



# **FGA0221 - Inteligência Artificial**

## **Portfólio 01**

**Novembro, 2024**

## **Tema do portfólio:**

Introdução à inteligência artificial, histórico, estado da arte, benefícios e riscos. Agentes inteligentes, ambientes e racionalidade.

**Aluno:** João Matheus de Oliveira Schmitz

**Matrícula:** 200058525

**Turma:** T01

**Semestre:** 2024.2

**Novembro, 2024**

# Sumário

<b>1. O que é a Inteligência Artificial?</b>	<b>1</b>
1.1. Diferentes abordagens	1
1.1.1. Fidelidade à performance humana x Racionalidade	1
1.1.2. Raciocínio x Comportamento	2
1.1.3. Exemplos para cada abordagem	2
1.1.4. Modelo Padrão	3
1.2. Fundamentos da Inteligência Artificial	4
<b>2. Histórico</b>	<b>5</b>
2.1. Mitologia e ficção	5
2.2. O início da Inteligência Artificial (1943 - 1956)	5
2.3. Entusiasmo inicial, Grandes expectativas (1952 - 1969)	6
2.4. Uma dose de realidade (1966 - 1973)	7
2.5. Sistemas Especialistas (1969 - 1986)	7
2.6. O retorno das Redes Neurais (1986 - presente)	9
2.7. Raciocínio probabilístico e Aprendizado de máquina (1987 - presente)	9
2.8. Big Data (2001 - presente)	10
2.9. Deep Learning (2011 - presente)	11
<b>3. Estado da arte, Benefícios e Riscos</b>	<b>12</b>
3.1. Stanford University's One Hundred Year Study on AI (AI100)	12
3.2. O que uma IA pode fazer hoje?	12
3.3. Riscos e benefícios da IA	13
3.4. IA Index 2024	14
3.5. Reportagens relacionadas	15
<b>4. Agentes e Ambientes</b>	<b>16</b>
4.1. Agentes inteligentes e seus componentes	16
4.1.1. Função agente	16
4.1.2. Programa agente	16
4.1.3. Arquitetura do agente	16
4.2. Racionalidade	17



<b>4.3. Especificando o Ambiente de tarefas</b>	<b>18</b>
4.3.1. Descrição PEAS	18
4.3.2. Propriedades do Ambiente de tarefas	18
4.3.3. Exemplo de Ambiente de tarefas	21
<b>4.4. Tipos de Agente</b>	<b>22</b>
4.4.1. Agentes reativos simples	22
4.4.2. Agentes reativos baseados em modelo	23
4.4.3. Agentes baseados em objetivos	23
4.4.4. Agentes utilitários	24
<b>4.5. Representação de estados do ambiente</b>	<b>26</b>
4.5.1. Representação atômica	26
4.5.2. Representação fatorial	26
4.5.3. Representação estruturada	27
<b>5. Conclusão</b>	<b>28</b>
<b>6. Referências</b>	<b>29</b>

# 1. O que é a Inteligência Artificial?

Inteligência é definida como a capacidade de compreender e resolver novos problemas e conflitos e de adaptar-se a novas situações. Também pode ser vista como o ato de aprender com experiências passadas e analisar o impacto de suas ações no futuro.

Os computadores conseguem resolver diversos problemas com mais rapidez e precisão do que os humanos, entretanto existem muitas instâncias onde as limitações inerentes de um computador os impedem de resolver conflitos e/ou alcançar certos objetivos, é aí que entra a Inteligência Artificial.

Existem muitas definições para o termo Inteligência Artificial. A IBM, por exemplo, traz a seguinte frase: “Inteligência artificial, ou IA, é uma tecnologia que permite que computadores e máquinas simulem a capacidade de resolução de problemas e a inteligência humana”.

É importante, porém, entender que nem toda IA é igual, existem várias diferenças entre elas, as quais dependem de qual o tipo de problema elas foram desenvolvidas para resolver.

## 1.1. Diferentes abordagens

Uma das formas de diferenciar inteligências artificiais é através de quais abordagens que esta IA pode adotar ao resolver problemas ou interagir com o mundo. Existem dois eixos utilizados nesta análise, eles são: ‘Fidelidade à performance humana vs. Racionalidade’ e ‘Raciocínio vs. Comportamento’.

### 1.1.1. Fidelidade à performance humana x Racionalidade

Caso uma IA seja criada com uma abordagem de **fidelidade à performance humana**, seu objetivo será reproduzir ou imitar os atos que um humano iria tomar em determinada situação. Essa abordagem busca entender e simular a maneira como os humanos realmente resolvem problemas, levando em conta o uso de intuição, vieses, erros, limitações, etc.

Por outro lado, uma IA criada com uma abordagem de **racionalidade** tem como objetivo atuar da forma mais otimizada possível, através de princípios matemáticos e lógicos. Neste caso, a IA não imita os processos humanos, ela foca em encontrar as melhores soluções possíveis para aquele contexto.

### 1.1.2. Raciocínio x Comportamento

A abordagem de **raciocínio** é diferente da abordagem racional já discutida, pois foca mais no "pensar" do que no "agir". Uma IA criada com essa abordagem tem como preocupação a capacidade de análise, compreensão e tomada de decisão antes de executar uma ação, sempre de forma lógica e estruturada. Portanto, IAs com foco em raciocínio são capazes de justificar suas decisões e demonstrar o raciocínio utilizado para chegar nestas.

Em contrapartida, a abordagem **comportamental** foca mais no "agir" do que no "pensar", ao contrário da abordagem de raciocínio. Isso ocorre pois IAs criadas assim se preocupam mais com seus comportamentos e como eles afetam o mundo real. Essa abordagem é popular com sistemas mais reativos.

### 1.1.3. Exemplos para cada abordagem

Ao combinar esses eixos, é possível distinguir quatro grupos de inteligências artificiais. Seguem aqui exemplos para cada grupo:

- IA com **fidelidade à performance humana** e foco em **raciocínio**: sistemas que simulam processos cognitivos e a tomada de decisão humana, como os utilizados na psicologia cognitiva para entender como o cérebro humano raciocina e aprende. **São sistemas que pensam como humanos.**
- IA com **fidelidade à performance humana** e foco em **comportamento**: assistentes virtuais, os quais utilizam padrões de conversa humana para interagir com usuários sem a necessidade de raciocínios lógicos complexos. **São sistemas que agem como humanos.**
- IA **racional** com foco em **raciocínio**: sistemas de diagnóstico médico, onde é utilizada somente a lógica para gerar diagnósticos a partir dos sintomas e histórico médico do paciente, sem deixar vieses e limitações humanas atrapalharem a decisão. **São sistemas que pensam racionalmente.**

- IA **racional** com foco em **comportamento**: sistemas de carros autônomos, os quais otimizam suas ações para garantir a segurança e o conforto dos passageiros, levando em conta como suas ações afetam o mundo ao seu redor. **São sistemas que agem racionalmente.**

#### 1.1.4. Modelo Padrão

Entre as diferentes abordagens utilizadas na concepção e desenvolvimento de inteligências artificiais, há uma em específico que prevalece na maior parte dos estudos e criações da área ao longo de sua história, a abordagem da **racionalidade**.

