa que fornece serviços para as empresas envolvidas no estudo e corriqueiramente tem contato com o ambiente organizacional e com as pessoas entrevistadas.

Os seguintes requisitos foram considerados para a seleção dos entrevistados: Estarem ativamente envolvidos no projeto de implantação de IA e estarem disponíveis para as entrevistas e possíveis questionamentos de verificação. Desta forma, foram realizadas 7 entrevistas semiestruturadas em profundidade, individuais e utilizando a ferramenta Microsoft Teams. Todas as entrevistas foram gravadas com o consentimento dos entrevistados de forma a garantir que informações relevantes não sejam excluídas (CHRISTENBERRY, 2017). Para preservar o anonimato solicitado pelas empresas, os entrevistados neste estudo de caso não foram identificados.

No **Erro! Fonte de referência não encontrada.** abaixo ficam esclarecidos detalhes dos entrevistados. O roteiro de entrevistas e o termo de autorização das entrevistas constam nos Anexo I e II, respectivamente.

QUADRO 2 - LISTA DE ENTREVISTADOS

ENTREVISTADO	PAPEL	EMPRESA	CARGA HORÁRIA	DATA DA ENTREVISTA	HORARIO
Entrevistado 1 (E1)	Engenheira de controle e automação do setor de processos.	ALFA	49:28 min	30/03/2022	10:00
Entrevistado 2 (E2)	Líder da equipe de ciência de dados para a América latina.	ALFA	45:44 min	31/03/2022	16:00
Entrevistado 3 (E3)	Líder da equipe de ciência de dados para o Brasil.	ALFA	26:10 min	30/03/2022	17:00

Entrevistado 4 (E4)	Gerente de aplicações e sistemas globais	BETA	30:17 min	13/05/2022	9:00
Entrevistado 5 (E5)	Líder do programa Champion de transformação digital	BETA	53:32 min	19/05/2022	16:00
Entrevistado 6 (E6)	Diretor de transformação digital	BETA	42:17	15/07/2022	16:00
Entrevistado 7 (E7)	Líder de IT global	GAMA	49:10 min	27/05/2022	15:00

FONTE: A autora (2022).

3.7 VALIDADE E CONFIABILIDADE

Qualidade, adequação e rigor são algumas das variações acerca de validade em pesquisa qualitativa. Os critérios de validade e de confiabilidade assumem aspectos particulares em pesquisas qualitativas e isso se deve à característica interpretativa que leva à subjetividade do pesquisador (HAYASHI; ABIB; HOPPEN 2019; DE PAIVA JÚNIOR; DE SOUZA LEÃO; DE MELLO, 2011).

Os elementos validade e confiabilidade são essenciais para garantir a precisão e rigor em pesquisas científicas. Um dos pontos fortes da pesquisa qualitativa é a aplicação de estratégias de validade. O pesquisador verifica a precisão dos resultados empregando alguns procedimentos e tem como base determinar se a precisão dos resultados existe a partir do ponto de vista do pesquisador, do participante ou dos leitores de um relato, enquanto a confiabilidade qualitativa indica se a abordagem do pesquisador é consistente entre diferentes pesquisadores e diferentes projetos (CRESWELL, 2014).

“Para garantir a validade e confiabilidade da pesquisa, a apresentação do processo é de fundamental importância. Dessa forma, melhores histórias e melhores construções podem ser alcançadas”, afirma Seuring (2008, p. 135). “Casos múltiplos têm maior validade externa do que casos únicos”, afirma Voss (2010, p.202). Existem riscos em todas as pesquisas, mas estes podem ser um pouco mitigados quando eventos e dados são comparados entre os casos. Os estudos de casos múltiplos

podem reduzir a profundidade do estudo quando os recursos são limitados e tendem a aumentar a validade externa e ajudar a proteger contra o viés do observador (VOSS, 2010).

Desta forma, o presente estudo buscou seguir boas práticas em pesquisa qualitativa, como as recomendadas por Eisenhardt (1989) em que os procedimentos e métodos adotados são apresentados de forma clara e detalhada, e a realização de feedback com os entrevistados para esclarecimentos sempre que necessário a fim de minimizar problemas de interpretação e vieses.

4 REVISÃO INTEGRATIVA

Por se adequar melhor aos objetivos deste trabalho a técnica escolhida foi a RI, pois permite a realização de análise, crítica e síntese da literatura de forma a integrar e criar novas perspectivas (TORRACO, 2005).

Uma RI é um método específico que resume o passado da literatura empírica e/ou teórica com o objetivo de traçar uma análise sobre o conhecimento já construído anteriormente sobre um dado tema a fim de fornecer uma compreensão mais abrangente de um fenômeno particular possibilitando a síntese de vários estudos já publicados, gerando novos conhecimentos (BOTELHO, CUNHA E MACEDO, 2011). A finalidade da RI é integrar perspectivas através do mapeamento e visão geral das publicações e não necessariamente cobrir todas as publicações existentes (SNYDER, 2019; TORRACO, 2016).

O primeiro passo foi realizar o planejamento da RI definindo as bases Scopus e Web of Science. Já os critérios de busca foram estabelecidos considerando os campos de busca, título e resumo, e as chaves de busca "human trust" OR "people trust" AND "artificial intelligence" OR "AI". A autora também reduziu a busca a "artigos", "publicados na língua inglesa" e sem delimitação de tempo. No dia 19/02/2022 foi realizada a busca que resultou em 22 artigos encontrados na base Scopus e 12 artigos encontrados na base WOS, totalizando 34 artigos. Destes, identificou-se 9 artigos repetidos, que foram removidos, totalizando 25 artigos encontrados.

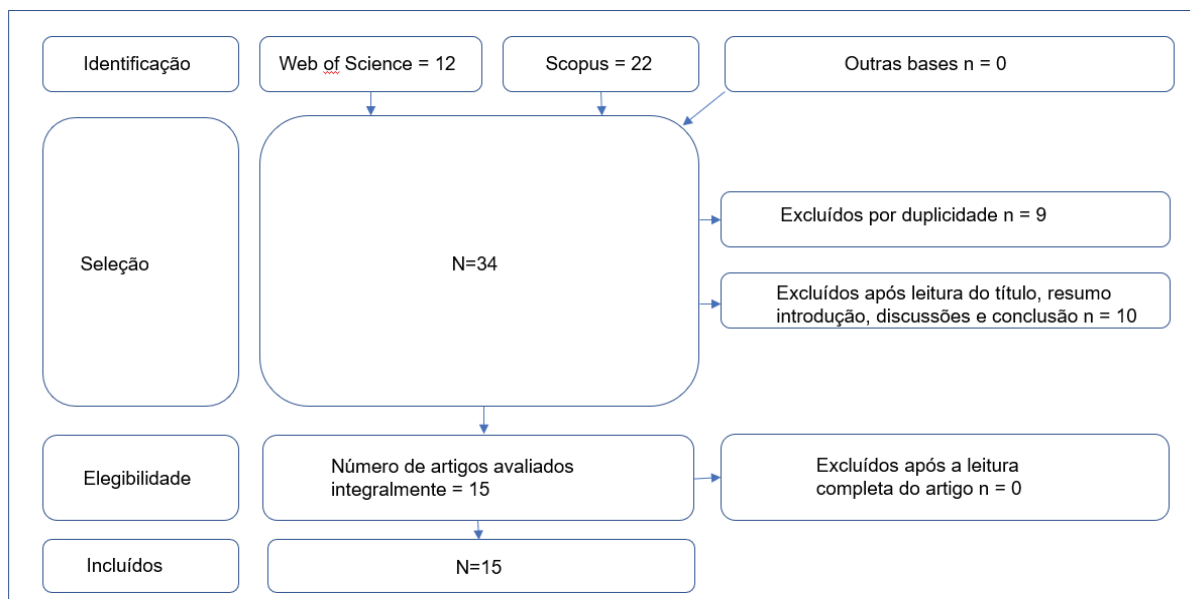
Considerando a recomendação de Snyder (2019), que sugere que o pesquisador deve utilizar uma abordagem qualitativa para avaliar os achados após seguir um processo de revisão sistemática para reunir as publicações, autora iniciou a etapa de leitura e análise dos 25 artigos, que foram revisados pela leitura do título, resumo, introdução, discussões e conclusões, excluindo aqueles que estavam fora do contexto adotado de confiança humana em IA. Por exemplo Kundu (2019) que possui foco específico sobre *Blockchain* e confiança em *Smart Cities*, **para tanto foi descartado desta RI**, resultando assim em 15 artigos selecionados.

A análise de conteúdo foi realizada com base nas etapas propostas por Bardin (2011): a) pré-análise através da leitura integral dos artigos b) exploração, codificação e categorização e c) interpretação e inferência e serviram de base para construção de um Quadro Teórico de Análise, conforme apresentado abaixo no QUADRO 4,

contemplando a evolução dos conceitos de confiança, IA e fatores cognitivos e emocionais.

As fases adotadas para RI podem também ser observadas na FIGURA 7 que ilustra o fluxograma Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (MOHER et al., 2009).

FIGURA 7 - FLUXO PRISMA



FONTE: A autora (2022).

A TABELA 1 abaixo apresenta a quantidade de publicações por ano. É possível observar que 46% das publicações encontradas foram realizadas em 2021 e 15% no primeiro trimestre de 2022 quando a busca foi realizada, demonstrando uma tendência de aumento do interesse em publicações para o tema estudado neste trabalho.

TABELA 1 - QUANTIDADE DE PUBLICAÇÕES POR ANO

Período	Contagem por ano
Março/2022	4
2021	12
2020	3
2019	4
2018	2
Total Geral	26

FONTE: A autora (2022).

4.1 TABULAÇÃO DOS DADOS

A tabulação dos 15 artigos elegidos foi realizada utilizando como ferramenta para registro, organização e análise de dados o software Microsoft Excel, estruturando os seguintes campos descritivos: identificador, ano de publicação título, autores, breve resumo, identificação dos conceitos de confiança, IA e fatores cognitivos e emocionais. A análise de conteúdo foi estruturada de acordo com as três fases propostas por Bardin (2011): a) Análise prévia e primeira leitura dos 15 artigos selecionados; b) Categorização e codificação dos dados; e c) Inferência e interpretação dos dados.

4.2 RESULTADOS E ANÁLISES DA RI

4.2.1 Quadros de análise

A partir da revisão integrativa e análise de conteúdo dos 15 artigos selecionados, foram elaboradas os Quadros 3 e 4. O Quadro 3 apresenta a síntese dos conceitos de confiança humana, inteligência emocional e fatores cognitivos e emocionais identificados na RI e suas respectivas referências. O Quadro 4 apresenta a evolução dos conceitos identificados na literatura.

QUADRO 3 - SÍNTESE DOS CONCEITOS E CONSTRUTOS IDENTIFICADOS NA RI

ID	ANO	TÍTULO	OBJETIVO	Confiança Humana	Inteligência Artificial	Fatores Cognitivos e Emocionais	AUTORES
1	2022	<i>Can people experience romantic love for artificial intelligence? An empirical study of intelligent assistants</i>	Contribuir com pesquisas relacionadas a IA de forma a oferecer uma compreensão do uso de assistentes inteligentes através da perspectiva do amor. Investigando aplicações de IA a partir de novos ângulos teóricos.	A disposição de confiança refere-se à propensão individual inerente a confiar ou desconfiar dos outros. É uma característica que governa como ele/ela vê a confiabilidade de qualquer outra entidade.	AI são as tecnologias de computador que permitem que as máquinas simulem funções cognitivas humanas, como aprender e resolver problemas.	A capacidade da computação cognitiva para realizar tarefas e a capacidade emocional de entender e interagir com os usuários também são consideradas as mais importantes para a experiência dos usuários com agentes de IA.	SONG, XU e ZHAO (2022)
2	2022	<i>What is it like to trust a rock? A functionalist perspective on trust and trustworthiness in artificial intelligence</i>	Examinar a partir de princípios básicos de como os humanos podem pensar sobre tecnologia e por que alguns princípios básicos que já existem ainda podem ser válidos quando a tecnologia com a qual estamos interagindo é mais autônoma e potencialmente menos compreendida e menos sob controle do que as tecnologias existentes.	Confiança é uma heurística amplamente utilizada por humanos em uma variedade de situações. É o ato de optar por se colocar em uma situação de risco, onde os resultados são dependente das ações de outro.	IA é considerada meramente um conjunto de processos, modelos e ferramentas projetados e fabricados,	N/A	LEWIS e MARSH (2022)
3	2022	<i>Human confidence in artificial intelligence and in</i>	Este trabalho investiga a evolução da confiança humana em IA e em si	N/A	N/A	N/A	CHONG et al. (2022)

		<i>themselves: The evolution and impact of confidence on adoption of AI advice</i>	<p>mesma, e seu impacto na tomada de decisões IA-humana. Especificamente, este estudo explora um dos contextos de colaboração mais prevalentes conhecido como tomada de decisão assistida por IA, onde os humanos são responsáveis pela decisão final.</p> <p>Neste trabalho, o foco é no aumento da confiança humana justificada (JT) em uma rede neural convolucional profunda (CNN), por meio de explicações onde a confiança justificada é calculada com base em dados humanos.</p>						
4	2022	<i>CX-ToM: Counterfactual explanations with theory-of-mind for enhancing human trust in image recognition models</i>		N/A	N/A	N/A	N/A	N/A	AKULA <i>et al.</i> (2022)
5	2021	<i>Artificial Intelligence Is Trusted Less than a Doctor in Medical Treatment Decisions: Influence of Perceived Care and Value Similarity</i>	<p>Comparar o nível de confiança dos indivíduos em IA x em médicos humanos através de um experimento online para explorar a confiança das pessoas em um sistema que sugere um tratamento.</p>	<p>O termo "confiança" se refere a um estado psicológico em que a vulnerabilidade e a incerteza são componentes essenciais.</p>	<p>IA são sistemas que tem a capacidade de superar ou especialistas humanos em certas áreas de domínio.</p>	<p>A percepção de cuidado e outras experiências percebidas como sentimentos de emoção, incluindo empatia.</p>		YOKOI <i>et al.</i> (2021)	

