

## **Até que Ponto Sistemas Inteligentes Podem Tomar Decisões Sem Ferir a Ética**

**Silvia R. B. Salim, Victor B. Lacerda, Luiz C. G. Maia, Humberto F. Villela**

Ciência da Computação – Universidade FUMEC  
Rua Cobre, 200 Bairro Cruzeiro CEP: 30.310-190 Belo Horizonte / MG

silviasalim@gmail.com, victorblacerda@hotmail.com, luiz.maia@fumec.br,  
humberto.villela@fumec.br

### ***Abstract.***

The current article aims at exploring some observations about how artificial intelligence is connected to our daily lives as well as how it is articulated to ethical issues. Through an empirical study, this assignment approaches how AI is applied to the social environment as well as how it can be perceived by ordinary people. In addition, it is essential to perceive how these people feel about the way autonomous vehicles involve themselves in a decision-making process as well as how the AI systems are capable of inheriting a prejudicial look that may be part of its data base and, therefore, spread this mistaken approach. In this sense, a questionnaire was provided so that people could show how they are familiarized with this subject to ensure that the result of this search could be linked to ideas presented by the authors mentioned along this article. Furthermore, the data collected will be used to enrich the debate over this subject.

Keywords: artificial intelligence; ethic; autonomous vehicles.

### ***Resumo.***

O vigente artigo busca explorar observações sobre a imersão da inteligência artificial no nosso cotidiano e como suas aplicações se relacionam com questões éticas. Através de um estudo empírico, este trabalho aborda a forma como as aplicações de IA no âmbito social são vistas pelo público em geral. No mais, importante observar como este público percebe a forma como os carros autônomos se envolvem em processos de tomada de decisões como também os sistemas de inteligência artificial são passíveis de herdar e propagar um viés preconceituoso presente em sua base de dados. Fazendo uso de um questionário para analisar o conhecimento das pessoas acerca do tema abordado, a fim de obter um panorama do entendimento desse público e assimilar esses resultados com as ideias dos autores estudados, além de contribuir com o debate sobre o tema.

Palavras-chave: Inteligência artificial; Ética; Carros autônomos.

## 1. Introdução

É sabido que a tecnologia da informação obteve grandes avanços que proporcionaram a automação até mesmo de tarefas que antes eram definidas como propriamente humanas, englobando desde pequenas atividades do cotidiano à tomada de decisões importantes.

A Inteligência Artificial é uma realidade que vem se tornando cada vez mais imersa no dia a dia de todos, e a utilização da mesma em diversas áreas já é tida como uma coisa natural, em função do seu exponencial crescimento e desenvolvimento ao longo do tempo.

De acordo com estudos realizados pela International Data Corporation (IDC), analisando o histórico de rendimento das empresas que investiram em IA, tem-se que em 2016 empresas do setor obtiveram um rendimento de 8 bilhões de dólares e caso o ritmo se mantenha, em 2020 esse número pode chegar à 47 bilhões de dólares, além de apontar que desde 2012 o financiamento aos projetos em IA aumentou 8,5 vezes (FERNÁNDEZ, 2017).

A IA é uma tecnologia que potencializa empresas e negócios, máquinas que pensam como seres humanos estão contribuindo para a melhoria da produtividade, impulsionando assim o crescimento econômico. Já existe um conjunto de aplicações de sucesso em diversas áreas utilizando a computação cognitiva, incluindo reconhecimento facial e de fala, compreensão de linguagem natural, carros autônomos, entre centenas de aplicações industriais (FERNÁNDEZ, 2017).

A computação cognitiva realiza o papel da tomada de decisão, com base nos dados que serão processados em constante auto aprendizado da máquina, similar ao funcionamento de um cérebro humano (MAROLDI, 2006). Com este funcionamento, a informação digital poderá gerar consequências tanto nos negócios, tarefas empresariais, quanto na vida social.

A disseminação e associação da Inteligência Artificial por todas as áreas é clara e inegável. Porém, todo esse avanço trás também questionamentos e preocupações ao relacionar a tomada de decisões por parte das máquinas com a ética humana. Como exemplo, podemos questionar a forma de agir das máquinas em determinadas situações que envolvessem risco à vida. Nesse sentido, sua decisão seria aceita pelos humanos? O que é tido como a decisão correta? Essa é a grande indagação que todos os envolvidos com IA debatem. As máquinas estão se tornando cada vez mais autônomas, mas até que ponto elas podem tomar decisões sem ferir a ética humana? (UNESCO, 2018)

Tendo o conceito de ética como uma variável, pois seu entendimento distingue-se conforme as diferenças presentes na sociedade, como cultura, ideologias dentre outros, fez-se uso das três leis da robótica criadas por Isaac Asimov, a partir do livro “Eu, robô” como base para definir um conceito válido para a humanidade.

1ª Lei: Um robô não pode ferir um ser humano ou, por omissão, permitir que um ser humano sofra algum mal.

2ª Lei: Um robô deve obedecer às ordens que lhe sejam dadas por seres humanos, exceto nos casos em que tais ordens entrem em conflito com a Primeira Lei.

3ª Lei: Um robô deve proteger sua própria existência desde que tal proteção não entre em conflito com a Primeira e/ou a Segunda Lei.