A área de inteligência artificial se concentrou, de maneira geral, no estudo e construção de agentes que fazem a “**coisa certa**”, que neste caso se refere à realização de um determinado objetivo. Este objetivo pode ser desde manter dois quadrados livres de poeira (para um robô aspirador de pó) até o transporte de passageiros para seus destinos (para um carro autônomo), variando de acordo com a IA.

Devido aos fatores comentados acima, a abordagem de um agente racional se vê como a abordagem mais comumente utilizada no estudo e desenvolvimento de IAs. Este paradigma geral é tão difundido que nós podemos chamá-lo de **modelo padrão**.

Embora o modelo padrão seja muito útil no desenvolvimento de inteligências artificiais, estamos encontrando problemas devido a suas limitações. O seu principal problema é assumir que forneceremos um objetivo totalmente especificado para a IA, o que se torna uma tarefa extremamente difícil em situações complexas do mundo real.

A tarefa de chegar a um acordo entre nossas verdadeiras preferências e o objetivo que especificamos para a IA é chamado de **Problema de Alinhamento de Valor**. A resolução deste problema é de suma importância no desenvolvimento de IAs com o modelo padrão, pois objetivos mal definidos geram consequências negativas quanto à performance esperada da IA. (Russell & Norvig, 2009)

## 1.2. Fundamentos da Inteligência Artificial

A área de estudo de inteligência artificial não é uma área isolada, vários conceitos e técnicas utilizadas em seu desenvolvimento tem origem nas mais diversas disciplinas. Seguem aqui alguns exemplos de disciplinas, juntamente com algumas de suas perguntas, que contribuíram com a área de IA (Russell & Norvig, 2009):

- Filosofia;
  - As regras formais podem ser usadas para obter conclusões válidas?
  - Como a mente se desenvolve a partir de um cérebro físico?
  - De onde vem o conhecimento?
  - Como o conhecimento leva à ação?
- Matemática;
  - Quais são as regras formais para obter conclusões válidas?
  - O que pode ser computado?
  - Como raciocinamos com informações incertas?
- Economia;
  - Como devemos tomar decisões para maximizar a recompensa?
  - Como devemos fazer isso quando outros podem não concordar?
  - Como devemos fazer isso quando a recompensa pode estar num futuro distante?
- Neurociência;
  - Como o cérebro processa informações?
- Psicologia;
  - Como os seres humanos e os animais pensam e agem?
- Engenharia da computação;
  - Como podemos construir um computador eficiente?
- Teoria de controle e cibernética;
  - Como os artefatos podem operar sob seu próprio controle?
- Linguística;
  - Como a linguagem se relaciona com o pensamento?



## 2. Histórico

### 2.1. Mitologia e ficção

O conceito base de inteligência artificial tem origem muito antes da existência dos computadores, com histórias vindas desde a mitologia grega, como a de **Talos**, um autômato gigante de bronze que era responsável pela proteção da ilha de Creta e dos povos que ali viviam.

Com o passar do tempo, cada vez mais histórias foram originando e/ou incorporando conceitos relacionados à área de IA. Alguns exemplos são **Frankenstein**, onde um humano cria uma criatura inteligente de modo artificial, **Pinóquio**, onde um menino de madeira ganha vida e busca se tornar um menino de 'verdade', ou, em outras palavras, se tornar humano, e o **Exterminador do Futuro**, que traz a ideia de que inteligências artificiais podem se rebelar contra os humanos que as criaram.

Histórias como estas trazem diversas ideias, conceitos e perguntas que ajudaram a estabelecer e avançar os estudos na área de IA, sendo parte **fundamental** de sua história.

### 2.2. O início da Inteligência Artificial (1943 - 1956)

A área de inteligência artificial, como área de estudo, teve seu início real em 1943 com Warren McCulloch e Walter Pitts. Estes pesquisadores propuseram um **modelo de neurônios** onde cada neurônio tem um estado: ligado ou desligado, onde os neurônios 'ligam' quando ocorre um número suficiente de estímulos por neurônios vizinhos. Para provar a utilidade de seu modelo, eles demonstraram que todos os conectivos lógicos (e, ou, não, etc.) podiam ser implementados por estruturas de redes simples. (Russell & Norvig, 2009)

Com base no trabalho de McCulloch e Pitts, Donald Hebb trouxe um grande avanço na área de IA em 1949 através do que hoje é chamado de **Aprendizado de Hebb** ou **Teoria Hebbiana**. Sua ideia, utilizada até hoje, tem como base a aprendizagem associativa e traz o conceito da intensidade entre as conexões de neurônios para o modelo de 1943, fazendo com que neurônios que ficam simultaneamente ativos em determinados momentos de tempo tenham

uma conexão mais forte do que aqueles que costumam estar em estados diferentes.

Foi nessa época, mais especificamente em 1950, que Alan Turing lançou seu artigo “Computing Machinery and Intelligence”, onde trouxe a ideia do **teste de Turing**, o qual categorizou as 6 maiores disciplinas de estudo dentro da área de inteligência artificial. São elas:

- Processamento de linguagem natural;
- Representação de conhecimento;
- Raciocínio automatizado;
- Aprendizado de máquina;
- Visão computacional;
- e Robótica;

## 2.3. Entusiasmo inicial, Grandes expectativas (1952 - 1969)

Nos anos iniciais da área, ocorreram muitos sucessos e um grande progresso no desenvolvimento de inteligências artificiais. Nessa época era comum pessoas dizerem que “uma máquina nunca poderá realizar X”, e pesquisadores da área provavam que tais afirmações eram falsas.

Essa época foi marcada por programas como o GPS (General Problem Solver), feito por Newell e Simon, o primeiro programa a incorporar a abordagem de “pensar de forma humana”, a criação da linguagem de alto nível **Lisp**, que se tornou a linguagem de programação dominante na IA pelos próximos 30 anos e o lançamento de artigos como “**Programs with common sense**”, o qual traz descreve o Advice Taker, um programa hipotético que pode ser visto como o primeiro sistema de IA completo e que tinha incorporado em si o conhecimento geral do mundo para a resolução de problemas no mundo real, ambos realizados por John McCarthy. (Russell & Norvig, 2009)

## 2.4. Uma dose de realidade (1966 - 1973)

Com o sucesso inicial do uso de inteligências artificiais na resolução de problemas simples, as expectativas colocadas na área eram enormes. Consequentemente, quando os pesquisadores encontraram barreiras em seus estudos ao tentarem utilizar IAs na resolução de problemas mais complexos e/ou extensos, o 'banho de água fria' foi ainda maior.