6	2021	<p><i>Effect of AI Explanations on Human Perceptions of Patient-Facing AI-Powered Healthcare Systems</i></p>	<p>Através de uma pesquisa empírica medir quantitativamente como as explicações da IA impactam as percepções dos consumidores de saúde ao usar sistemas de saúde baseados em IA voltados para o paciente e avaliar os distintos efeitos.</p>	<p>N/A</p>	<p>AI são sistemas modernos projetados para utilizar e aprender com grandes conjuntos de dados enquanto imita funções cognitivas humanas para processar informações de entrada para gerar resultados relevantes para apoio à decisão.</p>	<p>Percepções humanas de várias perspectivas, incluindo, confiança/convicção, semelhança humana, justificativa adequada e compreensibilidade/ explicações claras, informações sobre performance do modelo, transparência.</p>	<p>ZHANG <i>et al.</i> (2021)</p>
7	2021	<p><i>Employees' Trust in Artificial Intelligence in Companies: The Case of Energy and Chemical Industries in Poland</i></p>	<p>Por meio de um estudo quantitativo, verificar hipóteses examinando as ligações entre a confiança dos funcionários em IA na empresa e três outras variáveis latentes (confiança geral na tecnologia, confiança intraorganizacional e individual confiança de competência).</p>	<p>Pode-se presumir que a confiança é algum tipo de crença que a parte tem em relação ao comportamento futuro ou estados de o objeto de confiança. É uma variável complexa, multifacetada e multidimensional, latente e difícil de medir diretamente. Pode ser considerada um traço de personalidade e, portanto, uma variável interpessoal. Os objetos de confiança não são apenas pessoas, mas também</p>	<p>IA é um sistema ou uma máquina que imita a inteligência humana no desempenho de tarefas específicas e, além disso, pode melhorar (aprender) de forma interativa com as informações coletadas.</p>	<p>Percepção sob a ótica e avaliação das pessoas sobre consistência, confiabilidade, funcionalidade e suporte.</p>	<p>LAPINSKA <i>et al.</i> (2021)</p>

	8	<p><i>How to Evaluate Trust in AI-Assisted Decision Making? A Survey of Empirical Methodologies</i></p>	<p>Neste artigo é apresentado um levantamento dos métodos existentes para investigar a confiança na tomada de decisão assistida por IA e analisar o corpus ao longo dos elementos constitutivos de um protocolo experimental.</p>	<p>Uma atitude de que um agente atingirá o objetivo de um indivíduo em uma situação caracterizada por incerteza e vulnerabilidade.</p>	<p>Tem o papel de ajudar os humanos a tomar decisões com resultados incertos, como contratação, atribuição de tratamento ou processos de investigação.</p>	<p>N/A</p>	<p>VERESCHAK, BAILLY e CARAMIAUX (2021)</p>
	9	<p><i>Trust in artificial intelligence within production management—an exploration of antecedents</i></p>	<p>Neste artigo, uma análise dentro do gerenciamento de produção identifica possíveis variáveis antecedentes sobre a confiança na IA e avalia isso devido a cenários de interação com a IA.</p>	<p>Uma atitude de que um agente ajudará a atingir os objetivos de um indivíduo em uma situação caracterizada por incerteza e vulnerabilidade</p>	<p>IA é uma nova tecnologia que requer novas formas de interação e têm o potencial de assumir o trabalho do conhecimento para que as atividades humanas possam mudar o foco para a resolução de problemas e atividades de monitoramento e, mais amplamente, no trabalho criativo e emocional.</p>		<p>SASSMANNSHAUSEN et al. (2021)</p>

10	2020	<p><i>Artificial Intelligence and Human Trust in Healthcare: Focus on Clinicians</i></p>	<p>Apresentar fatores que moldam a confiança entre os médicos e a IA destacando os desafios críticos relacionados à confiança que devem ser considerados durante o desenvolvimento de qualquer sistema de IA para uso clínico.</p>	<p>A confiança é um mecanismo psicológico para lidar com a incerteza entre o conhecido e o desconhecido. Uma crença humana que é amplamente definido com base nas dimensões principais: benevolência, integridade e capacidade.</p>	<p>IA pode ser definida como uma tecnologia ou programa de computador que pode tomar decisões inteligentes para melhorar a tomada de decisão envolvendo incerteza e complexidade em sistemas.</p>	<p>Fatores humanos que influenciam, como: educação, experiências anteriores, preconceitos do usuário e percepção em relação automação, bem como propriedades de sistemas de IA incluindo controlabilidade, transparência e complexidade do modelo, riscos associados e outros.</p>	<p>ASAN <i>et al.</i> (2020)</p>
11	2020	<p><i>Human trust in artificial intelligence: Review of empirical research</i></p>	<p>Por meio de uma revisão bibliográfica identificar as formas de representação de IA e suas capacidades como antecedentes importantes para o desenvolvimento da confiança e propor um modelo que aborda os elementos que moldam os aspectos cognitivos e emocionais de confiança.</p>	<p>Confiança representa uma tendência a assumir riscos significativos devido a crença em um resultado positivo comum em diferentes disciplinas. Vulnerabilidade em relação às ações de outra parte com base na expectativa de que a outra irá realizar uma ação importante para quem confiou, independentemente da capacidade de monitorar ou controlar essa outra parte.</p>	<p>IA representa uma tecnologia altamente capaz e complexa que visa simular a inteligência humana.</p>	<p>Fatores cognitivos: tangibilidade, transparência, confiabilidade, características da tarefa, imediatismo. Fatores emocionais: tangibilidade, antropomorfismo e imediatismo.</p>	<p>GLIKSON e WOOLLEY (2020)</p>

12	2020	<i>The effects of personality and locus of control on trust in humans versus artificial intelligence</i>	Investigar o efeito das características de personalidade (<i>Big Five Inventory</i> (BFI) traços e locus de controle (LOC)) no comportamento de confiança, e até que ponto as pessoas confiam no conselho de um algoritmo baseado em IA, mais do que humanos, em um jogo de cartas de tomada de decisão.	A confiança é uma construção complexa e multifacetada. um aspecto essencial das relações interpessoais.	IA é uma interface capaz de apresentar às pessoas as informações em um ritmo suficiente e em uma forma adequada para que eles entendam e ajam.	Fatores relacionado ao meio ambiente tais como, colaboração, contexto da tarefa, recursos e capacidades, e fatores relacionados a humanos, características, estados, antropomorfismo.	SHARAN e ROMANO (2020)
13	2019	<i>A tale of two explanations: Enhancing human trust by explaining robot behavior</i>	Através de um experimento psicológico em que robôs aprendem através de demonstrações humanas, examinar quais formas de explicação melhor promovem a confiança humana nas máquinas e propor uma estrutura na qual explicações são geradas tanto do ponto de vista funcional quanto do mecanicista.	Confiança é a contrapartida dos humanos que geram explicações a partir de perspectivas funcionais que descrevem os efeitos ou objetivos das ações e de perspectivas mecanicistas que focam no comportamento como um processo.	IA é uma representação simbólica gerada a partir das observações de demonstrações humanas, onde IA aprende, executa e explica usando uma tecnologia.	Fatores cognitivos: uma crença ou uma avaliação com "boas razões racionais	EDMONDS et al. (2019)
14	2018	<i>Artificial intelligence: Why we must get it right</i>	Descrever os benefícios práticos de IA e dos computadores cognitivos que fazem IA funcionar e	N/A	AI é uma tecnologia de caixa preta em que mudando os pesos neurais profundamente no	Percepção de segurança, justiça, responsabilidade e transparência.	CURIONI (2018)

				delinear como a IBM e parceiros estão percorrendo o caminho no uso de transparente de tecnologias cognitivas.		aprendizado redes fazem suas próprias coisas, longe da vista, longe mente.		
15	2017	<i>People and intelligent machines in decision making</i>	Apresentar modelos que capturem as relações entre os níveis de serviço e a confiança examinando os sistemas ciber físicos em termos dos níveis de autonomia inteligente que eles podem possuir, incluindo questões especiais que surgem em sistemas altamente distribuídos e móveis e o advento do aprendizado de máquina avançado, considerando suas competências em conjunto com as máquinas, incluindo os erros e erros cometidos por humanos.	Confiança é crença firme na competência de uma entidade para agir de forma confiável dentro de um contexto específico.	IA tem a ver com máquinas e sistemas capazes de aconselhar ou tomando decisões normalmente tomadas por seres humanos.	Em uma abordagem cognitiva, os estados de crença são a base da confiança. A lógica difusa pode ser usada para fornecer uma base quase quantitativa para medir a confiança, tais como: disponibilidade, confiabilidade, segurança, usabilidade, acessibilidade, desempenho computacional e confiabilidade.		NYGARD et al. (2017)

FONTE: A autora (2022).

Łapińska et al. (2021) define confiança como um estado psicológico ou algum tipo de crença. Asan et al. (2020) definem confiança como um mecanismo psicológico ou uma crença humana. Já para Glikson e Woolley (2020), confiança representa uma tendência a assumir riscos significativos devido a crença em um resultado positivo. A semelhança dentre os termos favoreceu a integração e criação do conceito de confiança humana apresentado no **Erro! Fonte de referência não encontrada..**

A definição de Zhang et al. (2021) para IA são sistemas modernos projetados para utilizar e aprender com grandes conjuntos de dados enquanto imita funções cognitivas humanas. Já Yokoi et al. (2021) definem IA como sistemas que tem a capacidade desempenhar ou superar especialistas humanos em certas áreas de domínio. Ambas as definições convergem para integração e criação do conceito de inteligência artificial apresentado no **Erro! Fonte de referência não encontrada..**

Capacidade emocional de entender e interagir com os usuários é a definição dada por SONG et al., (2022) para fatores cognitivos e emocionais. Yokoi et al. (2021) define fatores cognitivos e emocionais como percepção de cuidado e outras experiências percebidas como sentimentos de emoção, incluindo empatia. Com isto, favorecendo a convergência para integração e criação do conceito de fatores cognitivos e emocionais apresentado no **Erro! Fonte de referência não encontrada..**

Diante deste contexto, a partir do

QUADRO 3, foi elaborado o **Erro! Fonte de referência não encontrada.** acerca da evolução dos construtos: confiança humana, inteligência artificial e fatores cognitivos e emocionais.

QUADRO 4 - EVOLUÇÃO DOS CONCEITOS

	CONCEITOS	REFERÊNCIA
Confiança Humana	Confiança humana é um aspecto essencial das relações interpessoais e se refere a um estado psicológico, uma crença, um mecanismo para lidar com a incerteza e uma tendência a lidar com riscos diante do que é conhecido e desconhecido. Representa comportamentos gerados a partir de perspectivas diversas cujo objeto não são apenas pessoas, mas também, instituições, sistemas computacionais e outros objetos.	SONG <i>et al.</i> , 2022; LEWIS e MARSH, 2022; YOKOI <i>et al.</i> , 2021; LAPINSKA <i>et al.</i> , 2021; SASSMANNSHAUSEN <i>et al.</i> , 2021; VERESCHAK <i>et al.</i> , 2021; SASSMANNSHAUSEN <i>et al.</i> , 2021; ASAN, BAYRAK, e CHOUDHURY, 2020; GLIKSON e WOOLLEY, 2020; SHARAN e ROMANO, 2020; EDMONDS <i>et al.</i> , 2019; NYGARD <i>et al.</i> , 2017.
Inteligência Artificial	IA é uma tecnologia moderna altamente de alta capacidade de processamento de dados baseada em conceitos computacionais complexos e abrangentes com a finalidade de melhorar a tomada de decisão em situações envolvendo incerteza. Baseada em sistemas computacionais, IA tende a imitar a inteligência humana na forma de robôs físicos, virtuais ou embutidos e com capacidade de aprendizagem interativa das funções cognitivas humanas para processar informações gerando resultados	SONG <i>et al.</i> , 2022; LEWIS e MARSH, 2022; YOKOI <i>et al.</i> , 2021; LAPINSKA <i>et al.</i> , 2021; SASSMANNSHAUSEN <i>et al.</i> , 2021; VERESCHAK <i>et al.</i> , 2021; SASSMANNSHAUSEN <i>et al.</i> , 2021; ASAN <i>et al.</i> , 2020; GLIKSON e WOOLLEY, 2020; SHARAN e ROMANO, 2020; EDMONDS <i>et al.</i> , 2019; CURIONI, 2018; NYGARD <i>et al.</i> , 2017.
Fatores Cognitivos e Emocionais	São fatores característicos do comportamento humano, relacionados a avaliações cognitivas ou racionais/situacionais ou emocionais/afetivas. Os fatores cognitivos e emocionais são diretamente afetados pelas crenças, culturas e ambiente em que as pessoas estão inseridas.	SONG <i>et al.</i> , 2022; YOKOI <i>et al.</i> , 2021; ZHANG <i>et al.</i> , 2021; LAPINSKA <i>et al.</i> , 2021; SASSMANNSHAUSEN <i>et al.</i> , 2021; SASSMANNSHAUSEN <i>et al.</i> , 2021; ASAN <i>et al.</i> , 2020; GLIKSON e WOOLLEY, 2020; SHARAN e ROMANO, 2020; EDMONDS <i>et al.</i> , 2019; CURIONI, 2018; NYGARD <i>et al.</i> , 2017

FONTE: A autora (2022).

4.2.2 Análises

Os 15 artigos selecionados foram revisados e categorizados por autor e por dimensão, considerando a popularidade de cada dimensão nos achados, conforme apresentado no QUADRO 5.