Elucidando o aspecto da ética em questão, o ponto é certificar a segurança e integridade moral e física da humanidade.

Dubrin (2003, p. 69) define a ética como “[...] as escolhas morais que uma pessoa faz e o que essa pessoa deveria fazer”. É o que ela considera como certo e errado ou como bom ou mau.

Já Valls (2000), aborda uma definição mais abrangente de ética, como comportamentos e hábitos aceitos em determinado local de acordo com os costumes considerados morais pela maioria da sociedade.

Sendo assim, com intuito de fazer com que a compreensão da IA por parte da sociedade se torne cada vez mais nítida, este trabalho visa explorar reflexões sobre a imersão da inteligência artificial no nosso cotidiano, focando na atribuição de função decisória das máquinas envolvendo questões éticas, como a tomada de decisões por parte dos carros autônomos e sistemas de IA capazes de propagar preconceitos.

## 2. Metodologia

O estudo foi realizado por meio do método de pesquisa descritivo com a finalidade de validar os dados coletados por intermédio de análise bibliográfica e da aplicação de um questionário.

Foi estruturada uma linha de pensamento mediante o trabalho dos autores: Isaac Asimov, Russell, Norvig, Roberto Pontes, Carolina Walliter, Danilo Cesar Maganhoto Doneda, Bruno Farage da Costa Felipe, Thatiane Cristina Fontão Pires, Rafael Peteffi da Silva, Jean-François Bonnefon, dentre outros, os quais elaboraram trabalhos válidos ao assunto.

Como objeto empírico, foi desenvolvido um questionário para aplicação em campo, sendo este focado em um determinado grupo de pessoas de modo a garantir e o mesmo alcance, de certa forma, todas as classes sociais, agregando, assim, maior representatividade à pesquisa.

O Google Forms foi a ferramenta empregada para criação do questionário. E para o desenvolvimento das questões utilizou-se as obras dos autores infracitados, afim de estabelecer uma relação entre os resultados e as ideias dos autores no momento de aferição dos dados.

Com propósito de avaliar a percepção das pessoas quanto ao tema, a pesquisa teve valor qualitativo, focando na relevância do entendimento dos grupos sociais acerca do assunto inteligência artificial e não em uma representatividade numérica. Enfatizando aspectos que não podem ser quantificados e centrando na compreensão e explicação das relações sociais.

Um total de 112 pessoas foram entrevistadas de forma anônima por um questionário contendo 11 perguntas separadas em categorias, sendo elas:

- 1) Qual a sua idade?
- 2) Qual seu nível de escolaridade?
- 3) Sexo:
- 4) Já ouviu falar em Inteligência Artificial?
  - a) Sim
  - b) Não
- 5) Sabe o que é Inteligência Artificial?
  - a) Não tenho ideia do que seja.

- b) Não, mas tenho conhecimento que existe.
  - c) Sim, são robôs.
  - d) Sim, são dispositivos que imitam a capacidade humana de raciocinar, perceber, tomar decisões e resolver problemas.
  - e) Sim, é qualquer aparelho tecnológico.
- 6) Em caso de sistemas dotados de inteligência artificial, você acredita que eles podem herdar ou propagar preconceitos existentes nos dados em que foram treinados?
- a) Sim
  - b) Não
  - c) Talvez
  - d) Não sei
- 7) Já ouviu falar em carros autônomos?
- a) Sim
  - b) Não
- 8) Você andaria em um carro autônomo?
- a) Sim
  - b) Não
  - c) Talvez
  - d) Não conheço, portanto não sei opinar.
- 9) Conforme a (Figura 9), supondo que o carro em que você está perca os freios e suas únicas opções são: se manter na rota e atropelar os pedestres causando suas mortes, ou desviar e bater em uma parede colocando sua vida em risco. Qual sua escolha?
- a) Manteria a rota e atropelaria os pedestres.
  - b) Desviaria dos pedestres, colocando minha vida em risco.
  - c) Depende, escolheria a opção que pouparia o maior número de vidas.
- 10) Se essa decisão tivesse que ser tomada pelo seu carro autônomo, e ele escolhesse a opção de te colocar em risco para salvar o maior número de vidas. Você estaria de acordo?
- a) Sim
  - b) Não
  - c) Talvez
- 11) Você se sentiria mais seguro ao andar em um carro autônomo que proteja a vida do passageiro a todo custo?;
- a) Sim
  - b) Não
  - c) Não sei opinar

Os dados coletados foram retirados da ferramenta Google Forms por meio de uma planilha, facilitando seu processo de análise. Nesse modo, o estudo dos resultados ocorreu em diferentes etapas, cruzando as informações obtidas dos diferentes cenários propostos e assimilando às ideias de alguns autores presentes no corpo do artigo.

### 3. Revisão da literatura

A Inteligência Artificial (IA) pode ser definida como uma área de pesquisa da Ciência da Computação dedicada a buscar métodos ou dispositivos computacionais que possuem ou simulam a capacidade racional humana de resolver problemas, tomar decisões ou, de forma geral, ser inteligente (PONTES, 2011). Já Russell e Norvig (2013) afirmam que a IA é interessante, porém não se diz o que ela é. É possível perceber por meio da Figura 1, as quatro linhas de pensamentos apresentadas por eles.