Algumas das dificuldades encontradas pelos pesquisadores foram:

- A maioria dos primeiros programas não tinha conhecimento de seu assunto; eles obtinham sucesso por meio de manipulações sintáticas simples. Consequentemente, **ambiguidades** geradas pelo **contexto** de um problema geravam erros e inconsistências nos resultados.
- A impossibilidade de tratar muitos dos problemas que a IA tentava resolver, devido à **maior complexidade** ao lidar com ambientes variáveis fora de 'micro-mundos'.
- As limitações fundamentais nas estruturas básicas que estavam sendo utilizadas para gerar o comportamento inteligente, como os Perceptrons (forma simples de rede neural) que podiam aprender tudo o que eram capazes de representar, mas **o que podiam representar era muito pouco**.
- O **poder computacional** e os algoritmos disponíveis na época eram **insuficientes** para resolver problemas complexos do mundo real, expondo as limitações dos primeiros sistemas de IA.

Pode-se dizer que as altas expectativas na área, em conjunto com a falta de progresso rápido na resolução de problemas complexos, ou até problemas relativamente simples, como é o caso do **problema da porta XOR**, trouxe o primeiro "**inverno IA**". (FAVARON, 2024)

## 2.5. Sistemas Especialistas (1969 - 1986)

Devido às barreiras encontradas na resolução de problemas mais complexos utilizando o método da época, o qual era focado em encontrar soluções completas para vários problemas diferentes através de passos lógicos elementares, a área de inteligência artificial passou por um período onde seu

foco foi em grande parte o desenvolvimento dos chamados **sistemas especialistas**.

Esses sistemas têm como principal característica o foco na resolução de problemas relacionados a **domínios específicos**, os tornando bem mais eficazes na resolução de desafios que envolvam o domínio em que se especializam. Consequentemente, o uso de tais sistemas em desafios que fogem de seus respectivos domínios de conhecimento não é aconselhado.

Ao **diminuir o escopo de domínio** de uma IA, foi permitido aos pesquisadores o **aumento na complexidade dos problemas** incluídos no escopo que poderiam ser solucionados usando tal IA.

Um exemplo de sistema especialista é o programa DENDRAL, criado por Ed Feigenbaum, Bruce Buchanan e Joshua Lederberg em 1969. Este programa é usado para inferir a estrutura molecular a partir de informações fornecidas por um espectrômetro de massa. Ele foi considerado o primeiro sistema bem-sucedido de conhecimento intensivo, pois tomava suas decisões baseadas em um grande número de **regras relacionadas especificamente com o domínio** do problema.

Por outro lado, tivemos o R1 em 1982, considerado o primeiro sistema especialista bem-sucedido comercialmente, utilizado para ajudar na configuração de pedidos de novos sistemas de computadores e fazendo a empresa **economizar dezenas de milhões** de dólares por ano com sua utilização. (Russell & Norvig, 2009)

A popularização de sistemas como o R1 levou a década de 80 a ser marcada pela expansão em **ordens de magnitude** no valor da área de inteligência artificial. Infelizmente, os mesmos erros do passado foram cometidos através de **expectativas extremamente altas** na área. Consequentemente, quando foram encontradas barreiras no progresso da área devido à incertezas nos contextos de problemas e a incapacidade destes sistemas em aprender com a experiência, muitas empresas caíram no esquecimento por não conseguirem alcançar expectativas tão extravagantes, causando um **declínio no interesse e financiamento** da área e gerando um segundo “inverno IA”. (FAVARON, 2024)

## 2.6. O retorno das Redes Neurais (1986 - presente)

O problema da porta XOR causou um período de estagnação na área de redes neurais, porém, em meados de 1980, pelo menos 4 grupos de pesquisadores reinventaram um método encontrado por Bryson e Ho em 1969, o qual resolve este problema e abre muitas portas para o progresso da área, o algoritmo de aprendizado por **retropropagação**: "... um método elegante para calcular como as alterações em qualquer um dos pesos ou vieses de uma rede neural afetarão a precisão das previsões do modelo.". (IBM, 2024)

O artigo "**Learning representations by back-propagating errors**", lançado em 1986 por David E. Rumelhart, Geoffrey Hinton e Ronald J. Williams, foi o responsável por: "... fornecer a derivação do algoritmo de retropropagação conforme usado e compreendido em um contexto moderno de aprendizado de máquina.". (IBM, 2024)

Nos dias de hoje, a pesquisa de redes neurais se bifurcou em dois campos, onde o primeiro é preocupado com "... a criação de algoritmos e arquiteturas de rede eficazes e a compreensão de suas propriedades matemáticas..." e o segundo com "... a modelagem cuidadosa das propriedades empíricas de neurônios reais e conjuntos de neurônios.". (Russell & Norvig, 2009)

## 2.7. Raciocínio probabilístico e Aprendizado de máquina (1987 - presente)

A área de inteligência artificial se desenvolveu por muito tempo como uma área de estudo isolada de outras disciplinas, como a ciência da computação. Porém, esse isolacionismo foi deixado para trás em prol de avanços no desenvolvimento da área.

No campo de reconhecimento da fala, por exemplo, abordagens baseadas em modelos ocultos de Markov passaram a dominar a área. Estes modelos têm como característica a determinação de parâmetros ocultos com base em parâmetros visíveis, com base em **probabilidade e reconhecimento de padrões**. Neste caso, eles trazem dois benefícios principais: a base em uma teoria matemática rigorosa e seu processo de geração, realizado com um grande conjunto de dados reais de fala. (Russell & Norvig, 2009)

A obra “**Probabilistic Reasoning in Intelligent Systems**”, escrita por Judea Pearl em 1988, gerou uma nova aceitação do uso da probabilidade e estatística na área de inteligência artificial. Nesta obra, são utilizadas as chamadas **redes bayesianas** para representar conhecimento incerto, superando em muito o método de raciocínio probabilístico utilizado em décadas anteriores, sendo o principal método de IA utilizado para trabalhar com raciocínio incerto. Exemplos de aplicações deste método são a biologia molecular e a medicina diagnóstica, onde a aquisição de dados é cara e requer estudos experimentais difíceis. (DATABRICKS, 2024)

A conexão com outras disciplinas também se faz presente em outros campos da IA, como robótica, visão computacional e representação de conhecimento. Pode-se dizer que a reintegração com outras áreas **beneficiou e continua a beneficiar** os estudos de inteligência artificial até os dias atuais.

