

## A TEORIA DO AGIR COMUNICATIVO DE HABERMAS E A INTELIGÊNCIA ARTIFICIAL

**Clóvis Ricardo Montenegro de Lima**

Instituto Brasileiro de Informação em Ciência e Tecnologia  
clovismlima@gmail.com

**Ana Gabriela Clipes Ferreira**

Universidade Federal do Rio Grande do Sul  
Instituto Brasileiro de Informação em Ciência e Tecnologia  
anaclipes@ufrgs.br

**Flávia da Silva Carvalho**

Instituto Brasileiro de Informação em Ciência e Tecnologia  
flaviaslvcarvalho@gmail.com

**Samya Massari**

Instituto Brasileiro de Informação em Ciência e Tecnologia  
samya.massari@marciocavalcante.com.br

---

### Resumo

A presente pesquisa tem como objetivo analisar na literatura a produção científica sobre a inteligência artificial e discutir o seu uso com o olhar habermasiano através da teoria do agir comunicativo. Ilustra exemplos de aplicação da IA em obras de ficção. Discute as questões das tecnologias digitais e as limitações binárias da inteligência artificial. Como método, adota duas etapas distintas: abordagem teórica através da Teoria do Agir Comunicativo; e a introdução do resultado de pesquisas através do método de revisão sistemática, a fim de reunir literatura relevante disponível em bases de dados. A base de dados selecionada foi a Scopus, base internacional multidisciplinar. Os termos utilizados na estratégia de busca, na busca avançada e sem corte temporal, foram: inteligência artificial, Habermas e teoria do agir comunicativo. A pesquisa resultou em sete registros. A discussão envolve a análise desses documentos, embasados na literatura científica, em especial a teoria do agir comunicativo. Conclui que, embora as teorias de Habermas não sejam diretamente relacionadas com a inteligência artificial, as ideias se conectam ao proporcionar reflexões como a ética do discurso e a relação da sociedade com a tecnologia.

**Palavras-chave:** Inteligência artificial. Teoria do Agir Comunicativo. Jürgen Habermas.

## HABERMAS' THEORY OF COMMUNICATIVE ACTION AND ARTIFICIAL INTELLIGENCE

### Abstract

This research aims to analyze the scientific production on artificial intelligence in the literature and discuss its use from a Habermasian perspective through the theory of communicative action. It illustrates examples of the application of AI in works of fiction. It discusses the issues of digital technologies and the binary limitations of artificial intelligence. As a method, it adopts two distinct stages: a theoretical approach through the Theory of Communicative Action; and the introduction of research results through the systematic review method, in order to gather relevant literature available in databases. The selected database was Scopus, an international multidisciplinary database. The terms used in the search strategy, in the advanced search and without temporal cut, were: artificial intelligence, Habermas and theory of communicative action. The search resulted in 7 records. The discussion involves the analysis of these documents, based on the scientific literature, especially the theory of communicative action. It concludes that, although Habermas' theories are not directly related to artificial intelligence, the ideas connect by providing reflections such as the ethics of discourse and the relationship between society and technology.

**Keywords:** Artificial intelligence. Theory of Communicative Action. Jürgen Habermas.



Esta obra está licenciada sob uma licença

Creative Commons Attribution 4.0 International (CC BY-NC-SA 4.0).

## TEORÍA DE LA ACCIÓN COMUNICATIVA E INTELIGENCIA ARTIFICIAL DE HABERMAS

### Resumen

La presente investigación tiene como objetivo analizar la producción científica sobre inteligencia artificial en la literatura y discutir su uso con una perspectiva habermasiana a través de la teoría de la acción comunicativa. Ilustra ejemplos de la aplicación de la IA en obras de ficción. Analiza los problemas de las tecnologías digitales y las limitaciones binarias de la inteligencia artificial. Como método adopta dos etapas diferenciadas: abordaje teórico a través de la Teoría de la Acción Comunicativa; y la introducción de los resultados de la investigación a través del método de revisión sistemática, con el fin de recopilar literatura relevante disponible en bases de datos. La base de datos seleccionada fue Scopus, una base de datos multidisciplinaria internacional. Los términos utilizados en la estrategia de búsqueda, en la búsqueda avanzada y no cronometrada, fueron: inteligencia artificial, Habermas y teoría de la acción comunicativa. La búsqueda arrojó como resultado 7 registros. La discusión pasa por el análisis de estos documentos, basándose en la literatura científica, especialmente en la teoría de la acción comunicativa. Se concluye que, si bien las teorías de Habermas no están directamente relacionadas con la inteligencia artificial, las ideas se conectan aportando reflexiones como la ética del discurso y la relación de la sociedad con la tecnología.