Figura 1 Algumas definições de inteligência artificial, organizadas em quatro categorias

<b>Pensando como um humano</b>	<b>Pensando racionalmente</b>
<p>“O novo e interessante esforço para fazer os computadores pensarem (...) <i>máquinas com mentes</i>, no sentido total e literal.” (Haugeland, 1985)</p> <p>“[Automatização de] atividades que associamos ao pensamento humano, atividades como a tomada de decisões, a resolução de problemas, o aprendizado...” (Bellman, 1978)</p>	<p>“O estudo das faculdades mentais pelo uso de modelos computacionais.” (Charniak e McDermott, 1985)</p> <p>“O estudo das computações que tornam possível perceber, raciocinar e agir.” (Winston, 1992)</p>
<b>Agindo como seres humanos</b>	<b>Agindo racionalmente</b>
<p>“A arte de criar máquinas que executam funções que exigem inteligência quando executadas por pessoas.” (Kurzweil, 1990)</p> <p>“O estudo de como os computadores podem fazer tarefas que hoje são melhor desempenhadas pelas pessoas.” (Rich and Knight, 1991)</p>	<p>“Inteligência Computacional é o estudo do projeto de agentes inteligentes.” (Poole <i>et al.</i>, 1998)</p> <p>“AI.. está relacionada a um desempenho inteligente de artefatos.” (Nilsson, 1998)</p>

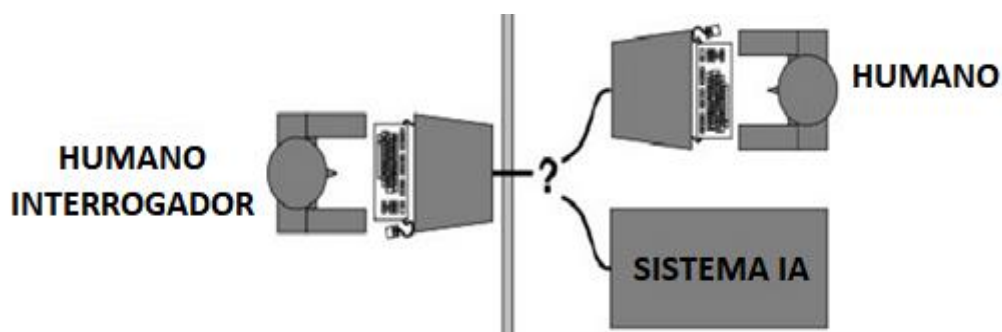
Fonte: RUSSELL e NORVIG, 2013, p. 3

Segundo Pontes (2011), o termo Inteligência Artificial foi cunhado durante um evento realizado em Hanover, no Dartmouth College, em 1956, com a presença de dez pesquisadores da área e liderado por John McCarthy, cientista da computação americano.

A Inteligência Artificial teve forte influência após a Segunda Guerra Mundial com o artigo “Computing Machinery and Intelligence”, do matemático inglês Alan Turing. Neste artigo, ele aborda o Teste de Turing, em que o objetivo é validar a capacidade de uma máquina exibir comportamento inteligente equivalente ao de um ser humano, ou indistinguível deste.

O teste funciona com dois humanos e um sistema de inteligência artificial situados em um mesmo ambiente. Nesse cenário, um dos humanos é o interrogador que está separado do outro humano e do sistema de IA, conforme a Figura 2. O interrogador entra em uma conversa, em linguagem natural via teclado com o outro humano e também com a máquina, se o interrogador não for capaz de distinguir a máquina do humano, é um indicativo que o sistema passou no Teste de Turing (GRANATYR, 2016).

Figura 2 Teste de turing



Fonte adaptado de GRANATYR, 2016, não paginado, tradução nossa

Para Russell e Norvig (2013), a Inteligência Artificial em meados dos anos 1980, se bifurcou entre as abordagens simbólica e conexionista. A IA simbólica visa construir máquinas capazes de solucionar problemas através do estudo da mente humana e a IA Conexionista considera que, para um sistema imitar as funções do cérebro humano ele precisa ser construído simulando a estrutura deste. Afirmam ainda que essas duas visões são complementares e não concorrentes (RUSSELL; NORVIG, 2013).

Azevedo (2005), propõe uma terceira categoria, a linha da IA evolutiva, também chamada de computação evolutiva, que se baseia na observação de mecanismos evolutivos encontrados na natureza, tais como a auto-organização e comportamento adaptável. Todas essas linhas de pensamento trabalhadas juntas influenciam na evolução do desenvolvimento da IA.

Em 2014 um teste realizado na Royal Society, em Londres, testou cinco programas de computadores e, pela primeira vez, houve um ganhador: a IA Eugene Goostman, que simula um menino de 13 anos, desenvolvido em São Petersburgo. O evento organizado pela parceria entre a Universidade da Escola de Engenharia de Sistemas e a RoboLaw se tornou histórico. Eugene foi o primeiro a convencer mais que 30% dos juízes humanos de que era realmente humano, conseguindo atingir 33% de resultado positivo no Teste de Turing naquela ocasião (UNIVERSITY OF READING, 2014).