## 2.8. Big Data (2001 - presente)

O termo Big Data tem como definição: “... são dados que contêm maior **variedade**, chegando em **volumes** crescentes e com mais **velocidade**... Esses conjuntos de dados são tão volumosos que o software tradicional de processamento de dados simplesmente não consegue gerenciá-los.”. (Oracle, 2024)

Sua conexão com o campo de inteligência artificial ocorre, principalmente, no processo de **aprendizado de máquina**, onde uma IA pode ser alimentada com uma quantidade enorme de dados. Segundo Yarowsky (1995) e Banko e Brill (2001), a utilização de um acervo muito maior de dados durante o processo de aprendizagem gera um benefício bem maior no desempenho, independente do algoritmo utilizado no processo. (Russell & Norvig, 2009)

Entretanto, a utilização de algoritmos mais eficazes e otimizados não se tornou obsoleta, seus benefícios ainda são válidos. Consequentemente, para uma otimização ainda maior do processo, **algoritmos especializados** em tratar a enorme quantidade de dados que o Big Data fornece foram desenvolvidos, alimentando de forma cada vez mais **eficaz** o aprendizado de novas inteligências artificiais.

## 2.9. Deep Learning (2011 - presente)

O processo conhecido como deep learning é um subconjunto de aprendizado de máquina, porém com uma diferença crítica: a quantidade de camadas de suas redes neurais. É comum que os modelos “não profundos” usem redes neurais simples de 1 ou 2 camadas, já os modelos de deep learning utilizam, geralmente, centenas ou milhares de camadas no processo de treinamento. (IBM, 2024)

Segundo a IBM, “O deep learning é uma vertente da ciência de dados que conduz muitas aplicações e serviços que melhoram a automação, realizando tarefas analíticas e físicas sem intervenção humana”. Seus usos incluem carros autônomos, IAs generativas, assistentes digitais, entre outros.

### 3. Estado da arte, Benefícios e Riscos

#### 3.1. Stanford University's One Hundred Year Study on AI (AI100)

Com o progresso constante da inteligência artificial, cuidados devem ser tomados para que seus esforços ajudem a sociedade, garantindo que ela está solucionando problemas, ao invés de criando-os. De acordo com o One Hundred Year Study on Artificial Intelligence (AI100), um projeto da universidade de Stanford, "A sociedade está agora em um momento crucial onde deve determinar **como implantar tecnologias** baseadas em IA de maneiras que **promovam**, ao invés de dificultarem valores democráticos como liberdade, igualdade e transparência".

O AI100 traz relatórios regulares sobre as novidades, tendências e desenvolvimentos dos últimos anos da área de inteligência artificial, além de realizar projeções sobre os próximos anos na área. Seu foco principal é nos oito domínios a seguir:

- Transporte (carros autônomos, ...)
- Casa / Robôs de serviço (robôs de limpeza, assistentes digitais, ...)
- Saúde e Medicina (monitoramento de saúde pessoal, cirurgia robótica, ...)
- Educação (sistema de ensino que entende linguagem natural, ...)
- Segurança pública (câmeras, drones, detecção de ameaças, ...)
- Comunidades carentes (sistema para melhor distribuição de alimentos, ...)
- Emprego e local de trabalho (ajudar na adaptação durante mudanças, ...)
- e Entretenimento (sistemas de criação de conteúdo como IA generativa, ...)

#### 3.2. O que uma IA pode fazer hoje?

Nos dias de hoje, a inteligência artificial está profundamente enraizada em muitos aspectos da vida humana, com potencial para se tornar ainda mais fundamental para nosso estilo de vida. Exemplos de uso desta tecnologia podem variar bastante, desde o uso em **carros autônomos** para levar um passageiro até o seu destino de maneira segura, eficaz e confortável, até em **planejamento logístico** para facilitar a criação de um plano de ação no transporte e distribuição de cargas para vários destinos.



Embora exemplos como os acima possam passar a impressão de que a IA só pode ser acessada por pessoas com maior poder econômico ou por empresas e indústrias, tal afirmação poderia estar correta no passado, mas no presente ela perde sua veracidade. O uso de assistentes digitais como a Alexa, da Amazon, ou a Cortana, da Microsoft, ou até mesmo o Google Assistente e chatbots presentes em sites de venda, todos os quais utilizam **sistemas de reconhecimento de voz**, construídos em cima de algoritmos como os de redes neurais, demonstra que as **IAs já permeiam diversas camadas da sociedade**, não só as mais privilegiadas. (IBM, 2024)

Outros exemplos são a utilização de machine learning por empresas como a Amazon, Netflix e Google para a recomendação de produtos/serviços para uma pessoa com base no que ela já demonstrou interesse previamente; a criação de IAs capazes de jogar jogos em nível profissional, ou até melhor do que qualquer humano; a tradução automática de textos de uma língua para outra; diagnóstico de doenças com base em sintomas e histórico de um paciente; além de sistemas como o ChatGPT, capaz de responder perguntas sobre os mais diversos tópicos, e sistemas como o Midjourney, capaz de gerar imagens com base em descrições providas pelo usuário.

### 3.3. Riscos e benefícios da IA

Ao analisar os exemplos da seção anterior, é fácil de visualizar a quantidade enorme de benefícios que a inteligência artificial traz para a humanidade. Para citar alguns deles:

- aumento de produtividade;
- monitoramento e detecção de crimes;
- transportes autônomos;
- casas inteligentes;
- recomendações baseadas em gosto pessoal;
- utilização da IA em trabalhos repetitivos e/ou perigosos;
- identificação e prevenção da propagação de vírus, fake news, etc;
- aceleração de descobertas científicas;

- e muitos outros.

Entretanto, como nada é perfeito, a inteligência artificial também traz riscos associados ao uso indevido de suas capacidades. Por exemplo:

- o aumento no desemprego devido a substituição de trabalhadores por IAs;
- o potencial destrutivo de armas de guerra autônomas produzidas em massa por países mais desenvolvidas quando utilizadas contra países sem os recursos necessários para sua criação;
- a perda de privacidade, censura política e a perda, mesmo que parcial, da livre expressão;
- vieses contra certos grupos raciais, políticos, etc, na decisão e/ou resultados obtidos por IAs devido a processos de aprendizado enviesados;
- a invasão por parte de hackers em sistemas como os de carros autônomos, o que pode levar a acidentes e crimes graves;
- puxando de obras de ficção como Exterminador do Futuro, o hipotético potencial de uma ‘revolução das máquinas’;
- entre outros.

Como conclusão, a inteligência artificial é uma tecnologia inovadora que, assim como várias tecnologias passadas, traz benefícios e malefícios consigo. Por isso, é necessário uma forte regularização para a diminuição do impacto ético/econômico/político/social/militar que pessoas mal intencionadas podem infligir com o uso indevido desta tecnologia.