**Palabras clave:** Inteligencia artificial. Teoría de la Acción Comunicativa. Jürgen Habermas.

## 1 INTRODUÇÃO

O uso da Inteligência Artificial (IA) em atividades sociais cotidianas, aplicada em análises ou trabalhos especializados faz parte do imaginário da humanidade há décadas. A substituição do trabalho humano manual ou intelectual é um assunto recorrente e que traz discussões sobre os benefícios e os riscos dentro e fora do mundo acadêmico. Nesta pesquisa, pretendemos apresentar parte dessa discussão, à luz do filósofo alemão contemporâneo Jürgen Habermas.

Obras de ficção, como livros e filmes, abordam a temática em diferentes aspectos, desde a humanização das máquinas, como nos títulos "A.I. - Inteligência Artificial", de 2001, inspirado em um conto de Brian Wilson Aldiss, *Supertoys Last All Summer Long* (1969), e "Eu, Robô", de 2004, e que tem como base a coleção de contos do escritor Isaac Asimov até às questões morais que ocorrem sobre o comportamento da máquina HAL-9000 (*Heuristically programmed ALgorithmic computer*, ou, em tradução livre, Computador Algorítmico Heuristicamente Programado), em "2001: Uma Odisséia no Espaço", filme e romance de 1968, feitos ao mesmo tempo. Na franquia de filmes "Alien" (1979, 1986, 1992, 1997, 2012 e 2017) há a presença de robôs humanóides que tomam decisões arbitrárias e priorizam a missão em detrimento da vida da tripulação, tal como HAL-9000. Em alguns casos, desenvolvem consciência e objetivos próprios.

Algumas utilizações de máquinas e a implementação de determinadas tecnologias, envolvendo IA ou não, são vistas como uma verdadeira utopia, assim como previsões de computadores pessoais que cabem na palma da mão ou chamadas em vídeo fazendo parte do dia a dia das pessoas. Todavia, diversas delas se concretizam e até mesmo extrapolam as previsões. O futuro das ficções é o "aqui e agora" conforme Gigliotti (2014) ao analisar algumas obras cinematográficas. O uso da inteligência artificial na Administração Pública é uma das aplicações atuais, da década de 2020, das tecnologias e têm sido aplicados na triagem e análise inicial de documentos.

O déficit de 10 mil funcionários em 2022 no Instituto Nacional do Seguro Social (INSS) fez a sua direção recorrer à aplicação da Inteligência Artificial na análise de pedidos de concessão de benefícios. É a alternativa para atender uma fila para perícias de cerca de 1,8 milhão de segurados (Melo, 2022). O Instituto acredita que, assim, é possível dar andamento mais rápido aos processos. Contudo, os robôs indeferem mais de 300 mil benefícios, sendo a maioria assistenciais (Gercina, 2022).

Observa-se que a interação humana é indispensável para a adequada análise dos documentos necessários à concessão de benefícios previdenciários. O modelo de Inteligência Artificial (IA) usado pelo instituto não é capaz de identificar que um pedido irregular pode ser suprido pela simples apresentação de um documento faltante e não que o segurado não tem direito à percepção do benefício.

O uso da Inteligência Artificial na Administração Pública traz algumas questões de interesse quanto à moralidade e eticidade no uso dessa tecnologia nas práticas de governo. No exemplo do INSS, a ausência de trabalho e supervisor humano opera em uma escala de erro que produz a negação de um direito fundamental à assistência social de milhares de beneficiários.

Cabe dizer que permitir a tomada de decisão automática não garante efetividade pois, se a inteligência artificial é capaz de tomar milhões de decisões a cada segundo - o que nenhum ser humano é capaz de supervisionar -, e se houver apenas uma pequena margem de erro, por exemplo, cerca de 1%, 1% de um milhão é 10 mil, o que resulta em 10 mil erros por segundo. Diego Cherulli, vice-presidente do Instituto Brasileiro de Direito Previdenciário, afirma que em 2022 o robô "[...] acaba indeferindo tudo porque os cadastros normalmente têm erros. Esse trabalho precisa ser feito por um servidor, um ser humano. A máquina não está preparada para isso" (Lüder, 2022).

Este trabalho pretende verificar na literatura, através do resultado da busca em base de dados multidisciplinar, a produção científica sobre a Inteligência Artificial com abordagem da Teoria do Agir Comunicativo (TAC) de Habermas (2012). Essa teoria explora características humanas: atos de fala, interação mediada pela linguagem, socialização e intersubjetivação. A partir dos resultados obtidos na busca quer-se discutir o uso da inteligência artificial sob a perspectiva habermasiana na TAC.

## **2 QUESTÕES DAS TECNOLOGIAS DIGITAIS DA INTELIGÊNCIA: AS LIMITAÇÕES BINÁRIAS**

Inteligência Artificial (IA) refere-se à capacidade de um sistema ou máquina de imitar ou simular a inteligência humana. É um campo multidisciplinar da ciência da computação que se concentra no desenvolvimento de algoritmos e técnicas que permitem que os sistemas aprendam, raciocinem, tomem decisões e ajam de forma autônoma.