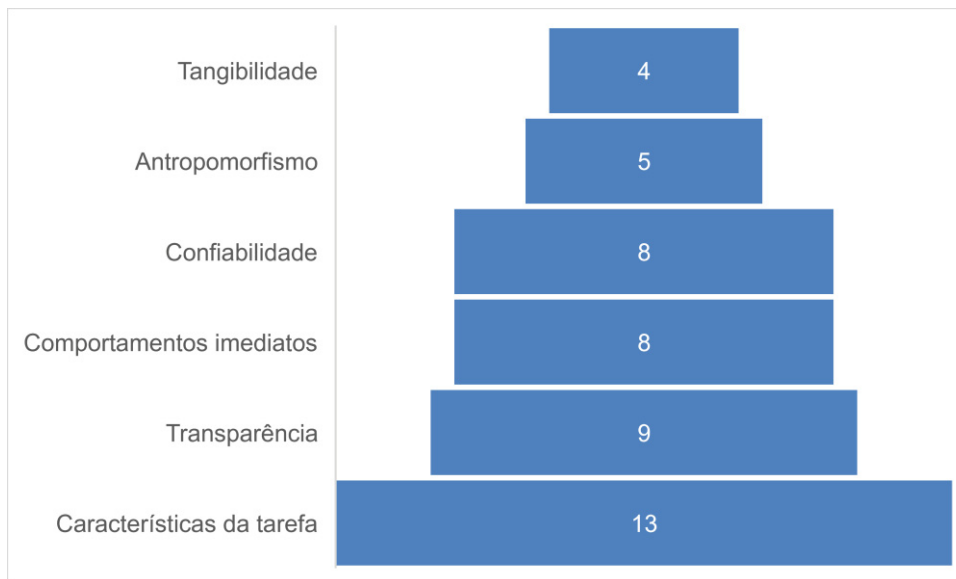
QUADRO 5 - SÍNTESE DOS FATORES COGNITIVOS E EMOCIONAIS IDENTIFICADOS NA RI

DIMENSÕES	FATORES COGNITIVOS E EMOCIONAIS QUE INFLUENCIAM A CONFIANÇA EM IA	Song <i>et al.</i> (2022)	Lewis e Marsh (2022)	Chong <i>et al.</i> (2022)	Akula <i>et al.</i> (2022)	Yokoi <i>et al.</i> (2021)	Zhang <i>et al.</i> (2021)	Lapinska <i>et al.</i> (2021)	Vereschak <i>et al.</i> (2021)	Sassmann shausen <i>et al.</i> (2021)	Asan, <i>et al.</i> (2020)	Glikson e Woolley (2020)	Sharan e Romano (2020)	Edmonds <i>et al.</i> (2019)	Curioni (2018)	Nygaard <i>et al.</i> (2017)
Tangibilidade	Presença ou representação humana física/virtual	X						X				X	X			
Transparência	Visibilidade e entendimento de como IA está sendo implantada		X		X		X		X		X	X	X		X	X
Confiabilidade	Confiabilidade do modelo	X	X		X		X	X		X	X	X				
Características da tarefa	Nível de complexidade de tarefas realizadas por IA	X	X	X			X	X	X	X	X	X	X	X	X	X
Comportamentos imediatos	Capacidade de resposta, adaptabilidade e comportamentos pró-sociais		X			X		X	X		X	X	X			X
Antropomorfismo	Similaridade de IA com humano	X						X				X	X			X

FONTE: A autora (2022).

A FIGURA 8 abaixo apresenta a distribuição de artigos por dimensão em ordem decrescente de popularidade e aponta que os fatores cognitivos e emocionais relacionados à dimensão características da tarefa apresentou maior popularidade. Os artigos foram analisados e classificados em ordem decrescente considerando a influência dos fatores cognitivos e emocionais nas dimensões abordadas de forma mais substancial em cada trabalho. A partir desta análise foi possível verificar que os fatores que apresentaram maior popularidade estão nas dimensões de transparência presente em pelo menos (60%) dos artigos e Características da tarefa presente em (87%) dos artigos.

FIGURA 8 - DISTRIBUIÇÃO DE ARTIGOS POR DIMENSÃO



FONTE: A autora (2022).

4.3 DISCUSSÕES E PROPOSIÇÕES

A partir dos achados da revisão da literatura formularam-se algumas proposições que serão apresentadas nas seções seguintes.

4.3.1 Fatores cognitivos e emocionais que influenciam confiança humana em IA nas organizações

Neste estudo, Asan et al. (2020) identificam que a confiança humana em IA depende de algumas variáveis e fatores como: transparência, complexidade do modelo, controlabilidade, experiência do usuário e experiências anteriores, percepções, preconceitos, dentre outros. Além destes, Zhang et al. (2021) apresenta que os efeitos de IA nas percepções humanas a partir de vários fatores e perspectivas, incluindo, confiança/convicção, semelhança humana, justificação adequada e compreensibilidade/ explicações claras, informações sobre performance do modelo, transparência. Confiança em IA pode ser influenciada por fatores relacionados ao meio ambiente (colaboração da equipe, tarefa / contexto) e ainda podem ser relacionados a fatores humanos como (características, estados, fatores cognitivos e emocionais) (SHARAN; ROMANO, 2020).

A partir do QUADRO 5 identificou-se um artigo em potencial: *Human trust in artificial intelligence: Review of empirical research* de Glikson e Woolley (2020) cujo abrange todas as dimensões dos fatores cognitivos e emocionais identificados na RI, em um modelo desenvolvido pelas autoras.

Com o objetivo de suportar a pesquisa atual e futuras pesquisas, Glikson e Woolley (2020) desenvolveram um modelo que ajuda a orientar a integração de conhecimento sobre as relações humano-IA considerando que IA pode ser apresentada aos seus usuários de diferentes formas e os aspectos relacionados a confiança são distribuídos em diferentes áreas categorizado os principais fatores em dimensões que cobrem todos os demais identificados, são elas: tangibilidade, transparência, confiabilidade, características da tarefa, imediatismo (fatores cognitivos) e tangibilidade, antropomorfismo e imediatismo (fatores emocionais) (GLIKSON; WOOLLEY, 2020). Com isto, optou-se pela adoção deste modelo no estudo de casos.

○ Erro! Fonte de referência não encontrada.

QUADRO 6, abaixo, apresenta o modelo proposto por Glikson e Woolley (2020).

QUADRO 6 - MODELO DE GLIKSON E WOOLLEY

TIPOS DE IA	ROBO FÍSICO				AGENTES VIRTUAIS		EMBUTIDA	
	Fatores cognitivos	Fatores Emocionais	Fatores cognitivos	Fatores Emocionais	Fatores cognitivos	Fatores Emocionais		
Dimensões								
Tangibilidade	Presença física aumenta a confiança em IA como robô físico do que em IA virtual.	Presença física de IA pode não apenas aumentar a confiança mas também induzir medo.	Presença virtual de IA aumenta confiança em relação a IA embutida.	A "persona" aumenta os gostos por IA e a confiança emocional.	O efeito do uso de IA não é claro.	Desconhecer o uso da IA pode evocar raiva. Emoções positivas podem ser induzidas por boa reputação de uma empresa.		
Transparência	Transparência pode aumentar a confiança, mas pesquisas empíricas neste sentido são escassas.	N/A	Transparência na confiabilidade de IA e explicações sobre como o algoritmo funciona aumentam a confiança.	N/A	Transparência de como o algoritmo funciona aumenta a confiança, especialmente para sistemas de gerenciamento altamente inteligentes.	N/A		
Confiabilidade	Baixa confiabilidade reduz estímulos de confiança, na maioria das vezes. Em alguns casos, quando o robô é percebido com alta inteligência de máquina, as pessoas tendem a seguir até mesmo um robô defeituoso.	N/A	A baixa confiabilidade diminui principalmente a confiança em estudos de laboratório e de campo onde a confiança inicial era muito alta.	N/A	Baixa confiabilidade diminui significativamente a confiança e recuperá-la é difícil e leva tempo.	N/A		

Características da tarefa	A confiança é maior para realização de tarefas técnicas do que das que requerem inteligência social.	N/A	Confiança em IA é maior do que em humanos para tarefas que requerem análise de dados.	N/A	Confiança nos humanos é maior do que na IA para tarefas que requerem inteligência social; alta autoconfiança modera a confiança na IA.	N/A
Comportamentos imediatos	Capacidade de resposta, adaptabilidade e comportamentos pró-sociais aumentam a confiança.	Comportamentos humanos induzem alta confiança emocional; robôs que comentem erros são mais apreciados do que os impecáveis.	Utilização de táticas de persuasão e personificação aumentam a confiança.	A similaridade com humanos aumenta o gosto por IA e a confiança, mas ainda depende da predisposição dos usuários.	A personalização melhora a confiança; acompanhamento constante do comportamento dos funcionários pode diminuir a confiança.	referência não encontrada
Antropomorfismo	N/A	A similaridade com humanos geralmente aumenta emoções positivas, mas também pode causar desconforto.	N/A	Aumenta a confiança, mas também cria alta expectativa em relação às habilidades da IA. Atratividade e personalização, como etnia ou semelhança facial com o usuário, aumentam a confiança.	N/A	referência não encontrada

Fonte: Adaptado de GLIKSON; WOOLLEY (2020).

Nesta revisão de literatura é possível observar a relação e influência da forma de IA e dos fatores cognitivos e emocionais para promoção de confiança humana nesta tecnologia nas organizações. Assim, entende-se que investir no desenvolvimento organizacional focando no nestes fatores durante a implementação de IA oferece grande potencial para promoção de confiança em IA e por conseguinte gera benefícios quanto ao uso da tecnologia reduzindo resistência e aumentando a aderência dos usuários.

A confiança inicial em IA do tipo Robô físico é relativamente baixa portanto, fatores como confiabilidade, transparência e as características da tarefa podem desempenhar um papel importante no desenvolvimento da confiança, afirmam Glikson & Woolley, (2020, p. 632). Com isto, deduz-se que:

Proposição 1. Considerar os fatores cognitivos e emocionais durante a implementação de IA na forma de Robô Físico irá favorecer o aumento de confiança humana nesta tecnologia nas organizações.

Algumas características de IA do tipo agentes virtuais podem aumentar a confiança, tais como: Visualização, forma e antropomorfismo, principalmente por aumentar significativamente as dos usuários, enquanto o nível real de inteligência de máquina de IA modera a direção da trajetória de confiança (GLIKSON; WOOLLEY, 2020). Com isto, deduz-se que:

Proposição 2. Considerar os fatores cognitivos e emocionais durante a implementação de IA na forma de Agentes Virtuais irá favorecer o aumento de confiança humana nesta tecnologia nas organizações.

IA do tipo embutida é geralmente incorporada a sistemas, mapas, dispositivos, ou seja, é “invisível” para os usuários, não tem representação física ou visual. Diante disto, muitos usuários não estão cientes ou ainda não conseguem reconhecer a existência de IA. Supondo que os usuários estejam cientes o nível percebido de especialização ou inteligência de máquina também desempenha um papel importante na confiança cognitiva, bem como o tipo de tarefa envolvida. À medida que o nível de inteligência da máquina aumenta, os fatores contextuais e centrados no usuário tornam-se mais importantes para a confiança. Tornar a tecnologia mais compreensível

ênfatizando os benefícios atuais e futuros, estratégia para o uso de UA e usabilidade obtém mais êxito na promoção de confiança (GLIKSON; WOOLLEY, 2020). Com isto, deduz-se que:

Proposição 3. Considerar os fatores cognitivos e emocionais durante a implementação de IA na forma Embutida irá favorecer o aumento de confiança humana nesta tecnologia nas organizações.

5 ESTUDO DE CASOS MÚLTIPLOS: ALFA, BETA E GAMA

O estudo de caso contribui de forma consistente com a criação de novas teorias e é considerado um dos métodos de pesquisa mais comuns em gestão de operações e administração pois sua principal vantagem é o aumento da chance de vincular causa e efeito (VOSS, 2010).

Eisenhardt (1989, p. 536) destaca que “a escolha dos casos é um tópico importante na concepção da teoria de estudo de casos. Um estudo de caso é uma unidade de análise na pesquisa de caso. É possível usar diferentes casos da mesma empresa para estudar questões diferentes, ou pesquisar a mesma questão em uma variedade de contextos na mesma empresa ou de várias empresas (VOSS, 2010).

A pesquisa de caso é o método que utiliza estudos de caso como base e tem pelo menos 3 pontos fortes:

- Estudo do fenômeno em seu ambiente natural, na prática;
- As questões “por que” e “o quê” podem ter respostas mais completas; e,
- Permite estudar variáveis ainda desconhecidas de fenômenos não compreendidos. (VOSS, 2010).

Com objetivo validar o os achados na RI, optou-se pela realização de estudo de casos múltiplos. A escolha pelo estudo de casos múltiplos se justifica pela possibilidade de aumentar a validade externa ajudando a proteger contra tendência do observador. Além disso, o estudo de casos múltiplos é geralmente considerado mais confiável do que o caso único pois pode ajudar a reforçar os achados na forma de replicações de cada caso, obedecendo a lógica da replicação e não da amostragem onde as comparações entre os casos permitem maior robustez no desenvolvimento (YIN, 2015).

A fim de responder a segunda questão de pesquisa “**Quais são os fatores (cognitivos e emocionais) que afetam a confiança humana em IA nas empresas do setor automotivo?**” e ainda colocar em prática os achados da RI testando o modelo proposto por Glikson e Woolley (2020), as perguntas das entrevistas semiestruturadas foram baseadas neste modelo. As entrevistas foram gravadas e transcritas em sua totalidade com objetivo de garantir todos os detalhes.

Para este estudo, como já mencionado, o setor escolhido foi o automotivo por ser um setor promissor para uso de IA nos próximos anos, e, para tanto, foram

selecionadas 3 empresas deste setor. No **Erro! Fonte de referência não encontrada.** abaixo são apresentadas informações das 3 empresas.

QUADRO 7 - INFORMAÇÕES SOBRE AS EMPRESAS ESTUDADAS

	ALFA	BETA	GAMA
Faturamento	€ 12,5 bilhões	R\$ 3,8 bilhões	US\$ 33,684 bilhões
Produtos	Caminhões, carros e furgões	Rodas	Carros, minivans, caminhonetes
Tempo no mercado	123 anos	78 anos	114 anos
Mercado	Multinacional	Multinacional	Multinacional
País	França	Brasil	Estados Unidos

Fonte: A autora (2022).

5.1 RESULTADOS

Este capítulo tem por objetivo apresentar os resultados, análises e discussões dos dados coletados para o estudo. Em harmonia com a metodologia descrita anteriormente, foram feitas entrevistas semiestruturadas e as informações coletadas foram então submetidas a leitura minuciosa com objetivo de organizar e identificar padrões para melhor interpretação e a apresentação dos resultados. A construção das discussões podem ser verificadas na última etapa da análise, com vistas a identificar respostas às questões propostas nesta pesquisa.

Esta seção foi estruturada considerando o formato da entrevista semiestruturada, conectando as informações de todos os entrevistados das três empresas selecionadas no estudo de casos múltiplos. Para tanto, temos as subseções: Contexto geral, tipos de IA e Fatores cognitivos e emocionais que influenciam confiança humana em IA. A fim de preservar a identidade das empresas e entrevistados, as respostas foram indicadas relacionando entrevistado 1 como E1 , entrevistado 2 como E2 e assim sucessivamente conforme já apresentado no **Erro! Fonte de referência não encontrada.**

5.1.1 Resultados das entrevistas

Contexto Geral

Esta seção tem por objetivo compreender por que, quando e como as empresas optaram por adotar IA em seus processos produtivos e produtos.

Pergunta 1: Quando a empresa começou a adotar IA nos processos produtivos? E nos produtos?

Quando questionados em relação a adoção de IA nos processos produtivos e nos produtos, (100%) indicaram que a empresa ainda não adota IA nos produtos, apenas nos processos produtivos. Em relação a adoção nos processos produtivos, para todas as empresas esta é ainda uma iniciativa recente. A empresa Alfa relatou que adota desde 2020, Beta desde 2019 e Gama desde 2018.