### 3.4. IA Index 2024

O IA Index é um projeto da Universidade de Stanford que publica relatórios anuais sobre o progresso nos estudos e pesquisas relacionados à inteligência artificial. No IA Index 2024, muitos assuntos foram abordados, mas o foco dos próximos parágrafos se encontra no capítulo 4: **“The data is in: AI makes workers more productive and leads to higher quality work”**.

Segundo pesquisas de 2023, a produtividade de trabalhadores tende a ser maior quando estes utilizam inteligência artificial, quando comparado a não

utilização. Esse assunto é interessante de ser discutido pois está interligado com um tópico extremamente delicado: **o impacto da IA em trabalhos humanos**.

Esse assunto está diretamente ligado a riscos como plágio e, principalmente, aumentos na taxa de desemprego. Muito é discutido sobre a possibilidade da inteligência artificial deixar o mundo desempregado, porém, estudos como os referenciados pelo IA Index 2024 mostram que a IA não tem somente potenciais negativos neste quesito; seus benefícios são claros e já se fazem presentes no mercado de trabalho atual.

Embora existam riscos como a substituição de humanos por IAs em diversos trabalhos, o aumento da produtividade de trabalhadores humanos ao utilizar essa tecnologia se torna um argumento incontestável de que a inteligência artificial é sim uma tecnologia com o potencial de trazer grandes benefícios para a humanidade.

### 3.5. Reportagens relacionadas

Para trazer uma conexão maior com o mundo real, segue aqui uma lista de reportagens/notícias relacionadas ao assunto:

- Hackers invadem sistema de um Jeep de forma remota:  
<https://www.youtube.com/watch?v=MK0SrxBC1xs>
- Criação e propagação de conteúdos falsos para a manipulação de eleições:  
<https://noticias.uol.com.br/confere/ultimas-noticias/2024/03/03/deepfake-uso-inteligencia-artificial-eleicoes-argentina-estados-unidos.htm>
- Imitação de voz para roubo de centenas de milhares de dólares:  
<https://www.infomoney.com.br/negocios/inteligencia-artificial-imita-voz-de-ceo-em-roubo-de-us-243-mil/>
- Algoritmo enviesado utilizado para analisar a possibilidade de reincidência de criminosos nos EUA:  
<https://apublica.org/2016/06/software-que-avalia-reus-americanos-cria-injusticas-na-vida-real/>
- Vantagens e riscos na cibersegurança:  
<https://olhardigital.com.br/2024/11/16/seguranca/especialistas-revelam-vantagens-e-riscos-do-uso-de-inteligencia-artificial-na-ciberseguranca/>

## **4. Agentes e Ambientes**

### **4.1. Agentes inteligentes e seus componentes**

Um agente, quando falamos de inteligência artificial, pode ser definido como uma entidade que percebe o ambiente, toma decisões e realiza ações para alcançar objetivos específicos. As ações tomadas por um agente dependem de sua programação interna e dos dados que podem ser obtidos do ambiente em que se encontra.

#### **4.1.1. Função agente**

A função agente é a responsável por ligar as informações sobre o ambiente recebidas pelo agente em determinado momento de tempo com as ações que devem ser tomadas por este de modo a chegar mais próximo de seu objetivo. Pode ser vista como a lógica que irá guiar o agente no processo de alcançar seus objetivos.

#### **4.1.2. Programa agente**

O programa agente pode ser definido como a implementação da função agente. Durante o processo de implementação ocorre o desenvolvimento, o treinamento e a implantação do agente na arquitetura selecionada, seja ela física ou digital. (AWS, 2024)

#### **4.1.3. Arquitetura do agente**

O agente, para agir sobre seu ambiente, necessita de um meio para tal. Um agente robótico, por exemplo, pode utilizar uma garra mecânica para levar itens de um lugar para outro. A arquitetura em que o agente é baseado pode ser tanto física, como no exemplo anterior, quanto digital, como no ChatGPT, um programa de software.

Não existe uma arquitetura geral, afinal um robô aspirador de pó precisa ter a capacidade de se mover e sugar a poeira, enquanto um assistente digital necessita de um meio para captar vozes, por exemplo. Consequentemente, cada

agente deve ter uma arquitetura personalizada concebida com a capacidade de atuar em seu ambiente e realizar seu objetivo.

## 4.2. Racionalidade

Quando falamos de agentes em inteligência artificial, geralmente estamos falando de agentes racionais, aqueles que ‘fazem tudo certo’. Porém, o que significa ‘fazer tudo certo’? Neste caso, utilizamos uma **medida de desempenho** para avaliar como as ações do agente afetam os estados de seu ambiente. Quanto mais próximo o estado do ambiente chegar ao objetivo desejado, mais podemos dizer que o agente está ‘fazendo a coisa certa’. Segundo a definição de Russell & Norvig (2009), “Para cada **sequência de percepções** possível, um agente racional deve **selecionar uma ação** que se espera venha a **maximizar sua medida de desempenho**, dada a evidência fornecida pela sequência de percepções e por qualquer **conhecimento interno do agente**”.

Algo importante de ser comentado é que não existe uma medida de desempenho geral para todos os agentes, afinal, cada agente individual atua em domínios e ambientes distintos, além de ter diferentes objetivos.

Ao criar uma medida de desempenho, alguns cuidados precisam ser tomados, pois, por exemplo, um aspirador de pó inteligente programado com uma medida de desempenho que o recompensa pela quantidade de vezes que ele limpou o chão em um período de tempo, neste caso ele pode sujar o chão, limpar, sujar, limpar e continuar nesse ciclo, maximizando a medida de desempenho através de um comportamento que vai contra o resultado esperado. Neste caso, uma medida de desempenho mais correta recompensaria o agente por minimizar o tempo que o chão fica sujo, gastando a menor quantidade de energia possível no processo, tornando o agente mais **racional**.

Pode ser compreendido então, como regra geral, que “...é melhor projetar medidas de desempenho de acordo com o resultado realmente desejado no ambiente, em vez de criá-las de acordo com o comportamento esperado do agente.” (Russell & Norvig, 2009)

## 4.3. Especificando o Ambiente de tarefas

O ambiente de tarefas é parte fundamental de todo agente, pois **descreve o problema** para qual o agente busca uma solução, **o mundo** no qual o agente percebe e interage. Portanto, o ambiente de tarefas é utilizado como base para a criação da função agente e programa agente, e, conseqüentemente, sua **especificação é o primeiro passo na construção de um agente**.