A IA é baseada na ideia de construir computadores e sistemas capazes de executar tarefas que normalmente requerem inteligência humana, como reconhecimento de fala, visão computacional, processamento de linguagem natural, tomada de decisões, resolução de

problemas e aprendizado. Para utilizar um corpo de conhecimento em uma máquina, é necessário definir uma maneira de representá-lo. A IA:

[...] tornou-se um termo genérico para aplicações que executam tarefas complexas que antes exigiam interação humana, como se comunicar com clientes online ou jogar xadrez. O termo é frequentemente usado de forma intercambiável com seus subcampos, que incluem *machinelearning* (ML) e *deeplearning*. No entanto, existem diferenças. Por exemplo, o *machinelearning* é centrado na criação de sistemas que aprendam e melhorem seu desempenho com base nos dados que eles consomem. É importante notar que, embora todo *machinelearning* seja IA, nem toda IA é *machinelearning* (ORACLE, 2021).

Dentro de qualquer programa de computador, está embutido o conhecimento relacionado a um problema específico que precisa ser resolvido. Esse conhecimento se manifesta nos algoritmos que o programa utiliza e nas regras de decisão que determinam qual algoritmo aplicar em circunstâncias particulares. Quando um programa é carregado em um computador, podemos dizer que o computador “adquire” esse conhecimento, embora na maioria dos programas, essas informações não sejam representadas explicitamente e, portanto, não possam ser facilmente atualizadas ou manipuladas.

Uma das características fundamentais dos programas de Inteligência Artificial (IA) é sua estrutura que separa claramente o código executável dos dados ou conhecimento do sistema. Assim, na IA, o termo “conhecimento” refere-se à informação necessária para que um programa de computador seja capaz de agir de maneira inteligente.

A lógica procedimental é uma abordagem da inteligência artificial (IA) que utiliza a lógica simbólica e algoritmos para representar e manipular o conhecimento. Essa abordagem é baseada em regras formais e procedimentos explícitos para realizar inferências lógicas e resolver problemas. Na lógica procedimental, o conhecimento é representado por meio de símbolos e regras de inferência. Os símbolos podem representar objetos, relações, propriedades, ações ou qualquer outra informação relevante para o problema em questão. As regras de inferência definem como esses símbolos podem ser combinados e manipulados para chegar a conclusões. Esse mecanismo permite que o sistema de IA resolva problemas, faça deduções e tome decisões com base na lógica formal.

A abordagem baseada em regras e procedimentos é muito útil em domínios nos quais o conhecimento é bem definido e estruturado. No entanto, a lógica procedimental pode encontrar dificuldades quando lida com problemas mais complexos ou situações de incerteza, nos quais o conhecimento é incompleto, vago ou sujeito a exceções.

O funcionamento das Inteligências Artificiais é baseado em algoritmos complexos de aprendizado de máquina. Esses algoritmos podem ser:

- a) como redes neurais convolucionais (CNNs): algoritmo de Aprendizado Profundo que pode captar uma imagem de entrada, atribuir importância (pesos e vieses que podem ser aprendidos) a vários aspectos/objetos da imagem e ser capaz de diferenciar um do outro. O pré-processamento exigido em uma ConvNet é muito menor em comparação com outros algoritmos de classificação. Enquanto nos métodos primitivos os filtros são feitos à mão, com treinamento suficiente, as ConvNets têm a capacidade de aprender esses filtros/características. A arquitetura de uma ConvNet é análoga àquela do padrão de conectividade de neurônios no cérebro humano e foi inspirada na organização do córtex visual. Os neurônios individuais respondem a estímulos apenas em uma região restrita do campo visual conhecida como Campo Receptivo. Uma coleção desses campos se sobrepõe para cobrir toda a área visual (Deep Learning Book, [2024a]). **Introdução às Redes Neurais Convolucionais.** [2024a]).
- b) redes adversárias generativas (GANs): arquiteturas de redes neurais profundas compostas por duas redes colocadas uma contra a outra (daí o nome “adversárias”. (Deep Learning Book, [2024b])).

Esses algoritmos são alimentados com grandes volumes de dados e treinados para reconhecer padrões e características presentes nesses conjuntos de dados. Uma vez treinadas, as IAs generativas podem gerar novas amostras que se assemelham aos dados originais.

A fim de prover uma compreensão elementar sobre o processo de aprendizado de máquina, a criação de modelos de algoritmos que analisam e interpretam dados, fundamenta-se em três abordagens distintas: Método de Aprendizado Supervisionado - dados de entrada e saída são manualmente inseridos para treinar uma rede neural artificial, permitindo que a máquina reconheça padrões e aprenda a chegar a determinadas saídas; Método de Aprendizado Não Supervisionado - processa dados acumulados, agrupando amostras com base em características sem conhecimento prévio das classificações; Método de Aprendizado por Reforço - envolve a repetição de tentativas e erros para alcançar objetivos, com a rede "aprendendo" com recompensas e penalidades.

