

CURSO

# INTELIGÊNCIA ARTIFICIAL NA SAÚDE



CURSO  
INTELIGÊNCIA  
ARTIFICIAL NA  
SAÚDE



Secretaria  
da Saúde



GOVERNO  
PERNAMBUCO  
ESTADO DE PERNAMBUCO

## EXPEDIENTE

### **Governo do Estado de Pernambuco**

Governadora | Raquel Teixeira Lyra Lucena

### **Secretaria Estadual de Saúde de Pernambuco**

Secretária | Zilda do Rego Cavalcanti

### **Secretaria Executiva de Gestão do Trabalho e da Educação na Saúde**

Secretária Executiva | Bruno Alves Carneiro

### **Escola de Governo em Saúde Pública de Pernambuco**

Diretora | Célia Maria Borges da Silva Santana

Gerente | Luciana Camêlo de Albuquerque

### **Coordenação de Educação Permanente em Saúde da Escola de Governo em Saúde Pública de Pernambuco**

Coordenadora | Emmanuely Correia de Lemos

### **Equipe Técnica da Coordenação de Educação Permanente**

Andréa Fernanda de Oliveira | Bárbara Paloma Marques de Luna | Natália Nunes de Lima | Nathalia Ingrid dos Santos Silva Lucena | Roberta Natalie de Andrade Santos | Wanda das Graças Silva de Barros

### **Residentes da Coordenação de Educação Permanente**

Beatriz da Silva Lemos | João Lucas Pessoa de Vasconcelos | Maria Clara Freitas Monteiro | Matheus Gomes Magalhães Leiros

### **Coordenação de Ensino à Distância da Escola de Governo em Saúde Pública de Pernambuco**

Coordenadora | Lorena Franco Sobral

### **Equipe Técnica da Coordenação de Ensino à Distância**

Adnan Gomes de Moraes | Anefátima Figueiredo | Arnaldo César Alencar da Boaviagem | Roberta Mendes Batista de Oliveira | Romero Sarmento Serrano

### **Autor**

Álvaro Farias Pinheiro

### **Design e Diagramação**

Romero Sarmento Serrano



Ficha Catalográfica elaborada pela Biblioteca Nelson Chaves (ESPPE), com os dados fornecidos pelo autor.

P452c	Pernambuco. Secretaria Estadual de Saúde. Secretaria Executiva de Gestão do Trabalho e Educação na Saúde. Escola de Governo em Saúde Pública de Pernambuco. Coordenação de Educação Permanente em Saúde da Escola de Governo em Saúde Pública.
	Curso inteligência artificial na saúde / Secretaria Estadual de Saúde de Pernambuco. Secretaria Executiva de Gestão do Trabalho e da Educação na Saúde. Escola de Governo em Saúde Pública de Pernambuco. Coordenação de Educação Permanente em Saúde da Escola de governo em Saúde Pública de Pernambuco. _ Recife: SES/ESPPE, 2025. 88 p.: il.
	1. Gestão de Informação. 2. Inteligência Artificial – Saúde. 3. Tecnologia em Saúde. 4. Saúde Pública. 5. Curso. I. Título.
ESPPE / BNC	CDU- 004.8: 37:614 (813.4)

Bibliotecária Responsável: Anefátima Figueiredo – CRB-4/P-1488



# INTRODUÇÃO

Sejam bem-vindos ao Curso Autoinstrucional Inteligência Artificial na Saúde, promovido pela Escola de Governo em Saúde Pública de Pernambuco (ESPPE).

Autor: Álvaro Farias Pinheiro atua como Diretor Geral de Inovação e Inteligência Artificial em Saúde (DGIAS) na Secretaria Estadual de Saúde de Pernambuco (SES/PE). Doutor em Engenharia da Computação pela Universidade de Pernambuco (UPE-PPGEC), com pós-doutorado em Inteligência Artificial na UFPE, Mestre, MBA, Especialista e Bacharel em Computação. Membro da Câmara de Governo Digital de Pernambuco (CGD/PE) e Analista em Gestão de TIC (AGTIC) na Agência Estadual de Tecnologia da Informação e Comunicação (ATI). Segue o LinkTree: <https://linktr.ee/alvarofpinheiro>.

Este curso tem como objetivo apresentar o papel da Inteligência Artificial na saúde pública, abordando como essa tecnologia pode transformar processos, otimizar o atendimento, contribuir para a tomada de decisões e ampliar o acesso e a qualidade dos serviços oferecidos à população. E tem como público prioritário os(as) agentes públicos da Secretaria Estadual de Saúde.

Neste curso, você será introduzido aos conceitos fundamentais de hiperautomação, explorando como a integração de Artificial Intelligence (AI) – Inteligência Artificial, Machine Learning (ML) – Aprendizado de Máquina, Deep Learning (DL) – Aprendizado Profundo e Large Language Model (LLM) – Grande Modelo de Linguagem, está revolucionando a área da saúde pública. Serão abordadas as principais classes de problemas solucionáveis com essas tecnologias, como agrupamento, associação, classificação, regressão, otimização e recomendação.