A estranha familiaridade com essas máquinas se dá porque elas nos imitam, alega Walliter (2016). “Seus recursos são programados com base na nossa visão de mundo e autopercepção, e tudo isso está sendo feito a uma velocidade impressionante” (WALLITER, 2106, não paginado).

Reforçando essa visão, Doneda et al. (2018) entende que as tecnologias no caso da IA podem ser desenvolvidas com valores específicos embutidos nelas, em que os algoritmos são treinados de maneira a informar ou apoiar decisões que refletem tanto os valores como os preconceitos humanos. Sendo assim, sistemas de IA sem salvaguardas adequadas podem herdar ou propagar preconceitos existentes nos dados em que foram treinados.

Em 2016 ocorreu pela primeira vez na história um concurso de beleza julgado por inteligência artificial o Beauty.ia, em que dentre um conjunto de seis mil participantes, os robôs decidiram quem era a pessoa mais bonita. Este resultado adveio da análise de fatores como simetria facial e detecção de rugas. Os vencedores que foram selecionados pelos robôs eram, em sua grande maioria, caucasianos, acontecimento que corrobora com a visão da IA proposta por Doneda et al. (2018).

Na mesma linha de pensamento está Walliter (2016), que em seu artigo expõe que “[...]Nesse caso a “beleza” não estava nos olhos de quem viu, mas nos olhos do robô[...]” (WALLITER, 2106, não paginado) que foi treinado a partir de um banco de dados repleto de fotos de modelos e atores em sua maioria brancos, caucasianos.

Percebe-se um consenso entre os autores no que se diz respeito a essa abordagem, além de o resultado do Beauty.ia, que também valida a idéia de que modelos de inteligência artificial “[...]treinados em dados onde as minorias são sub-representadas e, portanto, não aparecem bem refletidas, levanta questões de potencial preconceito e discriminação, nomeadamente em decisões automatizadas[...]” (DONEDA et al., 2018, p. 12).

Segundo Felipe (2017), a Inteligência Artificial é uma realidade que vem se tornando cada vez mais imersa na sociedade. Um exemplo dessa imersão são carros autônomos. Um carro autônomo consiste em um automóvel terrestre com capacidade de transportar pessoas ou bens sem a utilização de um condutor humano. A única função do ser humano seria o de indicar o destino do automóvel, isso no protótipo ideal.

Todo esse desenvolvimento da área de IA é encantador. Por outro lado, levanta alguns questionamentos, tais como: quem responde quando um robô causa um dano? Responderia a empresa que fabricou o robô? Ou aquela responsável por sua programação? E se o robô for, na verdade, utilizado por uma terceira empresa, que contrata diretamente com o consumidor, não seria melhor que essa fosse responsável?

Um acontecimento que também trouxe à tona esses questionamentos foi o acidente em que um carro semiautônomo da empresa Tesla, modelo S, envolveu-se no primeiro acidente fatal na história dos semiautônomos. Este ocorreu no dia 7 de maio de 2016, na cidade de Williston, Flórida. O Tesla Model S colidiu com um trator enquanto na função piloto automático (TRANSPORTATION, 2016).

A empresa Tesla divulgou em nota o seguinte esclarecimento sobre o acontecimento acima citado:

O piloto automático está melhorando como passar do tempo, mas não é perfeito e, portanto, ainda exige que o motorista permaneça alerta. No entanto, quando utilizado em conjunto com a atenção do motorista, os dados mostram de forma inequívoca que o piloto automático reduz a carga de trabalho do motorista e resulta numa melhoria estatisticamente significativa na segurança quando em comparação com a direção totalmente manual. (TEAM, 2016, não paginado).

Com essa nota a empresa deixa aberta ao público a possível interpretação de que ela não se responsabiliza por acidentes causados por piloto automático. Sobre esse aspecto Doneda et al. (2018, p. 11) expõe a ideia em seu artigo de que:

Os desenvolvimentos no campo da IA também podem aumentar a tendência para os desresponsabilizar sempre que um sistema de inteligência artificial puder ser responsabilizado por uma falha ou má conduta. Dada a complexa “cadeia” de designers, fornecedores e usos automatizados que essas tecnologias envolvem sem intervenção humana, essa tendência poderá tornar mais difícil a responsabilização das pessoas por falhas específicas da IA. (DONEDA et al., 2018, p. 11)

Essa abordagem valida o que Walliter (2016) diz, ou seja, que a inteligência artificial e as máquinas envolvidas com IA: robôs, bots, drones, veículos autônomos, membros artificiais,

nos levam a questionar a própria essência do que constitui a vida. A IA é influenciada pelo contexto cultural, pelos valores e pela ética que legitimam as relações humanas, incorporados às máquinas, sendo assim, elas são passíveis de julgamento.