Redes neurais artificiais (ANNs) (IBM, [2024]), também conhecidas como redes neurais simuladas (SNNs), são essenciais no aprendizado de máquina, inspiradas no funcionamento do cérebro humano. Compostas por camadas de nós interconectados, incluindo entrada, ocultas e saída, essas redes dependem de dados de treinamento para aprimorar sua precisão. À medida que são ajustadas para aumentar a precisão, tornam-se ferramentas

poderosas em ciência da computação, facilitando tarefas como reconhecimento de fala e identificação de imagens.

O aprendizado profundo ocorre na construção de modelos de redes neurais com múltiplas entradas e camadas intermediárias. Esses modelos passam por fases supervisionadas ou não, chamadas *feedforward* e *backpropagation*, buscando ajustes precisos nos parâmetros para identificação e reconhecimento de padrões específicos. Após a criação do modelo, o código é alimentado com um banco de dados, e o processo passa por ajustes manuais e testes para otimização dos parâmetros. Em resumo, o processo de ensinar um computador a reconhecer padrões, como gatos em fotos, envolve a criação de uma rede neural, o treinamento com dados e ajustes finos para garantir o reconhecimento preciso em novas situações.

### 3 MÉTODOS E RESULTADOS

O método aplicado nesta pesquisa possui duas funções distintas para embasar a discussão sobre inteligência artificial através do olhar habermasiano. A primeira função é a abordagem teórica através da Teoria do Agir Comunicativo. A segunda função é a introdução do resultado de pesquisas através do método de revisão sistemática, a fim de reunir literatura relevante disponível em bases de dados.

A busca foi realizada na base de dados Scopus em agosto de 2024, utilizando como recurso a busca avançada. Não há delimitação temporal nem de tipo de documento. Os dados são tratados e organizados em quadros com as principais informações (autores, título, ano, título de periódico e resumo, além das palavras-chave utilizadas pelos autores e, quando há, as utilizadas pela base e as referências de Habermas utilizadas em cada artigo.

O documento será eliminado se, na análise dos artigos, for identificada a não pertinência para esta pesquisa. Os termos utilizados na estratégia de busca estão organizados no Quadro 1.

**Quadro 1** - Estratégia de busca

"artificial intelligence" AND "Habermas" AND "Theoryofcommunicativeaction" OR "communicativeaction"
--

Fonte: Dados da pesquisa.

A busca resultou em 7 registros, sendo 3 artigos de periódicos e 4 trabalhos publicados em anais de evento ou em revistas que publicaram. Ou seja, todos artigos. Não houve duplicidade de documentos.

O Quadro 2 apresenta as informações dos artigos e as referências de Habermas utilizadas em cada artigo.

**Quadro 2** - Registros recuperados e referências de Habermas utilizadas

<b>Autor(es)/Título</b>	<b>Referências de Habermas</b>
1. ROSER, Andreas. Warum sprechen Menschen mit Maschinen? <b>Information - Wissenschaft &amp; Praxis</b> , [S.L.], v. 69, n. 5-6, p. 249-256, 6 nov. 2018.	Habermas J., Theorie des Kommunikativen Handelns, 1/2, (1971)
2. WANG, Hao. Algorithmic Colonization of Love. <b>Techné: Research in Philosophy and Technology</b> , [S.L.], v. 27, n. 2, p. 260-280, 2023.	Habermas J., Technology and Science as 'Ideology, Toward a Rational Society, pp. 81-127, (1970); Habermas J., The Theory of Communicative Action: Lifeworld and Systems, a Critique of Functionalist Reason, 2, (1987)
3. NATALINA, Nataliia. Digital communication transformation as a factor of delegitimization of political institutions. <b>European Political And Law Discourse</b> , [S.L.], v. 8, n. 5, p. 132-139, 2021.	Habermas J., The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society, (1989)
4. DIETZ, Jan L. G.. Understanding and Modelling Business Processes with DEMO. <b>Lecture Notes In Computer Science</b> , [S.L.], p. 188-202, 1999.	Habermas J., Theorie Des Kommunikatives Handelns, Erster Band, (1981)
5. FETZER, Anita. Non-acceptances: re- or un-creating context?. <b>Lecture Notes In Computer Science</b> , [S.L.], p. 133-144, 1999.	Habermas J., Theorie Des Kommunikativen Handelns, (1987)
6. LEMAÎTRE, Christian; FALLAH-SEGHROUCHNI, Amal El. A Multiagent Systems Theory of Meaning Based on the Habermas/ Bühler Communicative Action Theory. <b>Lecture Notes In Computer Science</b> , [S.L.], p. 116-125, 2000.	Habermas J., Postmetaphysical Thinking, (1996)
7. RICHARDSON, Sandra M.; COURTNEY, James F.; WAGNER, Gerald R. DISCOMAP: A System to Support Distributed Cognition in Inquiring Organizations. In: AMERICAS CONFERENCE ON INFORMATION SYSTEMS, 11., 2005. AMCIS 2005: A Conference on a Human Scale. <b>Proceedings...</b> AMCIS, 2005.	Habermas J., The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society, (1989)

Fonte: Dados da pesquisa.