Pergunta 2: A decisão inicial quanto a adoção de IA foi global ou local?

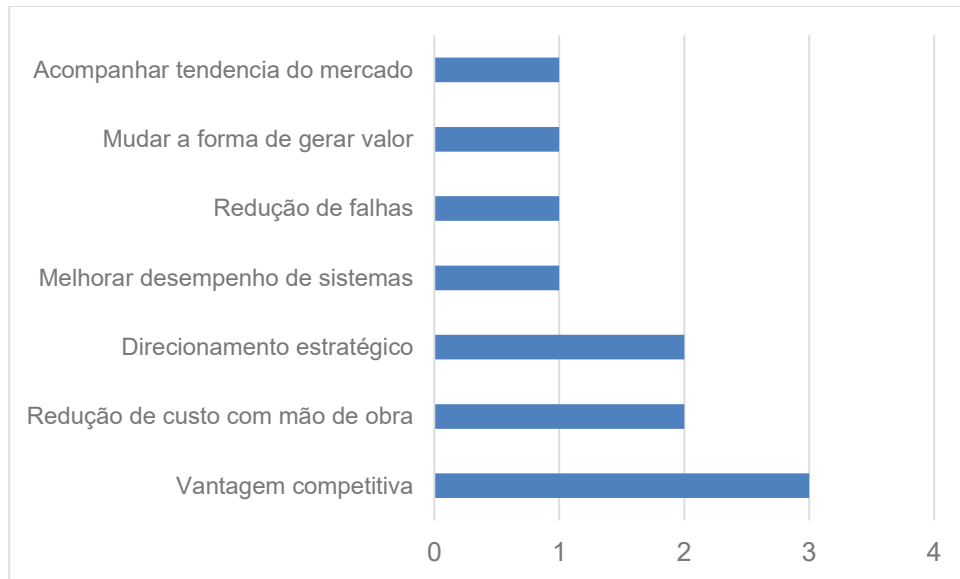
(100%) dos entrevistados responderam que foi uma iniciativa global.

Pergunta 3: Qual a relevância de IA para a empresa (alta, média, baixa)?

(71,5%) dos entrevistados responderam que a relevância é alta e (28,5%) dos entrevistados responderam que a relevância é média.

Pergunta 4: Quais os fatores internos e externos que estimularam a empresa a adotar IA nos seus processos produtivos e produtos? Descreva os desafios e oportunidades.

Os entrevistados relacionam fatores internos e externos que contribuíram para adoção de IA ressaltando que alguns entrevistados apontaram mais de um critério sendo vantagem competitiva, redução de custo com mão de obra e direcionamento estratégico os fatores mais comentados.



Fonte: A autora (2022).

“É claro que as empresas acabam uma assim, umas influenciando às outras, né? É uma questão fundamentalmente de se sustentar ao longo do tempo, então você precisa ter cada vez mais diferenciais competitivos e inteligência artificial acaba ajudando nesse sentido também” (E7).

“Eu acho que se a empresa parar no tempo e não fizer nada relacionado a novas tecnologias da transformação digital, inteligência artificial, ela vai acabar ficando para trás e o custo do seu do produto vai ficando mais caro. A tendência de uma empresa que não adotar inteligência artificial é sumir do mercado. É mais uma questão evolutiva estratégica” (E5).

Pergunta 5: Como foi a aceitação dos colaboradores em relação ao uso de IA nos processos produtivos (positiva, negativa ou neutra)?

(43%) dos respondentes consideraram positiva, (43%) neutra e (14%) negativa.

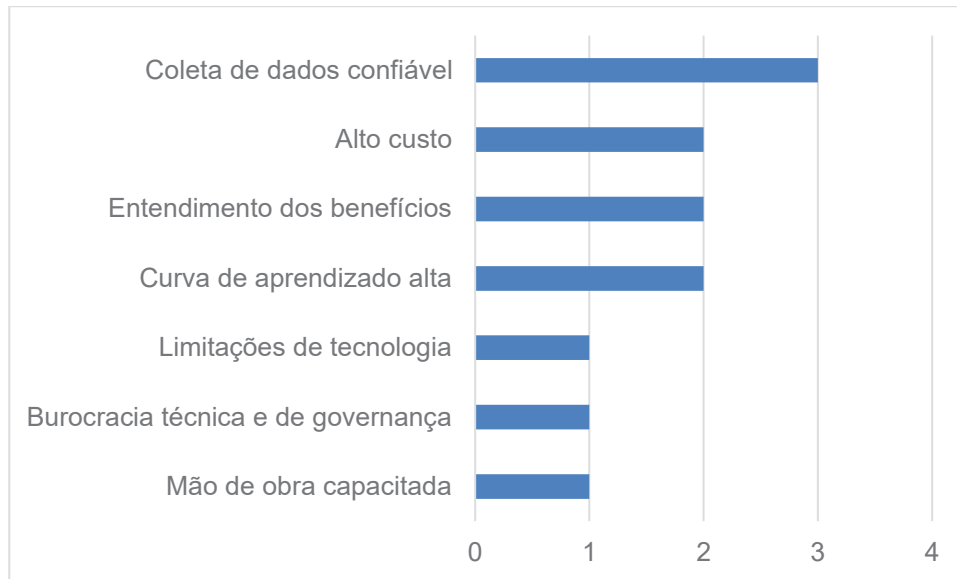
Pergunta 6: E como foi a aceitação dos clientes em relação ao uso de IA nos produtos?

Nenhuma das empresas respondeu que adota IA nos produtos (carros), apenas nos processos produtivos (fabricação).

Pergunta 7: Quais os principais desafios identificados em relação a adoção de IA nos processos produtivos? E nos produtos?

Os entrevistados relacionam como principais desafios a coleta de dados confiável, alto custo, entendimento dos benefícios e curva de aprendizado alta.

FIGURA 10 - DESAFIOS NA ADOÇÃO DE IA



Fonte: A autora (2022).

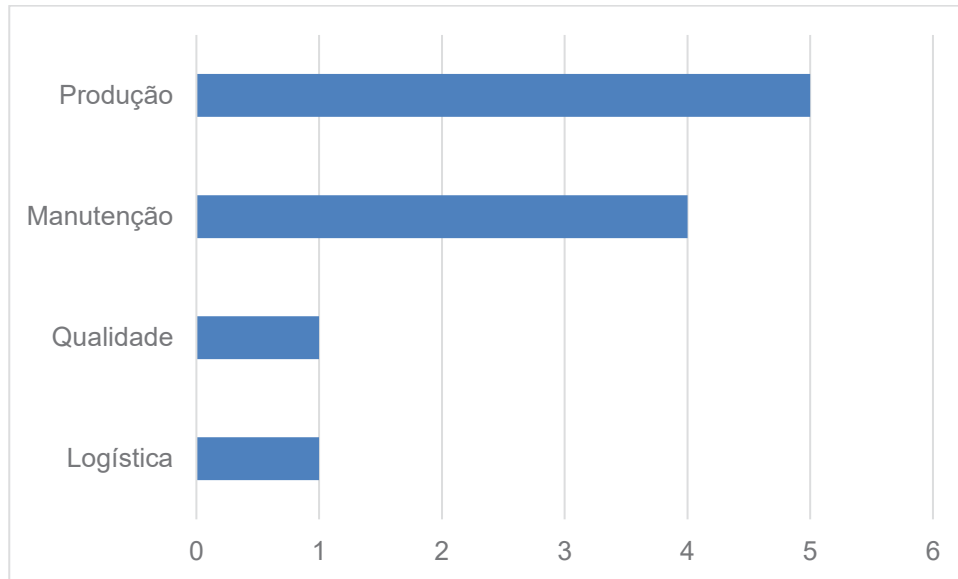
“Um outro desafio que eu vejo que faz parte é o da curva de aprendizado. Porque quando a gente está fazendo algo, tem algumas que participam e outras não. Elas entram sem conhecimento do tema, têm que se desenvolver e acabam levando mais tempo para ter esse desenvolvimento” (E2).

“E essa parte da coleta de dados é um grande desafio também, porque as fábricas que nós temos, elas são muito diferentes. As máquinas, as linhas são totalmente diferentes e a conexão conectividade que a gente tem nas plantas também. Então, só para dar um exemplo, tem máquina que tem mais de 100 anos em algumas plantas” (E4).

Pergunta 8: Quais áreas da empresa foram mais afetadas pela adoção do uso de IA nos processos produtivos? E nos produtos? (Enumere-as como operacionais, gerenciais/administrativas, organizacionais, intraorganizacionais). Destaque como elas foram afetadas?

Os entrevistados comentaram as áreas que foram mais afetadas pela adoção de IA. Importante ressaltar que alguns entrevistados apontaram mais de um critério sendo produção a área mais impactada.

FIGURA 11 - AREAS MAIS AFETADAS POR IA



Fonte: A autora (2022).

“O departamento que vai ter impacto é o departamento de manutenção. A maioria dos nossos projetos utilizando inteligência artificial é da parte de *analytics* para manutenção preditiva” (E5).

“A funilaria é uma área que tem altíssimo nível de automação, a altíssimo nível de integração, coleta de dados e todo um trabalho preditivo justamente para gerar inteligência e evitar problemas no processo produtivo através de manutenção preditiva (E7).

Pergunta 9: Qual é a percepção dos funcionários sobre a adoção de IA nos processos produtivos da empresa? (Desconfiança ou confiança)? E como você espera que esta tecnologia venha beneficiar as pessoas na empresa?

(71%) dos entrevistados comentaram que a percepção dos funcionários é de confiança e (29%) comentaram que a percepção é de desconfiança.

“Tem expectativa que novas soluções vão ajudar a resolver problemas” (E1).

“Esperam que vão ter benefícios” (E4).

“As pessoas veem como algo de tecnologia de ponta e se interessam, oportunidade de carreira” (E7).

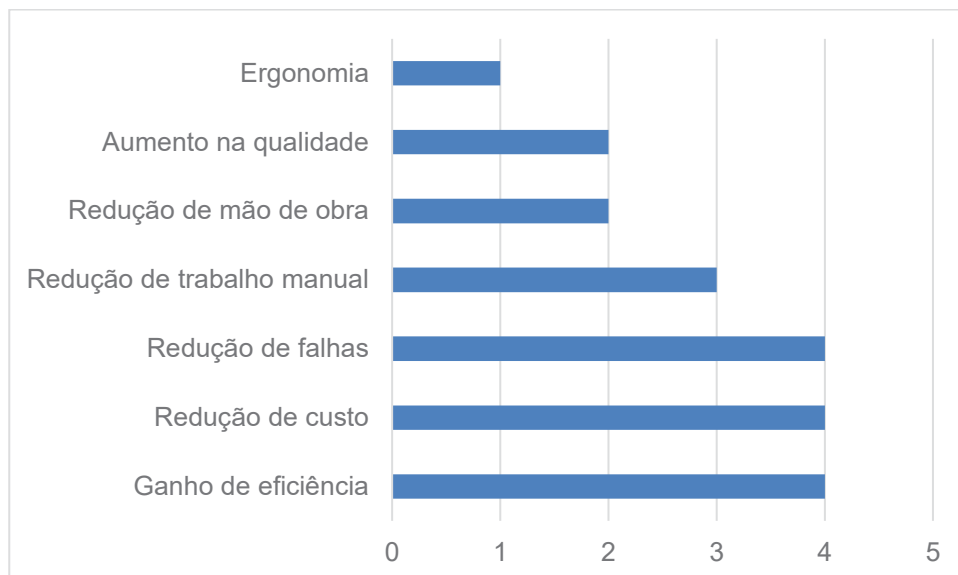
Pergunta 10: Qual é a percepção dos clientes sobre a adoção de IA nos produtos (Carros) da empresa? (Desconfiança ou confiança)?

Nenhuma das empresas relatou ter IA sendo adotada nos carros.

Pergunta 11: Quais benefícios puderam ser observados em relação ao uso de IA nos processos produtivos? (Exemplo: otimização do tempo de produção, redução de falhas/retrabalho, maior qualidade, aumento da transparência, redução de custo operacional, mais eficiência etc.)

Os três principais benefícios comentados pelos entrevistados foram: Redução de falhas, redução de custo e ganho de eficiência.

FIGURA 12 - BENEFICIOS DO USO E IA



Fonte: A autora (2022).

Pergunta 12: Quais benefícios puderam ser observados em relação ao uso de IA nos produtos (carros)? (Exemplo: maior aceitação pelos clientes, aumento nas vendas, diferencial competitivo, maior usabilidade etc.).

Nenhuma das empresas respondeu que adota IA nos produtos (carros), apenas nos processos produtivos (fabricação).

Pergunta 13: Quais os principais desafios identificados durante a implementação de IA nos processos produtivos no nível interno e externo? (Exemplo: alto custo da tecnologia, falta de mão de obra capacitada, ausência de políticas regulatórias para uso, operacionalização etc.).

Os entrevistados identificaram os principais desafios durante a implementação de IA. Importante ressaltar que alguns entrevistados apontaram mais de um critério sendo o alto custo com tecnologia o mais comentado.

FIGURA 13 - DESAFIOS DURANTE A IMPLEMENTAÇÃO DE IA



Fonte: A autora (2022).

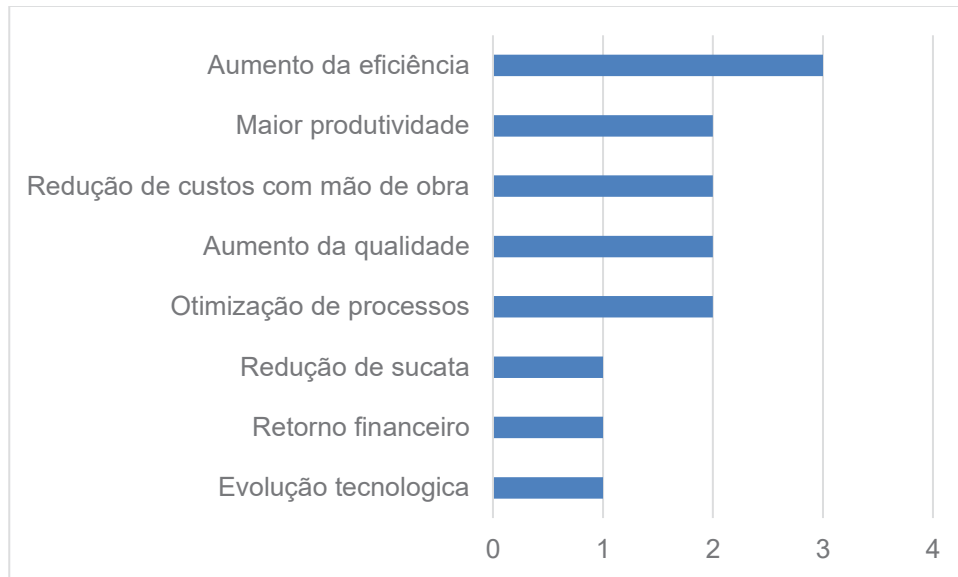
Pergunta 14: A Lei Geral de Proteção de Dados (LGPD) ou *General Data Proctetion Regulation* (GDPR) afetou de alguma forma o uso de IA nos processos produtivos ou produtos da empresa? Em caso afirmativo, quais medidas foram tomadas?