Além disso, o conteúdo apresenta uma visão abrangente sobre diferentes tipos de algoritmos utilizados para resolver esses desafios: algoritmos estatísticos, árvores de decisão, métodos evolucionários, técnicas baseadas em enxames e redes neurais artificiais. A proposta é instruir você a identificar e aplicar essas abordagens para aprimorar processos, automatizar tarefas complexas e tomar decisões mais precisas no contexto da saúde pública digital.

O curso é composto de quatro módulos: Transformação Digital na Saúde; Classes de Algoritmos e Tipos de Aprendizado de Máquina (Classes de Algoritmos); Engenharia de Prompt; e, Ferramentas Hardcode, Lowcode e Nocode.

O Módulo 1: O módulo terá como eixo central a transformação digital, entendida como o movimento estratégico pelo qual o setor público, em especial a saúde, precisa se adequar para incorporar tecnologias emergentes que ampliem eficiência, qualidade e capacidade de resposta. A transformação digital estabelece o contexto

necessário para compreender porque novas ferramentas e métodos passaram a ser essenciais, permitindo que gestores e equipes lidem com demandas crescentes, processos complexos e altos volumes de dados característicos da administração pública contemporânea.

A partir dessa contextualização, o módulo apresentará e definirá os conceitos fundamentais de hiperautomação, Inteligência Artificial (IA) e automação robótica de processos (RPA). A hiperautomação será explorada como uma abordagem ampliada da automação, que combina diversas tecnologias para automatizar processos de ponta a ponta, promovendo agilidade, precisão e inovação. Nesse cenário, a IA desempenha um papel estruturante, pois possibilita a análise inteligente de informações, o reconhecimento de padrões, a tomada de decisão baseada em evidências e a automação de fluxos que antes dependiam exclusivamente do trabalho humano.

Com isso, o módulo permitirá compreender de que forma RPA, IA e hiperautomação se articulam dentro da agenda de transformação digital, evidenciando como essas tecnologias podem aprimorar o atendimento ao cidadão, otimizar processos internos, detectar fraudes, qualificar triagens e acelerar respostas institucionais. Dessa forma, o conteúdo reforça que dominar esses fundamentos é indispensável para o desenvolvimento de uma administração pública mais eficiente, inovadora e orientada a resultados.

O Módulo 2: Para compreender a transformação digital na administração pública, é fundamental reconhecer os principais tipos de algoritmos utilizados em IA e suas diferentes formas de aprendizado. Os algoritmos de IA podem ser classificados conforme o tipo de problema que buscam resolver e, especialmente, pelo modo como aprendem: aprendizado supervisionado, aprendizado não supervisionado, aprendizado por reforço e aprendizado semisupervisionado.

No aprendizado supervisionado, algoritmos são treinados com exemplos rotulados, permitindo que desenvolvam a capacidade de prever resultados a partir de novos dados. Já o aprendizado não supervisionado emprega dados sem rótulos, possibilitando a descoberta de padrões e agrupamentos sem intervenção direta. O aprendizado semisupervisionado combina características dos métodos supervisionados e não supervisionados, utilizando tanto dados rotulados quanto não rotulados durante o treinamento dos algoritmos. O aprendizado por reforço, por sua vez, envolve algoritmos que aprendem a partir de interações com o ambiente, recebendo recompensas ou penalidades conforme suas ações. No aprendizado semisupervisionado, os algoritmos utilizam uma combinação de dados rotulados e não rotulados, aproveitando o conhecimento dos exemplos identificados para melhorar a capacidade de generalização e extrair padrões dos dados ainda sem classificação.

Reconhecer essas diferentes abordagens é essencial para identificar estratégias adequadas de automação e inovação nos serviços públicos, promovendo a eficiência, a transparência e a capacidade de resposta às demandas da sociedade por meio da utilização inteligente de IA.

O Módulo 3: No contexto da transformação digital na administração pública, destaca-se a importância de aplicar a Engenharia de Prompt (técnica para criar comandos claros que servem para orientar a Inteligências Artificiais a gerar respostas mais precisas e úteis) com os mais avançados modelos de IA, conhecidos como Large Language Models (LLMs), como GPT e Gemini. A utilização estratégica dessas ferramentas permite potencializar soluções inovadoras para desafios enfrentados pelo setor público, tornando os processos mais inteligentes, flexíveis e acessíveis. Ao estruturar prompts de maneira eficiente, é possível direcionar essas IAs para gerar análises, automatizar tarefas complexas e apoiar a tomada de decisão baseada em grandes volumes de dados. Essa integração entre Engenharia de Prompt e LLMs representa um passo fundamental para que a administração pública alcance maior eficiência, promova a inovação e dê respostas mais rápidas e precisas às demandas da sociedade.