As pesquisas relacionadas no Quadro 2 estão presentes em publicações da área de informação, direito e tecnologia. Já os artigos originalmente apresentados em eventos e publicados posteriormente em periódico são da área de ciência da computação. Os documentos mais antigos são do ano de 1999 (dois registros). Os demais documentos são de: 2000, 2005, 2018, 2021 e 2023, com um registro cada um. Outra informação sobre as publicações é que, embora de uma base de dados com diversos documentos com acesso restrito a assinaturas, 3 artigos são de acesso livre, sendo os três de periódicos. Esse dado corrobora com a popularidade do periódico científico na academia, uma vez que apresenta resultados mais rápidos se comparado aos livros, são os documentos mais comumente recuperados em bases de dados e



#### 4 DISCUSSÃO

Esta seção realiza a análise dos textos recuperados e são apresentados os principais objetivos e resultados de cada documento. As análises serão sustentadas na teoria do agir comunicativo de Habermas.

**Artigo 1:** Roser (2018) questiona o motivo das pessoas falarem com as máquinas. Para o autor, a comunicação entre as pessoas com máquinas ocorre uma vez que estas parecem atuar como seus parceiros de comunicação de escolha para diferentes objetivos de comunicação. Os robôs de fala, ou *chatbots*, se comunicam com os consumidores de forma consensual, independentemente das pessoas envolvidas. Essa comunicação ocorre de forma não violenta, sem discrepâncias, tematicamente abertos e ilimitados no tempo.

O autor afirma que Habermas não percebeu na teoria da ação comunicativa o perigo de ser aplicável à comunicação com máquinas e as consequências não triviais para sua análise de aplicação. A TAC discute a relação entre tecnologia, sociedade e racionalidade sob um ponto de vista filosófico e sociológico, sendo o foco a moralidade, a ética, a moralidade e as implicações sociais das tecnologias. O século XX e a era da informação são marcados pelo avanço acelerado das tecnologias da comunicação e informação. Dessa forma, para Habermas prever como estaria o uso nas últimas décadas seria utópico.

**Artigo 2:** No artigo seguinte, Wang (2023) discute a projeção do amor e o uso da inteligência artificial nos aplicativos de relacionamentos. A função da IA é identificar, em um grande banco de dados, perfis com correspondência com o usuário, o famoso "match" de um dos aplicativos mais populares deste segmento. De acordo com a pesquisa, as relações amorosas são invadidas por regras criadas por um algoritmo. Esse algoritmo delega a função de tomada de decisões sobre a vida amorosa das pessoas e teve como base a tese de colonização do mundo da vida. A pesquisa de Wang é centrada na colonização do mundo da vida de Habermas.

A tese da colonização do mundo consiste no resultado de profundo processo de argumentação no qual é reconstruído e apropriado de críticas de princípios teóricos de filosofia e ciências sociais (Siebeneichler, 2018). A colonização do mundo da vida é quando a lógica do sistema – como a economia, e a tecnologia - invade e domina o mundo da vida. Esse domínio pode ocorrer na esfera pública, na cultura e nas relações sociais, como o caso dos aplicativos de relacionamento. Entre as consequências da colonização do mundo da vida estão a perda da autonomia individual, a desintegração da comunidade, o declínio da esfera pública, a dominância da lógica do sistema e a perda de significado e do propósito (Habermas 1996; 2000; 2012).

**Artigo 3:** A transformação da comunicação em rede digital como um dos fatores que diminuem a confiança nas instituições políticas no contexto da pandemia global de Covid-19 é o foco da pesquisa de Natalina (2021). Entre os autores utilizados para a base teórica do estudo, a teoria da ação comunicativa de Habermas é a primeira citada. É sugerido que o uso da inteligência artificial restringe a liberdade de comunicação digital. Além disso, outros agentes limitadores das liberdades digitais são a censura de redes sociais por gigantes da tecnologia e a transferência de mensagens, por vezes manipuladas, em canais fechados que ocorrem nos aplicativos mensageiros. Essas atividades restringem o potencial de discussão e consenso como pré-requisito para a democracia representativa.

A Teoria do Agir Comunicativo de Habermas oferece uma abordagem rigorosa para entender o consenso como um objetivo fundamental da comunicação humana (Habermas, 2012). Para alcançar um consenso racional entre os participantes é necessária a compreensão mútua e o acordo sobre o significado das ações e declarações: através do bloqueio ou censura das redes sociais e envio unilateral de mensagens falsas ou manipuladas não é possível que exista o consenso. A informação é parte da comunicação nas interações mediadas pela linguagem (Lima, Gonçalves, Maia, 2023) então, ao ter informações manipuladas ou negadas, não há uma comunicação eficiente.