Os entrevistados comentaram que a LGPD não afetou a implementação de IA, mas ela sempre deve ser consultada.

Pergunta 15: Quais os resultados esperados em relação ao uso de IA nos processos produtivos no curto, médio e longo prazo? E nos produtos? Detalhe os resultados já alcançados e os encaminhados.

O aumento de eficiência é um dos principais resultados esperados segundo o comentário dos entrevistados.

FIGURA 14 - RESULTADOS ESPERADOS COM ADOÇÃO DE IA



Fonte: A autora (2022).

“Para mostrar aos nossos clientes, e para os investidores da empresa que nós estamos evoluindo. Por ser um ambiente automotivo, que é ainda considerado bem conservador, essa utilização mostra para o mundo que nós estamos tentando fazer algo para melhorar, utilizando tecnologias novas, também traz benefícios para a empresa” (E4).

Tipos de IA

Glikson e Woolley (2020) definem abaixo os principais tipos de IA.

- Robô físico: representações mecânicas físicas com similaridade humana contendo uma gama de funções e capacidades, que podem estar remotamente ou fisicamente presentes executando tarefas para as quais foram programados.
- Agentes virtuais: representação inteligente com características humanas sem presença física, uma identidade distinta como *chat bot* ou avatar). Exemplo: Lu, o *chat bot* premiado do Magazine Luiza.
- Embutida: é uma forma de IA invisível para os humanos, não possui representação visual e por ser totalmente incorporada a sistemas e pode ser incorporada a produtos e serviços tais como mecanismos de busca, programas e aplicativos."

Pergunta 1: Considerando os tipos de IA descritos acima (Robô Físico, Agentes Virtuais e Embutida), quais são os tipos de IA mais utilizados nos processos produtivos da empresa?

(100%) dos entrevistados responderam que IA está sendo ou foi implementada na forma de IA embutida/*bot*. Atualmente os robôs físicos existentes são braços robóticos utilizados como automação de processos, sem inteligência, aprendizagem ou ainda forma humana. Dois entrevistados comentaram ainda que a empresa possui agentes virtuais no setor de tecnologia, mais focado no suporte a usuários.

Pergunta 2: Quais os tipos de IA mais utilizados nos produtos? Justifique os critérios de escolha de cada tipo.

Nenhuma das empresas respondeu que adota IA nos produtos (carros), apenas nos processos produtivos (fabricação).

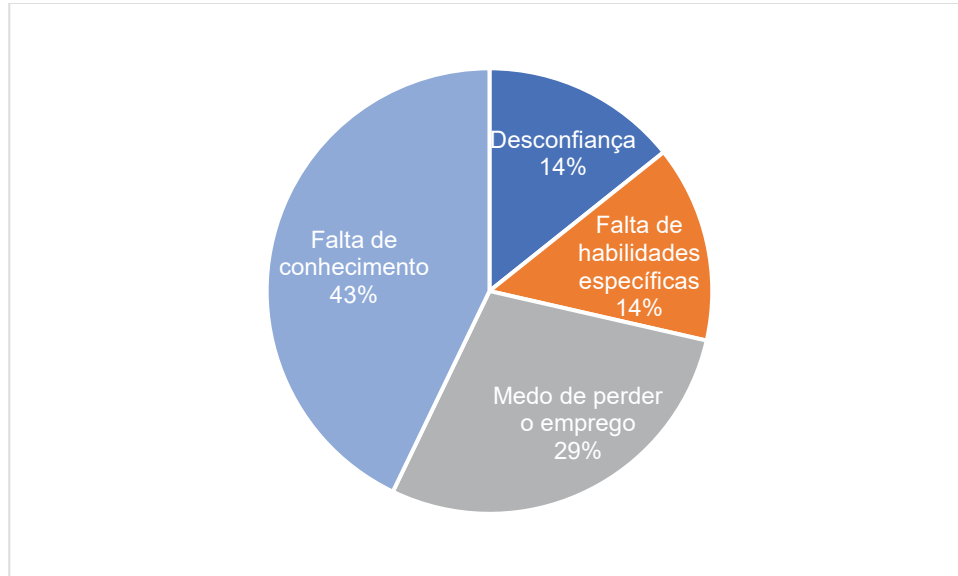
Pergunta 3: IA foi utilizada de forma pura ou em conjunto com outras tecnologias (Exemplo: *IoT*, *Big Data*, *Cloud* etc.)?

(100%) dos entrevistados responderam que IA foi ou está sendo implementada em conjunto com outras tecnológicas, principalmente *IoT*, *Big Data*, *Cloud*.

Pergunta 4: Quais os principais desafios identificados durante a implementação de cada tipo de IA nos processos produtivos em relação aos colaboradores usuários da tecnologia que atuam no dia a dia (desconfiança, medo de perder o emprego, resistência, baixa utilização, falta de conhecimento, necessidade de desenvolver habilidades específicas etc.).

(43%) dos entrevistados destacaram a falta de conhecimento como um dos principais desafios ao implementar IA por ser uma tecnologia muito nova.

FIGURA 15 - PRINCIPAIS DESAFIOS DURANTE A IMPLEMENTAÇÃO DE IA



Fonte: A autora (2022).

“Mas o que eu ouço é, senhor, os funcionários, eles têm receio de perder emprego por causa dessas automatizações” (E1).

Pergunta 5: Quais os principais desafios identificados durante a implementação de cada tipo de IA nos produtos (Carros) em relação aos clientes ou consumidores finais que adquirem veículos com IA como tecnologia embarcada? (Exemplo: alto custo da tecnologia, desconfiança dos usuários, baixa utilização etc.);

Nenhuma das empresas respondeu que adota IA nos produtos (carros), apenas nos processos produtivos (fabricação).

Fatores cognitivos e emocionais que influenciam confiança humana em IA.

Glikson e Woolley (2020) consideram que confiança é uma construção e pode ser influenciada por fatores cognitivos (avaliação racional e características situacionais) ou emocionais (baseado no afeto, conexões sociais), neste sentido, as autoras apresentam como tais fatores influenciam o grau de confiança em cada tipo de IA (Robô físico, Agentes virtuais e embutida) propõem em seu modelo as dimensões de confiança cognitiva (tangibilidade, transparência, confiabilidade, características da tarefa e comportamentos imediatos) e emocional (tangibilidade comportamentos imediatos e Antropomorfismo)."

Nesta seção o objetivo é identificar quais os fatores cognitivos e emocionais que influenciam confiança humana em IA de acordo com os tipos e IA que cada empresa te implantado.

Durante as entrevistas identificou-se que alguns entrevistados comentaram sobre os tipos Robô físico e Agentes virtuais, contudo, fora do contexto deste estudo e, para tanto, foram desconsiderados.

Pergunta 1: IA Embutida: Atualmente existe alguma forma de IA embutida sendo utilizados nos processos produtivos e produtos da empresa? Quais são?
Detalhar

Quando questionados sobre os tipos de IA adotado na empresa, (100%) dos entrevistados se referiram a IA embutida.

Pergunta 2: Dimensão tangibilidade. Desconhecer que IA está sendo utilizada pode evocar sentimentos negativos, desconforto, desconfiança, raiva. Porém, emoções positivas podem ser induzidas por boa reputação de uma empresa. Na sua opinião a reputação é um fator relevante para a adoção de IA?

(85%) dos entrevistados concordam que uma boa reputação da empresa é relevante para evocar confiança.

Pergunta 3: Dimensão Transparência. Compreender com transparência como o algoritmo funciona, por exemplo, pode aumentar a confiança em IA, especialmente para sistemas de gerenciamento altamente inteligentes. Vocês consideram ou consideraram este fator ao implementar IA?

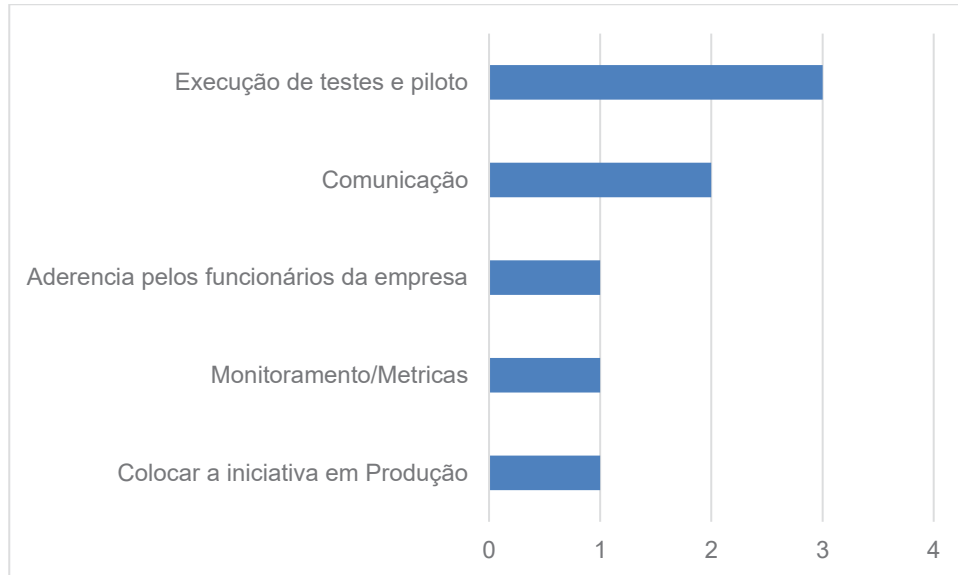
(100%) dos entrevistados concordam que transparência é essencial ao implementar novas tecnologias como IA.

Pergunta 4: Dimensão Confiabilidade. Baixa confiabilidade diminui significativamente a confiança e recuperá-la é difícil e leva tempo. Como a empresa busca mitigar esta situação? Existem ações mapeadas? Quais?

Quando questionados em relação a confiabilidade, (100%) concordam que este é um fator relevante e que impacta diretamente a confiança. Algumas das ações

que as empresas tomam para mitigar este fato estão relacionadas abaixo, sendo que a execução de testes e piloto é a mais comentada.

FIGURA 16 - AÇÕES PARA GERAR CONFIABILIDADE

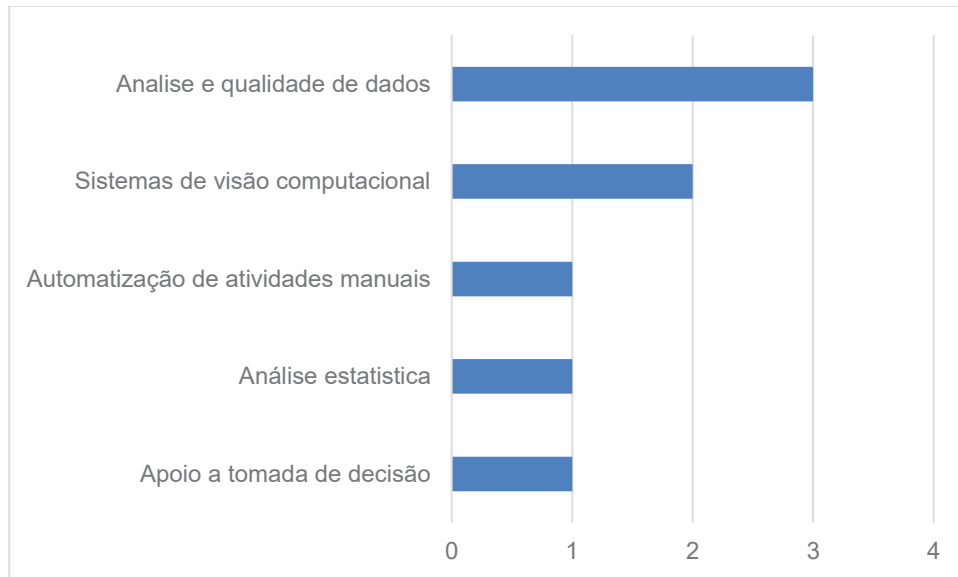


Fonte: A autora (2022).

Pergunta 5: Dimensão características da tarefa. Tarefas que requerem inteligência social tem maior aceitação e confiança em humanos do que em IA, porém, alta autoconfiança modera a confiança na IA, principalmente para tarefas especializadas. Na sua opinião, onde IA embutida tem sido mais utilizada pela empresa? Quais tipos de tarefa?

Os entrevistados apontaram tarefas de análise e qualidade de dados como as que tem sido mais utilizadas pela empresa.

FIGURA 17 - CARACTERÍSTICAS DA TAREFA MAIS COMUNS



Fonte: A autora (2022).

Pergunta 6: Dimensão comportamentos imediatos. A personalização melhora a confiança, porém, a sensação de acompanhamento/monitoramento constante do comportamento dos usuários de IA pode diminuir a confiança. A empresa busca promover alguma ação que ajude a mitigar estes casos?

Enquanto questionados (57%) dos entrevistados concordaram com a afirmação e relataram ações que a empresa toma para mitigar estes casos, já (43%) reportaram que a empresa não toma quaisquer ações.

“Em geral as pessoas têm receio de IA detectar erro no que estão executando” (E1).

“Todas as pessoas têm acesso a plataforma e acesso a esses dados” (E2).

“Executando treinamento, agindo com transparência” (E6).

5.2 ANÁLISES

Nesta seção, são analisados os dados identificados neste estudo com base nos elementos conceituais da teoria, em resposta aos objetivos específicos apresentados no capítulo 1.

Conforme apresentado no QUADRO 8 abaixo, em relação aos tipos de IA e aos fatores cognitivos e emocionais que influenciam confiança humana em IA no setor automotivo, através do estudo de casos foi possível identificar que atualmente nas

empresas estudadas existe apenas IA na forma Embutida implantada, ainda não existe IA do tipo Robô físico e Virtual.

QUADRO 8 - TIPOS DE IA POR EMPRESA

Empresa	Robô Físico	Agentes virtuais	Embutida
Alfa	Não possui IA na forma de Robô físico	Não possui IA na forma de Agentes virtuais	Possui IA na forma embutida
Beta	Não possui IA na forma de Robô físico	Não possui IA na forma de Agentes virtuais	Possui IA na forma embutida
Gama	Não possui IA na forma de Robô físico	Não possui IA na forma de Agentes virtuais	Possui IA na forma embutida

Fonte: A autora (2022).