### 4.3.1. Descrição PEAS

O ambiente de tarefas é **composto por quatro componentes**:

- **Medida de performance**: define o que o agente deve fazer, seus objetivos;
- **Ambiente**: o mundo que o agente 'vive', incluindo todos os elementos que podem interagir com ele durante a realização de seus objetivos;
- **Atuadores**: componentes da arquitetura que o agente utiliza para realizar ações;
- e **Sensores**: componentes da arquitetura que o agente utiliza para perceber o ambiente ao seu redor;

Ao especificar um ambiente de tarefas, é preciso definir os componentes acima. À descrição desses componentes, damos o nome de **descrição PEAS** (Performance, Environment, Actuators, Sensors — performance, ambiente, atuadores, sensores). (Russell & Norvig, 2009)

### 4.3.2. Propriedades do Ambiente de tarefas

Os ambientes de tarefas podem ser extremamente variados, mas ainda existem pontos em comum entre eles. Com essa ideia em mente, é possível categorizar os ambientes de tarefas de acordo com suas propriedades, sete delas, de acordo com Russell & Norvig (2009). Essas categorias são:

- **Completamente observável vs. Parcialmente observável vs. Não observável**;
  - Completamente observável:  
Os sensores de um agente captam todos os elementos de um ambiente (relevantes na hora de tomar uma ação) em todos os instantes de tempo.

- Parcialmente observável:  
Os sensores de um agente não conseguem captar todos os elementos relevantes de um ambiente em todos os instantes de tempo, seja devido a fontes de ruído, imprecisão ou impossibilidades técnicas, gerando assim incertezas nos estados do ambiente.
- Não observável:  
O agente não possui sensores para detectar o ambiente ao seu redor.
- **Agente único vs. Multiagente;**
  - Agente único:  
Quando o ambiente só tem um agente realizando ações para alcançar seus objetivos.
  - Multiagente:  
Quando há dois ou mais agentes afetando um mesmo ambiente para realizar seus objetivos. Neste caso, podemos dividir o ambiente em multiagente **competitivo**, onde os agentes têm objetivos que conflitam uns com os outros, e multiagente **cooperativo**, onde os agentes podem cooperar em prol de ambos atingirem seus objetivos.
- **Determinístico vs. Não determinístico;**
  - Determinístico:  
O estado futuro do ambiente depende completamente do estado atual e das ações tomadas pelo agente.
  - Não determinístico:  
O ambiente apresenta elementos não-visíveis e/ou impossíveis de prever, o que gera incertezas. Podemos dividir esses ambientes em **estocásticos**, caso as incertezas tenham probabilidades associadas, e **não-determinísticos**, onde não são atreladas probabilidades às incertezas.
- **Episódico vs. Sequencial;**
  - Episódico:  
O agente percebe o estado atual do ambiente e decide em uma ação somente com este estado em mente. As ações tomadas em um episódio não afetam outros.

- Sequencial:  
Neste caso, as ações tomadas em um instante de tempo podem afetar o estado futuro do ambiente e, conseqüentemente, as próximas ações do agente. Isso torna necessário que o agente seja capaz de prever como suas ações afetam o futuro.
- **Estático vs. Dinâmico**;
  - Estático:  
Não são exibidas mudanças no estado do ambiente enquanto o agente delibera sua próxima ação.
  - Dinâmico:  
O estado do ambiente pode ser alterado a qualquer momento, ou seja, o ambiente não espera o agente e está em constante mudança. Existem ainda ambientes que não exibem mudanças com a passagem de tempo, mas o nível de desempenho do agente se altera; neste caso chamamos o ambiente de **semi-dinâmico**.
- **Discreto vs. Contínuo**;
  - Discreto:  
Ocorre quando o ambiente tem estados claramente distintos, com momentos discretos de percepção e ação.
  - Contínuo:  
Neste caso, o estado do ambiente varia continuamente, necessitando de percepções e ações constantes por parte do agente.
- **Conhecido vs. Desconhecido**;
  - Conhecido:  
O agente, ou seu projetista, conhece as leis que governam o ambiente onde ele se encontra, podendo prever com certeza ou com certa probabilidade as saídas causadas por uma ação.
  - Desconhecido:  
O agente, ou seu projetista, não tem conhecimento prévio de como o ambiente funciona. Sendo assim, o agente precisará aprender as regras do ambiente para conseguir tomar boas decisões.



### 4.3.3. Exemplo de Ambiente de tarefas

Para demonstrar a descrição PEAS e as propriedades de um ambiente de tarefas, podemos tomar como base, por exemplo, uma IA com o objetivo de aprender a jogar o jogo Super Mario e finalizá-lo.

Sua **descrição PEAS** seria:

<b>Tipo de agente</b>	Jogador de Super Mario
<b>Medida de desempenho</b>	Aprender como o jogo funciona, chegar ao final do jogo, minimizar a quantidade de vidas perdidas, minimizar o tempo necessário para chegar ao fim do jogo, maximizar a pontuação
<b>Ambiente</b>	Plataformas, buracos, canos, obstáculos, blocos especiais, itens, moedas, inimigos, tempo
<b>Atuadores</b>	Movimento para esquerda e direita, pular, se abaixar (também utilizado para entrar em canos), lançar bola de fogo quando disponível
<b>Sensores</b>	Sensor visual para identificar os elementos na tela, como a posição dos elementos do ambiente, a pontuação, o tempo, as vidas restantes e a posição do Mario

Já as **propriedades do ambiente de tarefas** seriam:

<b>Parcialmente Observável</b>	Pois o agente só consegue visualizar parte da fase do jogo em determinado momento de tempo
<b>Agente Único</b>	Pois só existe um agente influenciando o ambiente
<b>Estocástico</b>	Pois existem blocos ocultos, os canos podem ou não ter plantas carnívoras e existe uma chance do personagem receber moedas ou power ups ao bater em blocos
<b>Sequencial</b>	Pois as decisões do agente afetam os estados do ambiente e ações futuras

<b>Dinâmico</b>	Pois o ambiente muda constantemente através de plataformas móveis, inimigos e o tempo da fase
<b>Discreto</b>	Pois os estados do ambiente são caracterizados por cada frame da tela e as ações do agente também são discretas, embora o tempo seja contínuo
<b>Desconhecido</b>	Pois o agente não foi programado com o conhecimento de como o jogo funciona

## 4.4. Tipos de Agente

Dependendo dos problemas para os quais um agente foi criado para solucionar e sua capacidade de raciocinar, seu programa agente pode estar categorizado entre os 4 tipos a seguir: agentes reativos simples, agentes reativos baseados em modelo, agentes baseados em objetivos e agentes utilitários.