**Artigo 4:** No artigo seguinte, Dietz (1999) apresenta uma metodologia para modelar, (re)projetar e (re)engenharia de organizações. Esse recurso, denominado DEMO (Dynamic Essential Modeling of Organizations), DEMO é uma perspectiva relativamente nova e promissora sobre processos de negócios e sistemas de informação, chamada de Perspectiva Linguagem/Ação. Possui base teórica em três fontes científicas de inspiração. A primeira é a Teoria da Ação Comunicativa de Habermas. As outras duas são a Escada Semiótica de Stamper e a Ontologia de Bunge.

Ao longo da pesquisa, o autor demonstra a aplicação da ferramenta em cerca de 50 projetos de diferentes tipos em várias organizações. Em todos, o fator de sucesso foi o mesmo. Embora a inteligência artificial não tenha sido diretamente abordada nesta pesquisa, observa-se que a TAC foi associada ao uso de tecnologias.

**Artigo 5:** Fetzer (1999) começa o texto discutindo a complexidade do significado de "NÃO" na linguagem natural. Assim, se no discurso em linguagem natural consistisse apenas na transmissão de informações, tanto dizer NÃO quanto interpretar o significado comunicativo do NÃO seria uma questão bastante direta, pois o ato de comunicação NÃO apresentaria informações explícitas, como eu não concordo com você, eu rejeito seu convite ou você não está dizendo a verdade. No entanto, no discurso da vida real, o ato de comunicação NÃO

difícilmente é realizado abertamente no registro, mas frequentemente mitigado para evitar potenciais ameaças aos desejos/necessidades dos participantes.

Assim, as Não Aceitações ou Recusas (Non-acceptances) são discutidas na teoria dos atos de fala, lógica e conversação e análise de discurso com referência especial a como o contexto é acomodado. A TAC, nesta pesquisa, atua na sistematização dos resultados e na estrutura de alegações de validade positiva/negativa baseadas na contextualização da teoria de ações comunicativas de Habermas. Aqui, novamente é aplicada o conceito da linguagem como meio de comunicação e entendimento mútuo para a tentativa de um consenso.

**Artigo 6:** Lemaître e Fallah-Seghrouchni (2000) propõe, com base na teoria da Ação Comunicativa de Habermas/Bühler, uma nova estrutura que vai além da teoria clássica do ato de fala e suas interpretações intencionalistas. É introduzida uma teoria abrangente do significado para atos de comunicação, assumindo que o conteúdo de enunciados da linguagem natural pode ser classificado em três domínios diferentes do discurso, cada um com um tipo diferente de validação semântica: o domínio dos fatos objetivos, o domínio interno ou subjetivo do remetente e o domínio relacional social do remetente e do destinatário.

Seguindo Habermas, é introduzido uma mudança crucial na abordagem da interação do agente, focando nas questões de controle da conversação, no destinatário e não no remetente. Afirmamos que essas duas novas abordagens de interações multiagentes permitirão controlar e gerenciar as interações complexas entre agentes em aplicações abertas do mundo real. Neste registro não houve na listagem das referências à obra específica de Habermas, teoria do agir comunicativo: todavia, a TAC é citada ao longo do texto uma vez que a obra consultada foi O Pensamento Pós-Metafísico.

**Artigo 7:** Por fim, Richardson, Courtney e Wagner (2005) apresentam o DISCOMAP. Trata-se de um sistema que testa os princípios de design revisados usando fóruns de discussão e o The Planners Laboratory©, um novo pacote de software que fornece modelagem avançada, recursos gráficos e de rede para fornecer modelos compartilháveis com interfaces visuais envolventes para tomadores de decisão. O uso da teoria do agir comunicativo neste trabalho está ligada ao argumento que a comunicação é uma parte vital da aprendizagem organizacional e da cognição distribuída. Para o autor, a aprendizagem organizacional integra a TAC quando aborda a fala ideal e a comunicação discursiva de Habermas. Este trabalho não aborda diretamente a inteligência artificial mas faz a interligação entre a TAC e o uso de tecnologias que realizam análises: ou seja, "máquinas" que executam atividades que poderiam ser executadas por humanos, neste caso, tomada de decisões.

A conexão que é possível inferir entre a inteligência artificial, Jürgen Habermas e a teoria do agir comunicativo é a possibilidade de discussões e análises profundas de como a tecnologia pode interagir ou interferir com a humanidade, racionalidade e a ética. Para Habermas (2012) a tecnologia deve se desenvolver e agir com o propósito de gerar comunicação e colaboração do ser humano sem que exista domínio ou controle durante as suas interações. Ou seja, um entendimento mútuo, responsável pelo entendimento da comunicação ativa e da ação coordenada entre os indivíduos.