Quanto aos fatores cognitivos e emocionais, os mais populares no modelo de Glikson e Woolley (2020), que são os das dimensões de transparência e características da tarefa por serem as mais comentadas ou com maior quantidade de ações sendo realizadas nas empresas do setor automotivo foram destacados. Inicialmente a confiança em IA é relativamente baixa, com isto, fatores como transparência e características da tarefa podem desempenhar um papel crucial no desenvolvimento de confiança GLIKSON; WOOLLEY (2020).

Os principais comentários dos entrevistados em relação às dimensões e fatores que influenciam confiança humana e IA são apresentados no QUADRO 9 a seguir.

QUADRO 9 - SÍNTESE DOS FATORES IDENTIFICADOS NO ESTUDO DE CASOS

DIMENSÕES	FATORES DA RI	EMPRESA ALFA	EMPRESA BETA	EMPRESA GAMA	EVIDÊNCIAS DOS FATORES NO EC
Características da tarefa	Nível de complexidade de tarefas realizadas por IA	Fator presente	Fator presente	Fator presente	"Tarefas complexas como análise de dados podem gerar mais confiança: "o pessoal tenta conferir, ver se está batendo com a realidade dos dados ali que ela está levantando. Então tem uma desconfiança no começo, mas depois melhora" (E1). "Como as tarefas são uns bem mais complexas, é um sentimento de que IA veio pra dar uma

				<p>turbinada nas nossas decisões e não e não pra roubar o lugar" (E2).</p> <p>"Tem muito mais confiança em IA nas questões técnicas" (E7).</p>
Transparência	Visibilidade e entendimento de como IA está sendo implantada	Fator presente	Fator presente	<p>Fator presente</p> <p>"Sim, é isso, é o nosso perfil da nossa empresa. A gente sempre tenta ser o máximo possível transparente com todos e naturalmente, a gente vê isso não só programa, mas nas ações como um todo, todas as funções. A gente procura ser sempre transparente, sempre. Então a gente a gente fala, a gente faz, a gente é procura expor o que que está acontecendo" (E1).</p> <p>"A gente sempre organiza a reuniões, envolve todo mundo dos setores para acompanhar o desenvolvimento. Então assim eles sabem que que está rodando por trás. Não é só uma caixa preta para ele" (E6).</p> <p>"Fazer as pessoas chegarem à correlação semelhante aqui. A gente tenta provocar sempre nesse sentido de pensamento racional, baseado em dados. E isso a gente é busca sempre em treinamentos e na transparência" (E6).</p> <p>"Como eu trabalho com desenvolvimento e eu trabalho com isso, isso é fundamental, você tem que explicar. E você tem que explicar também o que e como é</p>

					que funciona e tem que explicar as limitações. Se você não explica isso, é uma crença que é IA é um negócio mágico e vai resolver seu problema. E aí ela tem limitações, tem erro, tem acertos" (E3).
Tangibilidade	Presença ou representação humana física/virtual	Fator não presente	Fator não presente	Fator não presente	<p>"Eu acho que o nível de confiança entre IA embutida e Agentes virtuais é igual, porque na gente virtual, por mais que ela tem aquela, aquela figura e o nome tipo, a gente também não sabe que tá rolando por trás" [E1].</p> <p>"Existe os colaboradores que têm algum receio. Aí eu não sei dizer se é receio, porque não acreditam que vai ter um benefício ou muitas vezes até por medo. Muita gente ainda pensa que a inteligência artificial vai acabar com todo o trabalho, principalmente quando a gente vai para chão de fábrica" (E4).</p> <p>"Eu, por ter estudado um pouco de IA na especialização, tenho noção do que roda por trás. Mas muita gente, na hora que você fala inteligência artificial, pensa nos filmes do humanoide aprendendo e tudo mais, mas não tem noção exatamente do que é o que tem por trás. Conhecer isso muitas pessoas se fascinam já outras não se interessam em conhecer, saber o que que é" (E5).</p>
Confiabilidade	Confiabilidade do modelo	Fator presente	Fator presente	Fator presente	"Então eu acho que a confiabilidade no trabalho do time que faz a implementação e a confiabilidade do projeto, de como foi

					<p>feito são muito importantes. Todas as aplicações que a gente desenvolveu dentro das plantas tem esse trabalho em conjunto. Eles fazem parte, eles são os maiores interessados também fazer acontecer e fazer acontecer do da melhor forma possível" (E5).</p> <p>"com a iniciativa lá em produção alguns passos que têm que ter ali que você tenha garantia. Por exemplo, monitoramento da se continua performando, análise de dados e definir as métricas para garantir que ainda continua confiável" (E2).</p>
Comportamentos imediatos	Capacidade de resposta, adaptabilidade e comportamentos pró-sociais	Fator presente	Fator presente	Fator presente	<p>"A empresa é muito desconectada. Então quando nós queremos tomar alguma ação por conta da qualidade de ser alta para tomar alguma ação, a gente tem que ter acesso a esses dados. Essa massa de dados é muito complexa, então produção e engenharia também entram nesse bolo e são beneficiadas" (E4).</p> <p>"Em geral as pessoas têm receio de IA detectar erro no que estão executando" (E1).</p>
Antropomorfismo	Similaridade de IA com humano	Fator não presente	Fator não presente	Fator não presente	Não foram identificados tipos de IA com similaridade humana sendo utilizadas nas empresas estudadas.

Fonte: A autora (2022).

Conhecer os benefícios e desafios na implementação de IA pode contribuir para ampliar as chances de sucesso da implementação, bem como identificar e focar em estratégias específicas de acordo com a necessidade de cada organização. Para

tanto, após análises do conteúdo das entrevistas foram formulados os quadros 10 e 11 a seguir.

O QUADRO 10 a seguir apresenta os principais benefícios na implementação de IA.

QUADRO 10 – BENEFÍCIOS NA IMPLEMENTAÇÃO DE IA

BENEFÍCIOS	EMPRESA ALFA	EMPRESA BETA	EMPRESA GAMA	TRECHOS DAS ENTREVISTAS
Ganho de eficiência	X	X		"Ainda acho ela espera mesmo é que IA contribua com o aumento da eficiência da companhia" (E3)
Redução de custo	X	X		"É a redução de custo mesmo" (E2). É redução de custo até em alguns pontos a gente viu" (E4).
Redução de falhas	X	X	X	"Redução de falhas com certeza é" (E2) "Justamente na manutenção preditiva de falhas, uso máquinas na questão de prever potenciais problemas de qualidade. A então acho que o foco tem sido muito nesse lado" (E7).
Redução de trabalho manual	X	X		"Trabalhos massivos foram feitos de uma forma é" (E2).
Redução de mão de obra	X			"reduziu falha é dispensou alguns operadores, né? Porque o custo era muito alto para" (E1).
Aumento na qualidade		X	X	É otimização de tempo, com certeza aumento de qualidade de produtividade.
Ergonomia		X		"Pensar um pouquinho as automações que a gente fez, basicamente ergonomia. É um dos

				pontos principais pois tem bastante trabalho repetitivo, que foi eliminado" (E6).
--	--	--	--	---

Fonte: A autora (2022).

O QUADRO 11 sintetiza os principais desafios identificados na implementação de IA nas empresas.

QUADRO 11 - DEDAFIOS NA IMPLEMENTAÇÃO DE IA

DESAFIOS	EMPRESA ALFA	EMPRESA BETA	EMPRESA GAMA	TRECHOS DAS ENTREVISTAS
Mão de obra capacitada	X			"Além da falta de mão de obra que eu já comentei" (E1).
Burocracia técnica e de governança	X		X	"Tem tanta burocracia técnica quanto burocracias de governança" (E2)
Limitações de tecnologia	X			"Limitações da tecnologia às vezes lava a ter que refazer os processos" (E2)
Curva de aprendizado alta	X		X	"Um outro desafio que eu vejo é a curva de aprendizado" (E3).
Entendimento dos benefícios	X		X	"Você não coloca IA ali no meio e acha que aquilo vai resolver todos os problemas?" (E3).
Alto custo		X	X	"Claro que tem a questão do custo da tecnologia que nós já falamos" (E6) "Altos custos, sim, porque muitas ferramentas é que é até aqui. A gente está usando hoje o custo não é baixo, então a gente realmente tem que provar que aquilo funciona para avaliar o custo" (E4) "Também tenho custos de adoção de novas tecnologias"(E7)
Coleta de dados confiável	X	X	X	"A gente está lidando com conjuntos de dados bem grandes" (E1) "Análise de grande massa de dados não é mais fácil as pessoas acreditarem em IA do que não humano, por exemplo" (E2). "Essa massa de dados é muito complexo" (E5)

Dificuldade de integração de IA com sistemas Legados	X			"Um outro problema que eu estou observando é que os equipamentos que existem hoje em dia eles são com apostas, eles são. É software fechado, né? Não é software aberto, então, pra eu conseguir integrar a inteligência artificial com esses equipamentos é meio complicado" (E1)
Ausência de políticas regulatórias	X			Ausência de políticas regulatórias, mas não é bem uma ausência. Hoje, dentro da dentro da companhia elas estão mudando bastante, numa fase que está se adequando ainda ao uso. Então é a cada projeto novo tem alguma coisa diferente, e gente tem que descobrir como resolver" (E3).
Refazer processos	X	X		"A gente teve que desmanchar essas automações e refazer depois" (E5). "Então, precisa refazer os processos" (E1).
Tempo para implementar longo		X		"Acho que o tempo é um grande desafio" (E4).
Infraestrutura		X		"É ausência de infraestrutura" (E6)
Definição de estratégia clara		X		"É uma decisão estratégica e as empresas têm que tomar" (E5).

Fonte: A autora (2022).

Por conseguinte, pode-se compreender quais os principais fatores cognitivos e emocionais que influenciam confiança humana em IA e que o estudo de caso está alinhado com os benefícios de se considerar estes fatores durante a implementação e uso de IA nas organizações para aumentar as chances de sucesso. Por ser uma tecnologia nova e ainda pouco utilizada são muitos os desafios na implementação, porém o fortalecimento da confiança em IA por meio da educação, treinamentos e da transparência tendem a gerar bons resultados (LEWIS; MARSH, 2022).

5.3 DISCUSSÕES E PROPOSIÇÕES

Nesta seção, apresentam-se discussões que relacionam os resultados e as análises da RI ao estudo de caso.

5.3.1 Atendimento aos quadros de análise

A partir na Revisão Integrativa (RI) e análise de conteúdo dos 15 artigos selecionados, foram elaborados os Quadros 4 e 5 já apresentados em capítulo anterior sobre a RI. O Quadro 4 apresentou a evolução dos conceitos de confiança, inteligência artificial e fatores cognitivos e emocionais encontrados na RI. O quadro 5 apresentou a síntese dos fatores/dimensões identificados na literatura por autor. No estudo de caso foram investigados quais fatores/dimensões estavam presentes nas empresas sendo identificados como presentes os das dimensões: Características da tarefa, transparência, confiabilidade e comportamentos imediatos. Já as os fatores das dimensões tangibilidade e antropomorfismo não foram identificados.

5.3.2 Evidências sobre as proposições

Proposição 1. Considerar os fatores cognitivos e emocionais durante a implementação de IA na forma de Robô Físico irá favorecer o aumento de confiança humana nesta tecnologia nas organizações.

Conforme apresentado no Quadro 8, não foi possível identificar IA na forma de Robô físico em nenhuma das empresas estudadas.

Proposição 2. Considerar os fatores cognitivos e emocionais durante a implementação de IA na forma de Agentes virtuais irá favorecer o aumento de confiança humana nesta tecnologia nas organizações.

Conforme apresentado no Quadro 8, não foi possível identificar IA na forma de Agentes Virtuais em nenhuma das empresas estudadas.

Proposição 3. Considerar os fatores cognitivos e emocionais durante a implementação de IA na forma Embutida irá favorecer o aumento de confiança humana nesta tecnologia nas organizações.

Com a análise de conteúdo das entrevistas, triangulação dos dados encontrados e observação direta, pode-se compreender que atividades relacionadas a promoção de atividades envolvendo os fatores cognitivos e emocionais das

aumentam a confiança em IA, principalmente os fatores das dimensões transparência e características da tarefa, conforme apresentado no quadro 9.

Transparência tem um aspecto importante que inclui distintas explicações sobre como IA funciona ou porque determinada decisão foi tomada. Tornar isto compreensível para os usuários é importante mesmo quando eles têm pouco conhecimento técnico, pois os faz sentir parte da iniciativa minimizando a sensação de que IA irá “roubar” seu emprego ou fazer tudo sozinho (GLIKSON; WOOLLEY, 2020).

Conforme resultados apresentados no QUADRO 9, transparência é um fator presente em todas as empresas estudadas sendo considerado muito importante pela característica de envolver os usuários e gerar engajamento através de comunicação clara, reuniões de alinhamento e treinamentos.

Dessa forma, entende-se que quanto maior a transparência durante a implementação e uso de IA, maior o potencial de desenvolver confiança dos usuários e promover aderência ao uso.

6 CONSIDERAÇÕES FINAIS

Neste estudo, foram identificados alguns dos principais fatores cognitivos e emocionais que influenciam confiança humana em IA em distintas dimensões nas empresas do setor automotivo. Conhecer estes fatores é fundamental para realçar as oportunidades que existem para o desenvolvimento de confiança em IA, uma vez que esta tecnologia está se desenvolvendo muito rapidamente, possivelmente de forma exponencial, e como as pessoas vão se ajustar a isso é ainda um problema fascinante, comenta Kahneman (2021).

Promover confiança humana na utilização de IA é um fator chave para o sucesso da implantação desta tecnologia, principalmente porque os humanos temem serem substituídos pelas máquinas e este temor gera resistência criando barreiras no avanço desta tecnologia, porém, suas fortalezas combinadas elevam o potencial de aprendizagem melhorando suas capacidades e não o contrário (CURIONI, 2018).

Toda mudança requer preparação e passa por um período de adaptação, e neste sentido Glikson e Woolley (2020, P.633) destacam que “pesquisas existentes abordando a trajetória do ser humano e confiança em IA sugerem que a confiança inicial começa em um nível baixo e se desenvolve ao longo do tempo”. Com isto, conclui-se que, conhecer os fatores cognitivos e emocionais que influenciam confiança humana em IA oferece grande potencial para o desenvolvimento de confiança em IA nas organizações, favorecendo a aplicação e uso.