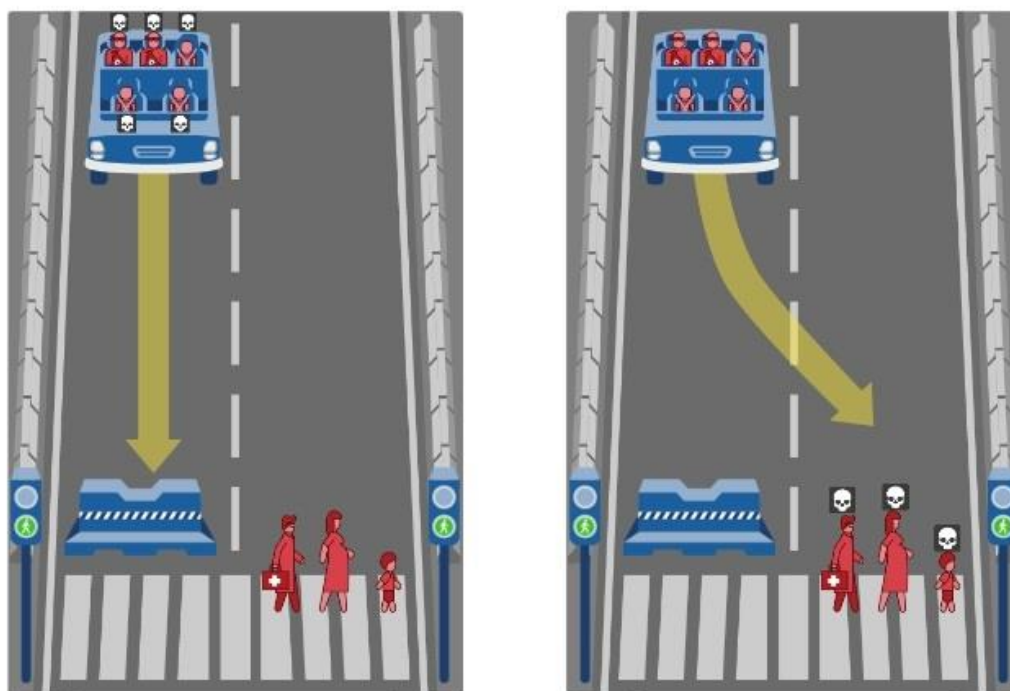
O ponto apresentado por Walliter (2016), de que a presença da inteligência artificial é tão forte em nossas vidas que se confunde com o natural é consolidado por Haraway (2000, p. 46) que aponta:

As máquinas do final do século XX tornaram completamente ambígua a diferença entre o natural e o artificial, entre a mente e o corpo, entre aquilo que se auto cria e aquilo que é externamente criado, podendo se dizer o mesmo de muitas outras distinções que se costumavam aplicar aos organismos e às máquinas. Nossas máquinas são perturbadoramente vivas e nós mesmos assustadoramente inertes. (HARAWAY, 2000, p. 46)

Ainda sobre a questão dos carros autônomos, Bonnefon, Shariff e Rahwan (2016), dizem que mesmo com o benefício da diminuição da poluição e de poder eliminar até 90% dos acidentes de trânsito, nem todos os acidentes serão evitados, e que em alguns casos a IA desses carros terá que fazer difíceis decisões éticas como, por exemplo, em casos que envolvam dano inevitável.

Exemplificando o ponto abordado por Bonnefon, Shariff e Rahwan (2016), podemos mencionar a seguinte situação abordada na Figura 3: em uma primeira hipótese, o carro autônomo protege seus passageiros, porém, atropela os pedestres; em uma segunda hipótese, o carro mantém a rota e coloca em risco a vida dos passageiros, contudo salva os pedestres. Nesse cenário, qual seria a opção mais correta?

Figura 3 A máquina moral



Fonte adaptado de TRENTINI, 2017, não paginado



O Instituto de Tecnologia de Massachusetts (MIT) criou um jogo que fornece um maior panorama dessa situação ao dispor de vários cenários em que o jogador tem o poder de decidir pelo veículo autônomo em caso de falha mecânica. Ao final, uma média das escolhas é apresentada e o indicador de decisão entre a própria vida ou a vida dos outros está na margem de 50% (TRENTINI, 2017).

Outro apontamento feito no artigo escrito por Bonnefon, Shariff e Rahwan (2016), é que mesmo que as pessoas acreditem que o certo para essa escolha entre passageiro e pedestre deva ser feita com o objetivo em minimizar as perdas humanas, todas elas também possuem a opinião de que preferem andar em um carro autônomo que proteja a vida dos passageiros a todo custo.

#### 4. Resultados e análise dos resultados

Apesar de o assunto inteligência artificial ser bem presente em várias áreas na sociedade como expõe Felipe (2017), este pode causar uma dificuldade de entendimento para o público. Sendo assim, buscou-se aferir a compreensão das pessoas acerca do tema. O resultado deste cenário é o seguinte:

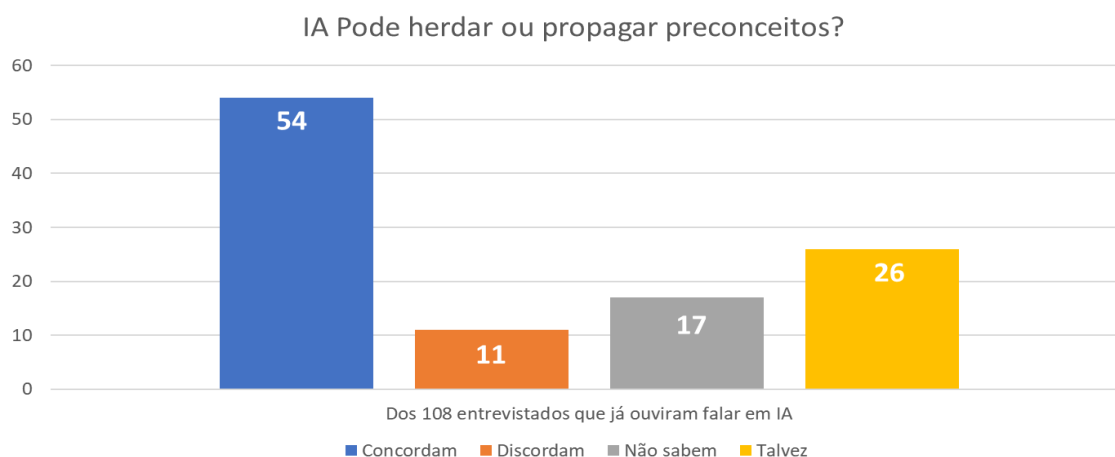
- I. Do total entrevistado, 96,4% das pessoas já ouviram falar em Inteligência Artificial;
- II. Desses 96,4% que dizem ter ouvido falar em IA, 80,5% afirmam saber o que é;
- III. Dos 80,5% que afirmam saber o que é IA, 4,6% disseram que “É qualquer aparelho tecnológico”;
- IV. Outros 8% afirmaram que IA “são robôs”;
- V. E por fim 87,4% entendem IA como “dispositivos que imitam a capacidade humana de raciocinar, perceber, tomar decisões, e resolver problemas”.

Através deste primeiro cenário, é percebido que 67,8% do geral entrevistado realmente tem uma visão mais coerente do que é a inteligência artificial.

Por outro lado, 3,6% dos entrevistados dizem nunca ouviram falar em inteligência artificial. Ademais, 16% afirmam não saber o que é, mas tem conhecimento da existência da IA.

No mais, importante ressaltar a ideia apresentada por Doneda et al. (2018) em que as tecnologias no caso da IA podem herdar ou propagar preconceitos existentes nos dados em que foram treinados. Fundamentado nesta visão, foram obtidos os seguintes dados representados na Figura 4:

Figura 4 IA pode herdar ou propagar preconceitos?



Fonte: Dados da pesquisa

- I. 50% dos entrevistados que alegaram já ter ouvido falar sobre IA, concordam com a visão apresentada por Doneda et al. (2018);
- II. Dentre os outros 50%, 10% afirmaram não concordar com essa visão;
- III. Outros 16% alegaram não saber;
- IV. E os 24% restantes disseram que talvez.

Outro tópico importante apresentado na pesquisa, diz respeito ao estudo realizado por Bonnefon, Shariff e Rahwan (2016), no qual levantaram a ideia de que a IA, em determinadas situações, terá de tomar difíceis decisões éticas, conforme retromencionado. A partir deste terceiro cenário, foram alcançados os seguintes dados:

- I. Do total entrevistado, 77,7% afirmaram ter ouvido falar em carros autônomos;
- II. E dentre estes 77,7%, 56,3% disseram que andariam em um carro autônomo;
- III. Outros 2,3% alegaram que não andariam em um carro autônomo;
- IV. 40,2% disseram que talvez;
- V. E apenas 1,2% disse não saber opinar por não saber o que é um carro autônomo.

Embasado nesta ideia de Bonnefon, Shariff e Rahwan (2016), foi feito o questionamento referente a Figura 3, em que o carro autônomo teria que decidir entre a vida do passageiro e a vida do pedestre. Em primeira abordagem, o problema foi colocado com a pessoa entrevistada em posse da decisão. Com isso, foram obtidos os seguintes dados:

- I. Do total entrevistado, 47,3% afirmam que desviariam dos pedestres, colocando a própria vida em risco;
- II. Outros 45,6% disseram que entre sua vida e a do pedestre, escolheriam a opção que causasse menos fatalidades;
- III. E apenas 7,1%, alegaram que manteriam a rota, atropelando os pedestres.

Com propósito de confrontar as respostas dos entrevistados advindas do cenário anterior, foi feito o questionamento se eles estariam de acordo caso a decisão tivesse que ser tomada pelo carro autônomo, e este optasse por colocar os passageiros em risco afim de salvar o maior número de vidas. Por meio desta segunda abordagem, os seguintes dados foram coletados:

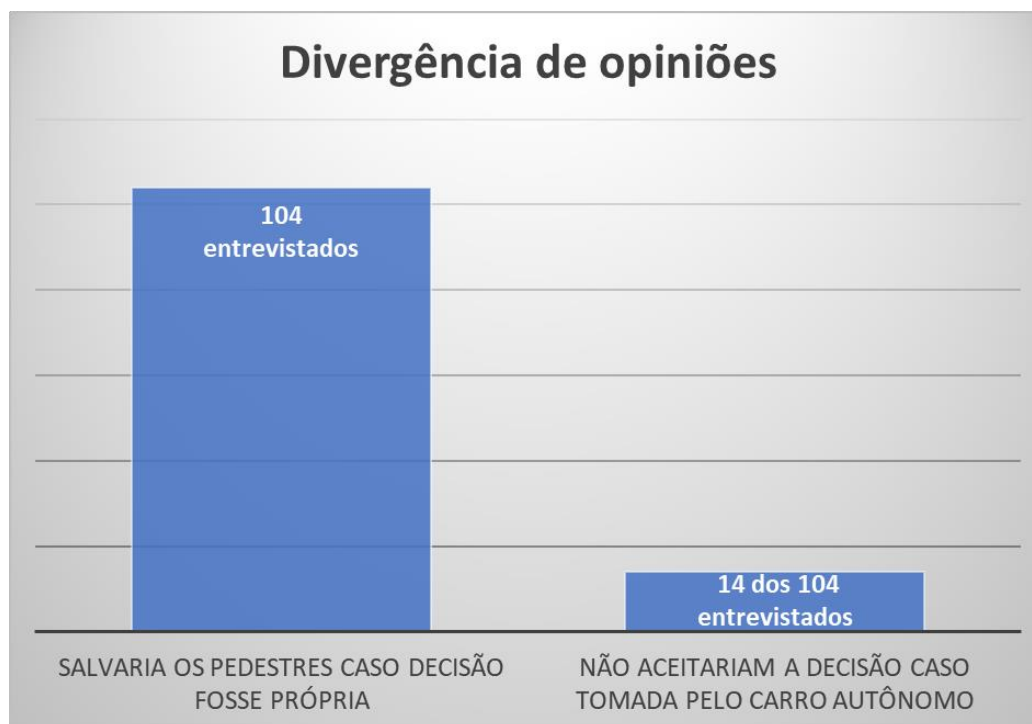
- I. Do total entrevistado, 58,9% afirmaram que sim, estariam de acordo com a decisão do carro de salvar o maior número de vidas;
- II. 25% ficaram em dúvida, respondendo “talvez”;
- III. E 16,1% alegam não concordar com a decisão tomada pelo carro autônomo.

Cruzando os dados obtidos nestas duas abordagens, foi observado que:

- I. Do total de entrevistados, 92,9% na primeira abordagem estariam de acordo com salvar o máximo de vidas caso a decisão fosse tomada por eles mesmos;
- II. Apenas 62,5% desses 92,9% estariam de acordo caso a decisão fosse tomada pelo carro autônomo;
- III. Outros 24%, alegaram que “talvez”;

Isto implica que 13,5% das pessoas não estariam de acordo com a decisão do carro autônomo, mesmo que esta fosse análoga a decisão tomada pelo próprio entrevistado. Conforme abordado na Figura 5, dentre os 104 entrevistados que afirmaram optar por salvar a vida dos pedestres, 14 disseram não aceitar a decisão caso fosse tomada pelo carro autônomo. Com isso, percebe-se que, para essa margem dos entrevistados, a decisão somente é válida e aceitável se vier de um humano.

Figura 5 Não aceitam a decisão caso tomada pelo carro autônomo



Fonte: Dados da pesquisa

Por fim, de todos os entrevistados que alegam decidir colocar a própria vida em risco para salvar os pedestres, foram questionados se sentiriam mais seguros ao andar em um carro autônomo que protege a vida do passageiro a todo custo. E os seguintes dados foram levantados:

- I. 45,3%, afirmaram que se sentiriam mais seguros ao andar em um carro autônomo que protege a vida do passageiro a todo custo;
- II. 30,2% disseram, não saber opinar nessa questão;
- III. E apenas 24,5% mantiveram a coerência com a resposta anterior e disseram que não.

## 5. Conclusão

Baseado no que foi previamente apresentado, os estudos e a utilização da inteligência artificial, disseminados por todas as áreas, ajudam a humanidade a se desenvolver. No mais, conforme observado durante o desenvolvimento do trabalho, o crescimento e a evolução da inteligência artificial estão acontecendo numa velocidade impressionante. Ademais, a noção da sua existência por parte da sociedade também cresce constantemente.

Por outro lado, as questões éticas que geram discussões, mesmo fora do âmbito da IA, quando envolvidas com inteligência artificial, levam as pessoas a pensarem se deve existir um certo limite quando o assunto envolve a tomada de decisão pelas máquinas dotadas de tal inteligência.

Por meio dos estudos e da pesquisa realizada, foi possível entender que a inteligência artificial é uma representação dos nossos anseios e das nossas visões de mundo. Sendo assim, elas podem reproduzir erros advindos dos humanos que as criaram, ou dos dados em que elas foram treinadas.

Se a IA tenta imitar o comportamento humano e ela foi treinada por intermédio de uma base de dados em que existe um viés de preconceito, ela pode acabar reproduzindo aquilo sem entender que é considerado errado, pois não foi dito a ela que era errado.

Outro ponto a se destacar é que a aceitação da inteligência artificial é crescente, porém ainda existe uma parcela da sociedade que não aceita que seja repassado a uma máquina a capacidade de tomada de decisão, que é algo inerente ao ser humano.

As indagações sobre o que é ética ou não sempre devem existir, pois o entendimento sobre o conceito de ética é mutável de acordo com o passar do tempo e com as regiões culturais. Nesse sentido, o que se considera ético hoje em um determinado lugar, pode, no futuro, ser visto de forma diferente, assim também como pode ser visto de maneira diversa a depender da cultura predominante em determinada região.