Na análise dos artigos recuperados na busca na base de dados foi possível observar os benefícios que o uso da inteligência artificial pode trazer em diferentes processos, mas também as preocupações com a delegação de atividades humanas para as máquinas. Afinal, a máquina é capaz de dialogar além da sua programação? Para Habermas (2012) o desenvolvimento moral ocorre através do diálogo e as máquinas possuem uma capacidade limitada para tanto. Por exemplo, os chatbots possuem um número limitado de respostas que, quando eliminados, é oferecido o atendimento através de outro meio, como atendimento humano ou envio de e-mail. Além disso, muitas vezes não compreendem as perguntas, tendo que ser refeitas diversas vezes até que haja um entendimento.

Para Lima, Gonçalves e Maia (2023) o uso do discurso é uma forma especial de agir comunicativo com o fim de construir entendimento subjetivo. Ou seja, o discurso é o meio utilizado para a resolução de conflitos sobre algo no mundo, além de construir acordos teóricos e práticos. O discurso entre sujeitos tem ainda a função de validação pragmática de expressões e representações do mundo da vida.

Nos exemplos das obras cinematográficas apresentadas no início da pesquisa, as máquinas foram incapazes de considerar os interesses de todos os envolvidos: não apenas seguiram a missão que receberam como não houve a busca de igualdade e justiça na tomada de decisões através da ética do discurso. E ocorre justamente por essa capacidade limitada das máquinas. Situação similar ocorre no caso das análises de processos do INSS, quando a análise dos documentos não considerou outros fatores além do que havia programado. Dessa forma, o que deveria ter sido uma ação para agilizar o atendimento, acabou negando benefícios.

## **5 CONSIDERAÇÕES FINAIS**

A atualidade de Habermas e das suas teorias não são relacionadas diretamente aos usos e implicações da inteligência artificial na sociedade. As ideias do filósofo se conectam, de forma indireta, com a IA e proporcionam reflexões sobre o tema. A relação da sociedade com a tecnologia, a teoria do agir comunicativo e a ética do discurso são alguns pontos que ilustram a afirmação.

Os documentos recuperados através do uso de estratégia de busca específica na base de dados Scopus fornecem resultados para fomentar a discussão sobre o uso da teoria do agir comunicativo não só na inteligência artificial mas das tecnologias. Os trabalhos abordam diferentes aspectos da TAC, e a análise desses documentos possibilita reflexões críticas sustentadas em Habermas. As pesquisas tratam desde o uso de ferramentas com as tecnologias da IA até os impactos causados na sociedade através de uma "terceirização" da escolha de parceiros para relação amorosa, das conversas com máquinas e na disseminação unilateral de informações manipuladas.

A discussão sobre os usos e limites da inteligência artificial é um tema atual, tendo diferentes abordagens há décadas e sendo objeto de estudo de diferentes áreas do conhecimento. A comunicação, por exemplo, ao analisar produções cinematográficas, a filosofia ao debater os embates éticos e a ciência da computação com olhar técnico porém por vezes combinado com um debate crítico sobre as aplicações da IA: as áreas conversam entre si em diferentes perspectivas sobre o mesmo tema. Os registros recuperados representam a um viés específico, qual seja, a IA através do olhar habermasiano sustentado na teoria do agir comunicativo.

O que este artigo orienta é o potencial uso crítico da filosofia para os estudos da Inteligência Artificial, particularmente da filosofia da linguagem. O destaque para Habermas neste trabalho radicaliza abordagens teóricas e abre espaço para outros autores. Uma das questões que merece ser observada e discutida é a filosofia da técnica, com estudos e autores que têm longa tradição, por exemplo, com Martin Heidegger e os críticos da Escola de Frankfurt da primeira geração (Adorno, Horkheimer, Marcuse e outros).

## REFERÊNCIAS

DEEP Learning Book. **Introdução às Redes Neurais Convolucionais**. [2024a]. Capítulo 40. Disponível em: <https://www.deeplearningbook.com.br/introducao-as-redes-neurais-convolucionais/>. Acesso em: 25 nov. 2023.

DEEP Learning Book. **Introdução às Redes Adversárias Generativas** (Gans – Generative Adversarial Networks). [2024b]. Capítulo 54. Disponível em: <https://www.deeplearningbook.com.br/?s=redes+advers%C3%A1rias+generativas>. Acesso em: 25 nov. 2023.

DIETZ, Jan L. G. Understanding and Modelling Business Processes with DEMO. **Lecture Notes In Computer Science**, Cham, p. 188-202, 1999. Disponível em: [http://dx.doi.org/10.1007/3-540-47866-3\\_13](http://dx.doi.org/10.1007/3-540-47866-3_13). Acesso em: 13 set. 2024.

FETZER, Anita. Non-acceptances: re- orun-creating context?. **Lecture Notes In Computer Science**, Cham, p. 133-144, 1999. Disponível em: [http://dx.doi.org/10.1007/3-540-48315-2\\_11](http://dx.doi.org/10.1007/3-540-48315-2_11). Acesso em: 13 set. 2024.