A partir da revisão integrativa, observou-se a necessidade de realização de estudos mais aprofundados sobre a influência dos fatores cognitivos e emocionais em relação a confiança humana em IA em função do baixo número de trabalhos encontrados na literatura.

Com o estudo de caso, foi possível compreender na prática como as organizações do setor automotivo tem lidado com confiança em IA e quais os fatores cognitivos e emocionais que mais influenciam e transparência foi a dimensão com maior popularidade

Recomenda-se que pesquisas futuras analisem também os fatores cognitivos e emocionais sob a perspectiva de longo prazo ao examinar o efeito da capacidade de resposta da IA a diferentes comportamentos de desuso e abuso no

desenvolvimento da confiança e na relação de trabalho humano-IA (GLIKSON; WOOLLEY, 2020).

Por fim, sugerem-se a condução de pesquisas envolvendo casos práticos nas organizações a fim de comparar os achados na literatura com o ambiente da vida real e como organizações que já usam IA em seus sistemas de gestão ou tomada de decisão lidam com confiança humano-IA.

6.1 CONTRIBUIÇÕES

A principal contribuição deste estudo é fornecer uma melhor compreensão dos fatores cognitivos e emocionais que afetam a confiança humana em IA e como empresas do setor automotivo podem enfrentar o desafio da complexidade intrínseca destes fatores. A partir dessa compreensão será possível gerar ações para melhorar a estratégia destas organizações no sentido de estarem preparadas para lidar com confiança humana em IA no setor automotivo.

Ajudar na identificação dos fatores cognitivos e emocionais que afetam a confiança humana em IA nos processos produtivos e produtos de empresas do setor automotivo de forma a promover ações para melhorar a percepção dos usuários, aumentar a resposta ao uso e fomentar a aquisição de produtos com IA por consumidores;

Contribuir com a literatura já existente sobre confiança e IA no setor automotivo. Futuros pesquisadores podem levar em consideração este estudo em outros setores;

Colaborar com o aprendizado contínuo requerido para o setor de tecnologia em função da complexidade e dinamismo advindo das constantes mudanças e necessidade de atualização.

6.2 IMPLICAÇÕES PARA GESTÃO E DECISÃO

Chong et al. (2022) destacam que a “confiança em IA é formada a partir da percepção de quem confia sobre a capacidade do objeto da confiança (neste caso, uma IA) de realizar uma determinada tarefa”. Compreender a dinâmica da confiança humana durante a tomada de decisão assistida por IA contribui para que empresas

do setor automotivo possam lidar melhor com os aspectos de confiança humana intrínsecos da utilização de IA em seu processo produtivo e produtos de forma a ajudar estas empresas se prepararem melhor para transformação digital no presente e acelerar o tempo de resposta a novos eventos no futuro (CHONG et al., 2022).

No estudo de caso, observou-se que a os fatores cognitivos e emocionais, principalmente voltados para transparência, influenciam positivamente na promoção de confiança humana em IA.

Pode-se ressaltar, por exemplo, que transparência na exposição de que IA está sendo implementada e como isto afetará a todos é chave para adesão dos funcionários. Deste modo, considerar transparência ao implementar IA é um primeiro passo que contribui para tomada de decisão envolvendo esta tecnologia.

6.3 RECOMENDAÇÕES PARA ESTUDOS FUTUROS

Pesquisas futuras podem abordar mais detalhadamente o comportamento de confiança e os fatores que influenciam considerando grupos de trabalho, convivência e distância geográfica, principalmente considerando o atual formato de trabalho remoto ou teletrabalho.

Considerando que IA é ainda uma tecnologia pouco conhecida e utilizada nas organizações e que a maturidade no uso de IA influencia diretamente na confiança dos usuários, recomendam-se estudos relacionados a maturidade no uso de IA nas organizações.

A adaptação dos funcionários ao uso de IA é um fator relevante para os avanços desta tecnologia e para o desenvolvimento de confiança. Por isto, estudos futuros que considerem uma abordagem de implementação de IA mais humanizada, considerando o ponto de vista dos funcionários afetados pode promover maior confiança e aderência.

Ainda, como estudos futuros, recomenda-se a calibração das expectativas iniciais dos indivíduos em relação a questão da transparência, especialmente nos casos em que a tecnologia ainda está em estágios iniciais de desenvolvimento e pode sofrer alterações ao longo da implantação devido aos desafios encontrados durante o caminho.

Por fim, recomenda-se a realização de estudos mais aprofundados sobre os fatores cognitivos e emocionais que influenciam confiança em IA em relação às

distintas formas de IA considerando uma variedade de setores em que IA é utilizada, especialmente no meio digital, redes sociais, e-commerce e metaverso.

6.4 LIMITAÇÕES DA PESQUISA

As limitações do presente estudo são estão relacionadas a:

Necessidade de extensão da coleta de dados para funcionários que usam IA no dia a dia uma vez que o presente estudo trouxe uma visão mais gerencial e estratégica da camada de gestão.

Por fim, a segunda limitação está no fato de ter sido considerado um estudo em empresas do setor automotivo. Tal limitação pode ser superada em novos estudos com uma amostragem de outros setores, observando as características e necessidades de cada um.

REFERÊNCIAS

- ASAN, O.; BAYRAK, A. E.; CHOUDHURY, A. Artificial Intelligence and Human Trust in Healthcare: Focus on Clinicians. **Journal of Medical Internet Research**, v. 22, n. 6, p. 1–7, 2020.
- BARDIN, L. Análise de conteúdo. 70a ed. São Paulo, 2011.
- Edições 70. BEIER, G.; KIEFER, J.; KNOPF, J. Potentials of big data for corporate environmental management: A case study from the German automotive industry. **Journal of industrial ecology**, v. 26, n. 1, p. 336–349, 2022.
- BEIER, G.; NIEHOFF, S.; HOFFMANN, M. Industry 4.0: a step towards achieving the SDGs? A critical literature review. **Discover sustainability**, v. 2, n. 1, p. 1–21, 2021.
- BEIER, G.; ULLRICH, A.; NIEHOFF, S.; REISSIG, M.; HABICH, M. Industry 4.0: How it is defined from a sociotechnical perspective and how much sustainability it includes—A literature review. **Journal of cleaner production**, v. 259, p. 120856, 2020.
- BOTELHO, L. L. R.; DE ALMEIDA CUNHA, C. C.; MACEDO, M. O método da revisão integrativa nos estudos organizacionais. **Gestão e sociedade**, v. 5, n. 11, p. 121–136, 2011.
- BRIDGE, F. **Artificial Intelligence Reshaping the Automotive Industry**. 2020. Disponível em: <https://www.futurebridge.com/industry/perspectives-mobility/artificial-intelligence-reshaping-the-automotive-industry/>. Acesso em: 30 abr. 2021.
- CHONG, L.; ZHANG, G.; GOUCHER-LAMBERT, K.; KOTOVSKY, K.; CAGAN, J. Human confidence in artificial intelligence and in themselves: The evolution and impact of confidence on adoption of AI advice. **Computers in Human Behavior**, v. 127, 2022.
- CHRISTENBERRY, T. Standalone Interviews Do Not Equal Qualitative Research. **Nurse Author & Editor**, v. 27, n. 4, p. 1–9, 2017.
- COOPER, D. R.; SCHINDLER, P. S. **Métodos de pesquisa em Administração**. 12ª ed. Porto Alegre: AMGH, 2016.

CRESWELL, J. W. **Projeto de pesquisa: métodos qualitativo, quantitativo e misto**. 3º ed. Porto Alegre: Bookman, 2014.

CURIONI, A. Artificial intelligence: Why we must get it right. **Informatik-Spektrum**, v. 41, n. 1, p. 7–14, 2018.

EDMONDS, M.; GAO, F.; LIU, H.; et al. A tale of two explanations: Enhancing human trust by explaining robot behavior. **Science Robotics**, v. 4, n. 37, p. 1–14, 2019.

EISENHARDT, K. M. Building theories from case study research. **Academy of management review**, v. 14, n. 4, p. 532–550, 1989.

EISENHARDT, K. M.; GRAEBNER, M. E. Theory building from cases: opportunities and challenges. **Academy of management journal**, v. 50, n. 1, p. 25–32, 2007.

GHOBAKHLOO, M.; FATHI, M.; IRANMANESH, M.; MAROUFKHANI, P.; MORALES, M. E. Industry 4.0 ten years on: A bibliometric and systematic review of concepts, sustainability value drivers, and success determinants. **Journal of Cleaner Production**, v. 302, p. 127052, 2021.

GIL, A. CARLOS. **Métodos e Técnicas de Pesquisa Social**. 7º ed. São Paulo: Atlas, 2019.

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.

GODOI, C.K; BANDEIRA-DE-MELLO, R.B.; SILVA, A. B. **Pesquisa qualitativa em estudos organizacionais: Paradigmas, estratégias e métodos**. 2º ed. São Paulo: Ed. Saraiva, 2010.

HAGENDORFF, T.; WEZEL, K. 15 challenges for AI: or what AI (currently) can't do. **AI & SOCIETY**, v. 35, p. 355–365, 2020.

HAYASHI JR, P.; ABIB, G.; HOPPEN, N. Validity in qualitative research: A processual approach. **The Qualitative Report**, v. 24, n. 1, p. 98–112, 2019.

KAHNEMAN, D. 'Clearly AI is going to win. How people are going to adjust is a fascinating problem'. **The Guardian**, 2021. Entrevista.

KOTHARI, C. R. **Research methodology: Methods and techniques**. New Age International, 2004.

ŁAPIŃSKA, J.; ESCHER, I.; GÓRKA, J.; SUDOLSKA, A.; BRZUSTEWICZ, P. Employees' trust in artificial intelligence in companies: The case of energy and chemical industries in Poland. **Energies**, v. 14, n. 7, 2021.

LEWIS, P. R.; MARSH, S. What is it like to trust a rock? A functionalist perspective on trust and trustworthiness in artificial intelligence. **Cognitive Systems Research**, v. 72, p. 33–49, 2022.

MARCH, J. G.; SIMON, H. A. **Organizations**. John Wiley & Sons, 1993.

MARTINS, G. DE A.; THEÓPHILO, C. R. **Metodologia da Investigação Científica para Ciências Sociais Aplicadas**. 3. ed ed. São Paulo: Atlas, 2007.

MOHER, D.; LIBERATI, A.; TETZLAFF, J.; ALTMAN, D. G.; GROUP*, P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. **Annals of internal medicine**, v. 151, n. 4, p. 264–269, 2009.

NAHAVANDI, S. Industry 5.0-a human-centric solution. **Sustainability (Switzerland)**, v. 11, n. 16, 2019. MDPI.

NOBRE, F. **On Cognitive Machines in Organizations**, 2005. PhD Thesis, Birmingham: University of Birmingham.

NOBRE, F. S. Core competencies of the new industrial organization. **Journal of Manufacturing Technology Management**, 2011.

NOBRE, F. S.; TOBIAS, A. M.; WALKER, D. S. The pursuit of cognition in manufacturing organizations. **Journal of Manufacturing Systems**, v. 27, n. 4, p. 145–157, 2008.

NOBRE, F. S.; TOBIAS, A. M.; WALKER, D. S. The impact of cognitive machines on complex decisions and organizational change. **AI & society**, v. 24, n. 4, p. 365–381, 2009.

NUTT, P. C.; WILSON, D. C. **Handbook of decision making**. John Wiley & Sons, 2010.

NYGARD, K.; CHOWDHURY, M.; BUGALWI, A.; KOTALA, P. People and Intelligent Machines in Decision Making. **Int. J. Comput. Their Appl**, 2017.

DE PAIVA JÚNIOR, F. G.; DE SOUZA LEÃO, A. L. M.; DE MELLO, S. C. B. Validade e confiabilidade na pesquisa qualitativa em administração. **Revista de Ciências da Administração**, v. 13, n. 31, p. 190–209, 2011.

PEREIRA, J. M. **Manual de metodologia da pesquisa científica**. 4º ed. Atlas, 2010.

PERES, R. S.; JIA, X.; LEE, J.; et al. Industrial artificial intelligence in industry 4.0-systematic review, challenges and outlook. **IEEE Access**, v. 8, p. 220121–220139, 2020.

RAISCH, S.; KRAKOWSKI, S. Artificial intelligence and management: The automation–augmentation paradox. **Academy of Management Review**, v. 46, n. 1, p. 192–210, 2021.

SABMANNSHAUSEN, T.; BURGGRÄF, P.; WAGNER, J.; et al. Trust in artificial intelligence within production management—an exploration of antecedents. **Ergonomics**, v. 0, n. 0, p. 1–18, 2021.

SEURING, S. A. Assessing the rigor of case study research in supply chain management. **Supply Chain Management: An International Journal**, v. 13, n. 2, p. 128–137, 2008. Emerald Group Publishing Limited. Disponível em: <<https://doi.org/10.1108/13598540810860967>>. Acesso em: 20 abr. 2021.

SHARAN, N. N.; ROMANO, D. M. The effects of personality and locus of control on trust in humans versus artificial intelligence. **Heliyon**, v. 6, n. 8, p. e04572, 2020. Elsevier Ltd. Disponível em: <<https://doi.org/10.1016/j.heliyon.2020.e04572>>. Acesso em: 20 abr. 2021.

SIMON, H. A. **Administrative behavior**. 4. ed. New York: The free Press, 1997.

SNYDER, H. Literature review as a research methodology: An overview and guidelines. **Journal of Business Research**, v. 104, n. August, p. 333–339, 2019.

SONG, X.; XU, B.; ZHAO, Z. Can people experience romantic love for artificial intelligence? An empirical study of intelligent assistants. **Information and Management**, v. 59, n. 2, 2022.

TORRACO, R. J. Writing Integrative Literature Reviews: Guidelines and Examples. **Human Resource Development Review**, v. 4, n. 3, p. 356–367, 2005.

TRIVIÑOS, A. N. S. A pesquisa qualitativa em educação. **São Paulo: Atlas**, 1987.

VERESCHAK, O.; BAILLY, G.; CARAMIAUX, B. How to evaluate trust in AI-assisted decision making? A survey of empirical methodologies. **Proceedings of the ACM on Human-Computer Interaction**, v. 5, n. CSCW2, p. 1–39, 2021.

VERGARA, S. CONSTANT. **Projetos e relatórios de pesquisa em administração**. 16º ed. São paulo, 2016.

VOSS, C. Case research in operations management. **Researching operations management**. p.176–209, 2010.

WANG, B.; ZHENG, P.; YIN, Y.; SHIH, A.; WANG, L. Toward human-centric smart manufacturing: A human-cyber-physical systems (HCPS) perspective. **Journal of Manufacturing Systems**, v. 63, p. 471–490, 2022.