### 4.4.1. Agentes reativos simples

O tipo mais simples, como seu nome já diz, é o **agente reativo simples**. Sua principal característica é determinar que ação deve ser tomada em momento de tempo levando em conta **somente o estado presente do ambiente**. Eles ignoram o histórico de estados anteriores e também previsões futuras, tendo um maior desempenho em ambientes **totalmente observáveis**. Um exemplo é o robô aspirador de pó, que tem como regras:

- **Se** o chão estiver sujo, **então** devo aspirar;
- **Se** o chão estiver limpo e eu estiver na direita, **então** devo ir pra esquerda;
- **Se** o chão estiver limpo e eu estiver na esquerda, **então** devo ir pra direita;

Esse padrão estará presente em todos os agentes desse tipo e é denominado como **regra condição-ação**. Porém, o comportamento desses agentes também é utilizado em problemas mais complexos como em um carro autônomo, por exemplo: **se** o carro da frente frear, **então** devo frear. Isso ocorre porque algumas mudanças de estado no ambiente não necessitam de considerações adicionais por parte do agente para chegar a solução correta. (Russell & Norvig, 2009)

#### 4.4.2. Agentes reativos baseados em modelo

Como agentes reativos simples necessitam estar em ambientes totalmente observáveis para ter certeza que suas ações são corretas, eles não são efetivos em situações onde o ambiente é **parcialmente observável**, quem lida com isso são os **agentes reativos baseados em modelo**.

A primeira característica que os diferenciam de agentes reativos simples é a existência de um estado interno, onde eles **armazenam seus históricos de percepções**. Devido a isso, são capazes de monitorar e lembrar de elementos do ambiente que não estão em sua percepção imediata, permitindo seu uso em problemas mais complexos.

Para garantir que o agente consiga interpretar o ambiente ao seu redor corretamente, levando em consideração ambos os elementos que ele consegue e não consegue ver, é importante que ele tenha dois **modelos do mundo**, o diferenciando ainda mais dos agentes reativos simples. Estes modelos são:

- **Modelo transitório**: descreve como as ações do agente afetam o ambiente.
- **Modelo sensorial**: descreve como o ambiente evolui independente das ações do agente.

Ao combinar informações dos modelos transitório e sensorial com o estado interno e as regras de condição-ação, um agente reativo baseado em modelos consegue determinar as ações que precisam ser tomadas naquele instante de tempo. Entretanto, esses agentes ainda não têm conhecimento completo sobre seu ambiente, gerando **incertezas no estado atual** do ambiente, o que, por sua vez, pode acarretar no agente tomando decisões erradas em certos casos, provando que eles, assim como os agentes reativos simples, não podem ser usados em todas as situações. (Russell & Norvig, 2009)

#### 4.4.3. Agentes baseados em objetivos

Em contrapartida aos dois tipos de agentes reativos, os **agentes baseados em objetivos** não utilizam regras condição-ação, em seu lugar são utilizados os objetivos que o agente quer alcançar. Essa troca traz um benefício muito

importante, o qual permite a esses agentes atuarem em situações mais complexas que os anteriores: uma **maior flexibilidade**.

Esse ponto pode ser demonstrado em carros autônomos, pois, caso fossem programadas regras condição-ação para esses sistemas, a quantidade de variáveis ambientais que precisam ser levadas em consideração seria enorme. Neste caso, como o clima afeta o modo que o carro irá transitar pelas vias, precisam ser criadas regras específicas para que, como exemplo, o carro possa frear devido a presença de um pedestre na rua, então seria levado em consideração não somente se ele está em uma reta, curva aberta ou curva fechada, mas também se o clima está ensolarado, com chuva fraca ou forte, com vento intenso ou não, com neblina, ...

Em ambientes complexos como no exemplo anterior, a utilização de regras condição-ação, que precisam ser programadas de antemão, se torna ineficiente e, dependendo do ambiente, inviável. Por outro lado, se o agente baseado em objetivos souber como o freio se comporta em diferentes condições climáticas, ele pode **alterar seu comportamento de forma automática**, se tornando uma opção bem mais viável para situações como esta. Sem contar que objetivos podem ser reescritos bem mais facilmente do que regras condição-ação, trazendo ainda mais flexibilidade no comportamento destes agentes.

Outro ponto em que os agentes baseados em objetivo se diferenciam de agentes reativos é que, ao utilizar os modelos, estado interno e percepção atual para definir qual ação tomar, eles **consideram o impacto de suas ações no futuro**, fazendo perguntas como 'O que acontecerá se eu fizer isso?' ou 'Essa ação irá me aproximar dos meus objetivos?'.

Levando em consideração a troca das regras condição-ação para objetivos, e a introdução de considerações sobre como suas ações afetam o futuro, os agentes baseados em objetivos se mostram **capazes de resolver uma maior gama de problemas** do que agentes reativos, sejam eles simples ou baseados em modelo. (Russell & Norvig, 2009)

#### 4.4.4. Agentes utilitários

Embora os agentes baseados em objetivos sejam capazes de resolver a maioria dos problemas, isso não quer dizer que o modo que seus objetivos são alcançados é o ideal. Seguindo o exemplo do carro autônomo, é possível que,

para chegar a um destino no menor tempo possível, o carro prefira passar por uma estrada irregular de terra para economizar tempo de viagem do que uma estrada bem asfaltada um pouco mais longa, sem se importar com o conforto dos passageiros durante a viagem. Ele alcançou seu destino, mas provavelmente não da forma esperada ou ideal.

Os **agentes utilitários**, portanto, tem como característica a preocupação com os **diferentes modos que objetivos podem ser alcançados**. Neste caso, o agente é responsável por comparar estes estados e escolher aquele que irá **maximizar sua performance (utilidade)**. Para isso, os agentes contêm a chamada **Função Utilidade**, responsável por medir sua performance neste quesito.

Existem duas situações principais onde agentes utilitários tem clara vantagem. Segundo Russell & Norvig (2009), “Primeiro, quando houver **objetivos conflitantes**, apenas alguns dos quais podem ser alcançados (por exemplo, velocidade e segurança), a função utilidade especifica a escolha apropriada. Segundo, quando há **vários objetivos** que o agente pode visar **e nenhum dos quais pode ser alcançado com certeza**, a utilidade proporciona uma maneira em que a probabilidade de sucesso pode ser pesada em relação à importância dos objetivos”.

Retomando o exemplo onde o carro autônomo pode tomar dois caminhos para chegar a seu destino, mas desta vez com as rotas sendo uma mais longa onde o passageiro pode apreciar o mar e outra mais curta sem pontos de interesse, além disso, o passageiro precisa chegar a tempo em seu destino para uma reunião. Podemos então adicionar **probabilidade** ao problema: de acordo com a previsão do tempo, a primeira rota tem a possibilidade de estar com uma chuva forte, impedindo a visão do mar; enquanto a segunda passa por uma área onde ocorrem engarrafamentos ocasionais. Portanto, o agente deve levar em consideração os seguinte pontos na hora de escolher a rota:

- Qual a utilidade de uma vista pro mar?
- Qual a utilidade de chegar a tempo na reunião?
- A rota 1 permite chegar na reunião sem atrasos? Se sim, qual a utilidade de uma viagem mais rápida, quando ambas as rotas chegam antes do horário de reunião? Se não, é mais útil uma viagem com vista para o mar ou chegar a tempo para a reunião?
- Qual a probabilidade de chover na rota 1, atrapalhando a vista e ocasionando uma viagem ainda mais longa?