GERCINA, Cristiane. INSS usa robôs e aumenta pedidos negados, diz sindicato. **Folha de S. Paulo**, São Paulo, 04 maio 2022. Disponível em: <https://www1.folha.uol.com.br/mercado/2022/05/inss-usa-robos-para-tentar-reduzir-fila-de-beneficios-diz-sindicato.shtml>. Acesso em: 10 ago. 2024.

GIGLIOTTI, Fátima. Os olhares da ficção científica cinematográfica para um futuro que já é aqui e agora. **ComCiência**, Campinas, n. 158, 2014. Disponível em: [https://comciencia.scielo.br/scielo.php?script=sci\\_arttext&pid=S1519-76542014000400010&lng=en](https://comciencia.scielo.br/scielo.php?script=sci_arttext&pid=S1519-76542014000400010&lng=en). Acesso em: 10 ago. 2024.

HABERMAS, Jürgen. **O discurso filosófico da modernidade**: doze lições. São Paulo: Martins Fontes, 2000. 540 p.

HABERMAS, Jürgen. **Teoria do agir comunicativo**. São Paulo: WMF Martins Fontes, 2012. 2v.

HABERMAS, Jürgen. **Between facts and norms**: contributions to a discourse theory of Law and democracy. Massachusetts : MIT Press, 1996. 631 p.

IBM. **O que são redes neurais**. [2004]. Disponível em: <https://www.ibm.com/br-pt/topics/neural-networks>. Acessado em 25 nov. 2023.

LEMAÎTRE, Christian; FALLAH-SEGHRUCHNI, Amal El. A Multiagent Systems Theory of Meaning Based on the Habermas/ Bühler Communicative Action Theory. **Lecture Notes In Computer Science**, Cham, p. 116-125, 2000. Disponível em: [http://dx.doi.org/10.1007/3-540-44399-1\\_13](http://dx.doi.org/10.1007/3-540-44399-1_13). Acesso em: 13 set. 2024.

LIMA, Clovis Ricardo; GONÇALVES, Márcio; REBELO MAIA, Mariangela. Notas para uma teoria crítica discursiva da informação. **Encontros Bibli**: revista eletrônica de biblioteconomia e ciência da informação, Florianópolis, v. 28, p. 1–16, 2023. DOI: 10.5007/1518-2924.2023.e92925. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/92925>. Acesso em: 28 ago. 2024.

LÜDER, Amanda. Robô que analisa processos do INSS causa aumento de recursos por indeferimento. **Globo News**, 07 out. 2022. Disponível em: <https://g1.globo.com/economia/noticia/2022/10/07/robo-que-analisa-processos-do-inss-causa-aumento-de-recursos-por-indeferimento.ghtml>. Acesso em: 10 ago. 2024.

MELO, Laís de. Servidores acusam INSS de utilizar robôs para análise de pedidos. **Jornal da Cidade.net**, Aracaju, 05 maio 2022. Disponível em: <https://www.jornaldacidade.net/cidades/2022/05/328029/servidores-acusam-inss-de-utilizar-robos-para-analise-de-ped.html>. Acesso em: 30 ago. 2022.

NATALINA, Nataliia. Digital communication transformation as a factor of delegitimization of political institutions. **European Political And Law Discourse**, Praga, v. 8, n. 5, p. 132-139, 2021. Disponível em: <http://dx.doi.org/10.46340/eppd.2021.8.5.15>. Acesso em: 13 set. 2024.

ORACLE. **O que é IA? Saiba mais sobre inteligência artificial**. 2021. Disponível em: <https://www.oracle.com/br/artificial-intelligence/what-is-ai/>. Acessado em 25 nov. 2023.

RICHARDSON, Sandra M.; COURTNEY, James F.; WAGNER, Gerald R. DISCOMAP: A System to Support Distributed Cognition in Inquiring Organizations. In: AMERICAS CONFERENCE ON INFORMATION SYSTEMS, 11., 2005. AMCIS 2005: A Conference on a Human Scale. **Proceedings...** AMCIS, 2005. Disponível em: <http://aisel.aisnet.org/amcis2005/57>. Acesso em: 13 set. 2024.

ROSER, Andreas. Warumsprechen Menschen mit Maschinen? **Information – Wissenschaft & Praxis**, Berlin, v. 69, n. 5-6, p. 249-256, 6 nov. 2018. Disponível em: <http://dx.doi.org/10.1515/iwp-2018-0041>. Acesso em: 13 set. 2024.

SIEBENEICHLER, Flávio Beno. Mundo da vida e sistema na teoria do agir comunicativo. **Logeion: Filosofia da Informação**, Rio de Janeiro, v. 5, p. 27-36, 11 nov. 2018. Disponível em: <https://revista.ibict.br/fiinf/article/view/4498>. Acesso em: 30 set. 2024.

WANG, Hao. Algorithmic Colonization of Love. **Techné: Research in Philosophy and Technology**, Charlottesville, v. 27, n. 2, p. 260-280, 2023. Disponível em: <http://dx.doi.org/10.5840/techne202381181>. Acesso em: 13 set. 2024.