XU, X.; LU, Y.; VOGEL-HEUSER, B.; WANG, L. Industry 4.0 and Industry 5.0—Inception, conception and perception. **Journal of Manufacturing Systems**, v. 61, p. 530–535, 2021.

YIN, R. K. **Estudo de Caso-: Planejamento e métodos**. Bookman editora, 2015.

YOKOI, R.; EGUCHI, Y.; FUJITA, T.; NAKAYACHI, K. Artificial Intelligence Is Trusted Less than a Doctor in Medical Treatment Decisions: Influence of Perceived Care and

Value Similarity. **International Journal of Human-Computer Interaction**, v. 37, n. 10, p. 981–990, 2021.

ZHANG, Z.; CITARDI, D.; WANG, D.; et al. Patients' perceptions of using artificial intelligence (AI)-based technology to comprehend radiology imaging data. **Health Informatics Journal**, v. 27, n. 2, 2021.

ZHANG, Z.; GENC, Y.; WANG, D.; AHSEN, M. E.; FAN, X. Effect of AI Explanations on Human Perceptions of Patient-Facing AI-Powered Healthcare Systems. **Journal of Medical Systems**, v. 45, n. 6, 2021.

ANEXO I: Roteiro de Entrevistas

PROCEDIMENTOS INICIAIS

Nome da empresa:

Dados de identificação do entrevistado:

Setor:

Função:

Tempo de atuação na empresa:

Tempo na função:

Grau de escolaridade:

Idade:

QUESTÕES PARA A ENTREVISTA

Parte 1 – Contexto Geral

Contexto do porquê, quando e como a empresa optou por adotar IA em seus processos produtivos e produtos.

1. Quando a empresa começou a adotar IA nos processos produtivos? E nos produtos?
2. A decisão inicial quanto ao uso de IA foi global ou local?
3. Qual a relevância de IA para a empresa (alta, média, baixa)?
4. Quais os fatores internos e externos que estimularam a empresa a adotar IA nos seus processos produtivos e produtos? Descreva os desafios e oportunidades.
5. Como foi a aceitação dos colaboradores em relação ao uso de IA nos processos produtivos (positiva, negativa ou neutra)?
6. E como foi a aceitação dos clientes em relação ao uso de IA nos produtos?
7. Quais os principais desafios identificados em relação a adoção de IA nos processos produtivos? E nos produtos?
8. Quais áreas da empresa foram mais afetadas pela adoção do uso de IA nos processos produtivos? E nos produtos? (Enumere-as como operacionais,

gerenciais/administrativas, organizacionais, inter-organizacionais). Destaque como elas foram afetadas?

9. Qual é a percepção dos funcionários sobre a adoção de IA nos processos produtivos da empresa? (Desconfiança ou confiança)? E como você espera que esta tecnologia venha beneficiar as pessoas na empresa?

10. Qual é a percepção dos clientes sobre a adoção de IA nos produtos da empresa? (Desconfiança ou confiança)?

11. Quais benefícios puderam ser observados em relação ao uso de IA nos processos produtivos? (Exemplo: otimização do tempo de produção, redução de falhas/retrabalho, maior qualidade, aumento da transparência, redução de custo operacional, mais eficiência etc.)

12. Quais benefícios puderam ser observados em relação ao uso de IA nos produtos? (Exemplo: maior aceitação pelos clientes, aumento nas vendas, diferencial competitivo, maior usabilidade etc.).

13. Quais os principais desafios identificados durante a implementação de IA nos processos produtivos no nível interno e externo? (Exemplo: alto custo da tecnologia, falta de mão de obra capacitada, ausência de políticas regulatórias para uso, operacionalização etc.).

14. A Lei Geral de Proteção de Dados (LGPD) ou General Data Protection Regulation (GDPR) afetou de alguma forma o uso de IA nos processos produtivos ou produtos da empresa? Em caso afirmativo, quais medidas foram tomadas?

15. Quais os resultados esperados em relação ao uso de IA nos processos produtivos no curto, médio e longo prazo? E nos produtos? Detalhe os resultados já alcançados e os encaminhados.

Parte 2 – Tipos de IA

Nota: Glikson e Woolley (2020) definem abaixo os principais tipos de IA.

a) Robô físico: representações mecânicas físicas com similaridade humana contendo uma gama de funções e capacidades, que podem estar remotamente ou fisicamente presentes executando tarefas para as quais foram programados.

b) Agentes virtuais: representação inteligente com características humanas sem presença física, uma identidade distinta como chat bot ou avatar). Exemplo: Lu, o chat bot premiado do Magazine Luiza.

c) Embutida: é uma forma de IA invisível para os humanos, não possui representação visual e por ser totalmente incorporada a sistemas e pode ser incorporada a produtos e serviços tais como mecanismos de busca, programas e aplicativos.

1. Considerando os tipos de IA descritos acima (Robô Físico, Agentes Virtuais e Embutida), quais são os tipos de IA mais utilizados nos processos produtivos da empresa?

2. Quais os tipos de IA mais utilizados nos produtos? Justifique os critérios de escolha de cada tipo.

3. IA foi utilizada de forma pura ou em conjunto com outras tecnologias (Exemplo: IoT, Big Data, Cloud etc.)?

4. Quais os principais desafios identificados durante a implementação de cada tipo de IA nos processos produtivos em relação aos colaboradores usuários da tecnologia que atuam no dia a dia (desconfiança, medo de perder o emprego, resistência, baixa utilização, falta de conhecimento, necessidade de desenvolver habilidades específicas etc.);

5. Quais os principais desafios identificados durante a implementação de cada tipo de IA nos produtos em relação aos clientes ou consumidores finais que adquirem veículos com IA como tecnologia embarcada? (Exemplo: alto custo da tecnologia, desconfiança dos usuários, baixa utilização etc.);

Parte 3 – Fatores cognitivos e emocionais que influenciam confiança humana em IA.

Nota: Glikson e Woolley (2020) consideram que confiança é uma construção e pode ser influenciada por fatores cognitivos (avaliação racional e características situacionais) ou emocionais (baseado no afeto, conexões sociais), neste sentido, as autoras apresentam como tais fatores influenciam o grau de confiança em cada tipo de IA (Robô físico, Agentes virtuais e embutida) propõem em seu modelo as dimensões de confiança cognitiva (tangibilidade, transparência, confiabilidade,

características da tarefa e comportamentos imediatos) e emocional (tangibilidade comportamentos imediatos e Antropomorfismo).

Robô Físico

Atualmente existem Robôs Físicos sendo utilizados nos processos produtivos e produtos da empresa? Estes Robôs têm nome, aparência e comportamentos humanos? Quais? Detalhar.

1. Tangibilidade.

a. Fatores cognitivos: atualmente existem robôs com aparência humana nos processos produtivos ou produtos da empresa? Na sua opinião a presença física dos robôs contribui para o aumento do nível de confiança dos colaboradores/usuários em IA?

b. Fatores emocionais: a presença física de robôs, embora em alguns casos possa aumentar a confiança, em outros pode causar medo nos usuários/colaboradores? Em geral, como os colaboradores e usuários reagido em relação aos robôs (positivamente, negativamente, neutro).

2. Transparência.

a. Fatores cognitivos: na sua opinião transparência (fluxo contínuo de informação) é relevante para adoção de IA? Quais aspectos de transparência a empresa tem utilizado para aumentar a confiança dos colaboradores/usuários em IA?

3. Confiabilidade.

a. Fatores cognitivos: considerando que baixa confiabilidade reduz os estímulos de confiança, qual o grau de confiabilidade percebido em relação ao uso de robôs pelos colaboradores e usuários desta tecnologia na empresa (alto, médio, baixo)?

4. Características da Tarefa.

a. Fatores cognitivos: na sua opinião, geralmente a confiança percebida é maior em relação a tarefas técnicas do que as que requerem inteligência social?

5. Comportamentos imediatos.

a. Fatores cognitivos: a capacidade de resposta, adaptabilidade e os comportamentos pró-sociais aumentam a confiança em IA. Quais os principais comportamentos observados nos colaboradores e/ou clientes usuários de IA?

b. Fatores emocionais: geralmente os robôs que têm algum tipo de comportamentos humano geram maior percepção de confiança. Qual sua opinião em relação a isto (concorda, discorda). Por quê?

6. Antropomorfismo.

a. Fatores emocionais: robôs que tem semelhança com humanos geram maior percepção de confiança ou ainda podem gerar certo desconforto. Quais as principais características humanas dos robôs da empresa? Como os colaboradores/usuários percebem estas características (positivamente, negativamente ou neutro).

Agentes virtuais

Atualmente existem agentes virtuais sendo utilizados nos processos produtivos e produtos da empresa? Estes agentes têm nome, aparência ciberfísica e comportamentos humanos? Quais? Detalhar

1. Tangibilidade.

a. Fatores cognitivos: Você considera que a confiança é maior em agentes virtuais do que em IA embutida?

b. Fatores emocionais: Qual a percepção da empresa em relação a confiança dos colaboradores/usuários no uso de agentes virtuais? Ter presença de um personagem que lembra um humano aumentou a confiança dos usuários/colaboradores? Comentar.

2. Transparência.

a. Fatores cognitivos: a transparência (fluxo contínuo de informação) e explicações sobre como funciona o algoritmo por trás do agente virtual pode aumentar a confiança. A empresa já realizou alguma ação no sentido de promover informações mais técnicas sobre a tecnologia usada para os colaboradores e/ou usuários dos produtos?

3. Confiabilidade.

a. Fatores cognitivos: Considerando que baixa confiabilidade reduz os estímulos de confiança, qual o grau de confiabilidade percebido em relação ao uso de robôs pelos colaboradores e usuários (alto, médio, baixo)?

4. Características da Tarefa.

a. Fatores cognitivos: tarefas complexas como por exemplo análise de dados tendem a gerar maior confiança em IA. Você concorda? Quais os principais tipos de atividades desenvolvidas pela empresa?

5. Comportamentos imediatos.

a. Fatores cognitivos: táticas de persuasão e personificação aumentam a confiança em IA. A empresa considerou alguns destes comportamentos ao implantar IA nos processos ou produtos?

b. Fatores emocionais: a similaridade com humanos tende a aumentar o nível de confiança em IA, no entanto, ainda depende da predisposição dos usuários. Como você classifica a percepção dos colaboradores/usuários/clientes em relação ao uso de agentes virtuais nos processos produtivos e produtos da empresa.

6. Antropomorfismo.

a. Fatores emocionais: a similaridade com humanos, atratividade e personalização, como etnia ou semelhança facial com o usuário, aumenta a confiança, mas também cria alta expectativas em relação às habilidades da IA. Como foi o processo de decisão em relação a quais características considerar? Quais os benefícios e desafios encontrados?

IA Embutida

Atualmente existe alguma forma de IA embutida sendo utilizados nos processos produtivos e produtos da empresa? Quais são? Detalhar.

1. Tangibilidade.

a. Fatores emocionais: Desconhecer que IA está sendo utilizada pode evocar sentimentos negativos, desconforto, desconfiança, raiva. Porém, emoções positivas podem ser induzidas por boa reputação de uma empresa. Na sua opinião a reputação é um fator relevante para a adoção de IA?

2. Transparência.

a. Fatores cognitivos: compreender com transparência como o algoritmo funciona, por exemplo, pode aumentar a confiança em IA, especialmente para sistemas de

gerenciamento altamente inteligentes. Vocês consideram ou consideraram este fator ao implementar IA?

3. Confiabilidade.

a. Fatores cognitivos: baixa confiabilidade diminui significativamente a confiança e recuperá-la é difícil e leva tempo. Como a empresa busca mitigar esta situação? Existem ações mapeadas? Quais?

4. Características da Tarefa.

a. Fatores cognitivos: tarefas que requerem inteligência social tem maior aceitação e confiança em humanos do que em IA, porém, alta autoconfiança modera a confiança na IA, principalmente para tarefas especializadas. Na sua opinião, onde IA embutida tem sido mais utilizada pela empresa? Quais tipos de tarefa?

5. Comportamentos imediatos.

a. Fatores cognitivos: A personalização melhora a confiança, porém, a sensação de acompanhamento/monitoramento constante do comportamento dos usuários de IA pode diminuir a confiança. A empresa busca promover alguma ação que ajude a mitigar estes casos?

6. Antropomorfismo.

a. N/A.

**ANEXO II – Termo de Autorização
Discente/: Patrícia Valério Goulart**

Linha de Pesquisa: Competências Organizacionais e Decisão

Termo de Autorização de Entrevistas com Finalidade Acadêmica

Todas as entrevistas a serem conduzidas pela autora se restringem à finalidade de pesquisa pura e acadêmica e tem como objetivo a coleta de dados que fomentarão análises do estudo de caso da dissertação de mestrado de Patrícia Valério Goulart, mestranda do Programa de Pós-Graduação em Gestão de Organizações, Liderança e Decisão da Universidade Federal do Paraná.

O conteúdo das entrevistas possuem só e somente só finalidade de pesquisa acadêmica de mestrado, não interfere na conduta ética da empresa e da Universidade Federal do Paraná. Adicionalmente, as partes envolvidas (empresa, UFPR, autora e entrevistados(as)) assinarão por livre e espontânea vontade, de boa-fé, este Termo que estabelece autorização para realização das entrevistas.

Esta pesquisa possui como objetivo estudar as implicações de confiança humana em Inteligência Artificial em empresas do setor automotivo que adotam IA na forma robótica, virtual e embutida. Para tal, pretende-se realizar aproximadamente 5 entrevistas com colaboradores/funcionários da empresa estudada que aceitaram por livre e espontânea vontade, de boa-fé, concedê-las à autora. A autora garante que os nomes dos colaboradores serão tratados com extremo sigilo e não serão divulgados sob forma alguma. Também não serão divulgados em hipótese alguma, informações confidenciais da empresa, objeto de estudo, nem mesmo nome, razão social e localização.

CLÁUSULA 1ª. E por estarem de acordo com os termos aqui apresentados, as partes abaixo assinam o presente Termo por livre e espontânea vontade, sem qualquer tipo de coação ou indução em erro, o que fazem após ter tido plena ciência do conteúdo do instrumento/protocolo e das questões pertinentes às entrevistas que têm objetivos puramente acadêmicos, assinando o presente termo, que foi lido e achado conforme, dando por boas e verdadeiras as assinaturas lançadas.

Curitiba, dia ___ de março de 2022

Discente/Autora

Patrícia Valério Goulart

Professor Orientador

Dr. Farley Simon Mendes Nobre

Entrevistado