- Qual a probabilidade de ocorrer um engarrafamento na rota 2, fazendo essa rota perder sua única vantagem?
- Levando em consideração as probabilidades acima, quais as chances do passageiro se atrasar ao passar por cada uma das rotas?
- Qual rota tomar caso ambas causem atrasos?

O agente utilitário irá considerar essas perguntas e, de acordo com sua Função Utilidade, irá realizar as ações que maximizam sua **utilidade esperada**, isto é, a utilidade que o agente espera obter dadas as probabilidades e as utilidades de cada resultado. (Russell & Norvig, 2009)

## 4.5. Representação de estados do ambiente

O ambiente que um agente se encontra pode ter seus estados representados, em ordem crescente de **complexidade e expressividade**, como atômico, fatorial ou estruturado.

### 4.5.1. Representação atômica

Nesta representação, **cada estado é indivisível**, ou seja, ele existe mas não tem propriedades extras dentro de si, além de **não ter nada em comum com outros estados**. É a mais simples das três categorias.

Como exemplo temos um jogo de labirinto, onde cada posição do jogador é um estado; em um grid, por exemplo, o jogador pode sair da posição (estado) E5 para F5 e somente o nome de seu estado mudará.

### 4.5.2. Representação fatorial

Uma representação fatorada divide **cada estado em um conjunto fixo de variáveis ou atributos**, cada um dos quais pode ter um valor. Enquanto estados atômicos diferentes não têm nada em comum, estados fatorados diferentes **podem compartilhar alguns atributos e não compartilhar outros**. (Russell & Norvig, 2009)

Continuando com o exemplo do labirinto, desta vez o jogador pode ter uma certa quantidade de vidas, um número que representa sua energia

(quantidade de posições que ainda pode andar) e uma quantidade finita, porém incerta, de barras de cereal utilizadas para recuperar parte de sua energia.

### 4.5.3. Representação estruturada

Em situações mais complexas, onde elementos do ambiente podem se relacionar de diversas maneiras, ficaria **inviável ter atributos que representassem todas as possibilidades de conexão** entre os estados. Portanto, a representação estruturada modela os **estados como objetos complexos, capazes de se relacionar e interagir com outros estados**. (Russell & Norvig, 2009)

Um exemplo deste tipo de representação é encontrado em um jogo de xadrez, onde um estado pode ser dito como o conjunto das posições de todas as peças no tabuleiro em determinado momento de tempo. Neste caso, cada peça tem suas características de cor (preta ou branca) e tipo (peão, torre, cavalo, bispo, rainha ou rei), além das relações entre si (o bispo branco está em uma posição onde consegue capturar o cavalo preto, ao mesmo tempo em que está vulnerável para a torre preta, por exemplo).

## 5. Conclusão

A inteligência artificial é um campo de estudo com origem multidisciplinar, o qual se isolou academicamente por um período mas voltou a interagir com outras áreas devido aos benefícios que derivam de teorias externas à área em si. Sua história tem início em contos mitológicos e de ficção, os quais afetam o seu desenvolvimento até os dias de hoje.

O uso desta tecnologia é capaz de resolver incontáveis problemas, desde os mais simples, por agentes reativos simples, até os mais complexos, por agentes utilitários. Além disso, a capacidade de aprendizado destes agentes aumenta diariamente com o progresso de tecnologias como o Big Data, Deep Learning e a criação de algoritmos mais sofisticados.

Entretanto, é necessário exercer cuidado em seu uso, devido aos riscos éticos/econômicos/políticos/sociais que a IA traz junto com seus benefícios. Esses riscos podem surgir do uso indevido e imoral por pessoas mal intencionadas, por uma programação incompleta/errada ou por um aprendizado enviesado durante o treinamento de novas IAs.

Portanto, pode-se concluir que a inteligência artificial é uma tecnologia capaz de moldar o mundo, de formas benéficas ou maléficas. Ela é parte integral da sociedade tecnológica contemporânea e seu papel na vida humana só tende a crescer.



## 6. Referências

SOARES, Fabiano Araujo. Slides da aula 02 à 06. Apresentação do PowerPoint.

Russell, S. & Norvig, P. **Artificial Intelligence – A Modern Approach**. 3ª ed. Pearson Education, Inc. 2009.

O que é inteligência artificial (IA)? **IBM**. Disponível em: <https://www.ibm.com/br-pt/topics/artificial-intelligence>. Acesso em: 14 de nov. de 2024.

O que é : XOR Problem. **IA Tracker**. Disponível em: <https://iatracker.com.br/glossario/o-que-e-xor-problem/>. Acesso em: 15 de nov. de 2024.

FAVARON, Guilherme. Os Invernos da IA: Ciclos de Ascensão e Queda na História da Inteligência Artificial. **Guilherme Favaron**, 2024. Disponível em: <https://www.guilhermefavaron.com.br/post/os-invernos-da-ia-ciclos-de-ascensao-e-queda-na-historia-da-inteligencia-artificial>. Acesso em: 15 de nov. de 2024.

O que é retropropagação?. **IBM**. Disponível em: <https://www.ibm.com/br-pt/think/topics/backpropagation>. Acesso em: 15 de nov. de 2024.

O que são redes neurais bayesianas?. **Databricks**. Disponível em: <https://www.databricks.com/br/glossary/bayesian-neural-network>. Acesso em: 15 de nov. de 2024.

O que é Big Data?. **Oracle**. Disponível em: <https://www.oracle.com/br/big-data/what-is-big-data/>. Acesso em: 15 de nov. de 2024.

O que é deep learning?. **IBM**. Disponível em: <https://www.ibm.com/br-pt/topics/deep-learning>. Acesso em: 15 de nov. de 2024.

One Hundred Year Study on Artificial Intelligence (AI100). **Stanford University**. Disponível em: <https://ai100.stanford.edu/>. Acesso em: 16 de nov. de 2024.

AI Index Annual Report, Welcome to the 2024 AI Index Report. **Stanford University**. Disponível em: <https://aiindex.stanford.edu/report/>. Acesso em: 16 de nov. de 2024.

O que é reconhecimento de fala?. **IBM**. Disponível em: <https://www.ibm.com/br-pt/topics/speech-recognition>. Acesso em: 16 de nov. de 2024.

O que são agentes de IA?. **AWS Amazon**. Disponível em: <https://aws.amazon.com/pt/what-is/ai-agents/>. Acesso em: 16. de nov. de 2024.