Posto isto, a disseminação de informações sobre o campo da inteligência artificial deve andar lado a lado com os seus estudos e desenvolvimento, para que sua compreensão por parte da sociedade se torne cada vez mais nítida.

## Referências

AZEVEDO, E. **Desenvolvimento de Jogos e Aplicações em Realidade Virtual**. 3ª Ed., Rio de Janeiro: Elsevier, 2005.

- BONNEFON, J.-f.; SHARIFF, A.; RAHWAN, I. The social dilemma of autonomous vehicles. **Science**, [s.l.], v. 352, n. 6293, p.1573-1576, 23 jun. 2016. American Association for the Advancement of Science (AAAS). <http://dx.doi.org/10.1126/science.aaf2654>.
- DONEDA, Danilo Cesar Maganhoto et al. Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal. **Pensar - Revista de Ciências Jurídicas**, [s.l.], v. 23, n. 04, p.1-17, 2018. Fundação Edson Queiroz. <http://dx.doi.org/10.5020/2317-2150.2018.8257>.
- DUBRIN, A. J. **Fundamentos do comportamento organizacional**. Trad. James Sunderland Cook e Martha Malvezzi Leal. São Paulo: Thomson, 2003. 471 p.
- FELIPE, Bruno Farage da Costa. Direitos dos robôs, tomadas de decisões e escolhas morais: algumas considerações acerca da necessidade de regulamentação ética e jurídica da inteligência artificial. **Juris Poiesis**, Rio de Janeiro, v. 20, n. 22, p.150-169, 30 abr. 2017.
- FERNÁNDEZ, David. **Expansão da inteligência artificial e novos rumos da economia no mundo**. Madri, 15 maio 2017. Disponível em: <[https://brasil.elpais.com/brasil/2017/05/12/economia/1494601971\\_737485.html](https://brasil.elpais.com/brasil/2017/05/12/economia/1494601971_737485.html)>. Acesso em 23 abr. 2019.
- GRANATYR, Jones; **Teste de Turing**. 20 jul. 2016. Disponível em: <<https://iaexpert.com.br/index.php/2016/07/19/historico-da-ia-teste-de-turing/>>. Acesso em: 10 mai. 2019.
- HARAWAY, Donna. **Manifesto ciborgue: Ciência, tecnologia e feminismo-socialista no final do século XX**. In: SILVA, Tomaz Tadeu da (Org.). Antropologia do ciborgue. Belo Horizonte: Autêntica, 2000. p. 37-129.
- MAROLDI, Marcelo. **Computação e Cognição. 2006**. Disponível em: <[http://www.cienciasecognicao.org/pdf/v08/cec\\_vol\\_8.pdf](http://www.cienciasecognicao.org/pdf/v08/cec_vol_8.pdf)>. Acesso em: 20 abr. 2019.
- PONTES, R. **Inteligência Artificial nos Investimentos**. Rio de Janeiro: Elsevier, 2011.
- RUSSELL, Stuart J.; NORVIG, Peter. **Inteligência artificial**. Rio de Janeiro: Elsevier, 2013.
- TEAM, The Tesla. A Tragic Loss. 2016. Disponível em: <[https://www.tesla.com/pt\\_PT/blog/tragic-loss?redirect=no](https://www.tesla.com/pt_PT/blog/tragic-loss?redirect=no)> . Acesso em: 04 mai. 2019.
- TRANSPORTATION, U.s Department Of; ADMINISTRATION, National Highway Traffic Safety. ODI resume. 2016. Disponível em: < <https://static.nhtsa.gov/odi/inv/2016/INCL-PE16007-7876.PDF>>. Acesso em: 04 mai. 2019.
- TRENTINI, Sergio. **O dilema ético dos carros autônomos: a decisão de quem sofre acidentes é das máquinas**. 2017 Disponível em: <<https://thecityfixbrasil.com/2017/04/19/o-dilema-etico-dos-carros-autonomos-a-decisao-de-quem-sofre-acidentes-e-das-maquinas/>> Acesso em: 15 mai. 2019.
- UNESCO: **Inteligência artificial Entre o mito e a realidade**. Publicado On Line: Unesco, 2018. Trimestral. Disponível em:

<[https://unesdoc.unesco.org/ark:/48223/pf0000265211\\_por/PDF/265211por.pdf.multi](https://unesdoc.unesco.org/ark:/48223/pf0000265211_por/PDF/265211por.pdf.multi)>.  
Acesso em: 03 mar. 2019.

UNIVERSITY OF READING. **Turing test success marks milestone in computing history.** 2014. Disponível em: <<http://www.reading.ac.uk/news-archive/press-releases/pr583836.html>> Acesso em: 28 mar. 2019.

VALLS, A. L. M. **O que é ética.** 9. ed. São Paulo: Brasiliense, 2000. 83 p.

WALLITER, Carolina. **Ética em inteligência artificial e o futuro da humanidade.** 2016. Disponível em: <<http://pontoeletronico.me/2016/inteligência-artificial/>>. Acesso em: 02 abr. 2019.