O Módulo 4: Com o avanço da hiperautomação e da Inteligência Artificial, aplicar algoritmos de IA utilizando as mais diversas ferramentas desde plataformas Hardcode, Lowcode até Nocode tornou-se fundamental para solucionar uma ampla gama de desafios no setor público. Segue uma breve explicação das categorias de ferramentas com o detalhamento sendo dado no módulo 4. Ferramentas hardcode são caracterizadas pelo desenvolvimento tradicional, onde a programação é feita de forma manual e detalhada por desenvolvedores, exigindo conhecimento aprofundado em linguagens de código. Já as ferramentas lowcode permitem a criação de soluções com pouca escrita de código, utilizando interfaces gráficas e componentes pré-configurados para acelerar o desenvolvimento e facilitar ajustes. Por fim, ferramentas nocode possibilitam que usuários sem experiência em programação criem aplicações inteiras apenas por meio de interfaces visuais, eliminando completamente a necessidade de codificação.

Essas soluções tecnológicas permitem abordar problemas que vão desde a análise preditiva de demandas sociais, passando pelo processamento inteligente de grandes volumes de dados, até a automação de tarefas administrativas e operacionais. Ao integrar diferentes métodos de aprendizado de máquina com ferramentas inovadoras, a administração pública potencializa a eficiência, a transparência e a capacidade de resposta, adaptando-se rapidamente às necessidades da sociedade e promovendo a inovação nos serviços governamentais.

Ao compreender essa jornada de inovação e as inúmeras possibilidades proporcionadas pela hiperautomação e pela Inteligência Artificial, é possível visualizar uma transformação profunda nos serviços públicos, especialmente na saúde. Tais avanços demandam uma abordagem estratégica, capaz de alinhar tecnologia, processos e pessoas de maneira ética, transparente e orientada a resultados. Nesse contexto, o primeiro módulo do curso oferece as bases conceituais e práticas essenciais para explorar o potencial com a visão estratégica, preparando o terreno para a aplicação efetiva dessas tecnologias na melhoria do atendimento ao cidadão e na modernização da gestão pública.



## **Curso Autoinstrucional em Inteligência Artificial na Saúde - Apresentação do Curso**

[Clique Aqui](#) 

## SUMÁRIO

<b>Módulo 1: A Transformação Digital na Saúde</b>	06
Vídeo Aula - Módulo 1	31
<b>Módulo 2: Classes de Problemas e Tipos de Aprendizado</b>	32
Vídeo Aula - Módulo 2	44
<b>Módulo 3: Engenharia de <i>Prompts</i></b>	45
Vídeo Aula - Módulo 3	66
<b>Módulo 4 - Ferramentas Digitais na Saúde</b>	67
Vídeo Aula - Módulo 4	82

# MÓDULO 1

## A TRANSFORMAÇÃO DIGITAL NA SAÚDE



## APRESENTAÇÃO

A incorporação da Inteligência Artificial (IA) e da hiperautomação à saúde pública representa um salto transformacional na construção de sistemas mais eficientes, integrados e inteligentes, capazes de responder com agilidade às demandas crescentes da sociedade contemporânea. Ao unir tecnologias como automação robótica de processos (RPA), aprendizado de máquina, processamento de linguagem natural (PLN) e análise avançada de dados clínicos e administrativos, a saúde pública passa a dispor de instrumentos estratégicos para enfrentar desafios complexos desde o gerenciamento massivo de informações até a otimização de fluxos assistenciais e a tomada de decisão baseada em evidências.

Sob uma perspectiva estratégica, a IA aplicada à saúde redefine paradigmas de gestão, promovendo uma alocação mais inteligente de recursos, ampliando a personalização do atendimento ao cidadão e reduzindo barreiras burocráticas. O potencial de automatizar desde operações rotineiras até análises preditivas de riscos epidemiológicos e desfechos clínicos cria possibilidades para o monitoramento em tempo real, a prevenção de agravos e a resposta rápida a emergências sanitárias.

Este módulo oferece uma visão abrangente e aplicada sobre como a hiperautomação inteligente, quando implementada de forma ética, transparente e centrada no paciente, pode impulsionar a inovação, fortalecer a governança pública e aumentar a resiliência dos sistemas de saúde. Mais do que uma revolução tecnológica, trata-se de uma transformação cultural e institucional orientada para resultados concretos, sustentáveis e humanizados, em que a IA se consolida como aliada essencial da saúde pública do futuro.

## HIPERAUTOMAÇÃO

Hiperautomação (Hyperautomation, do inglês) é um conceito emergente no campo da transformação digital que representa a automação extensiva de processos de negócios por meio da integração de múltiplas tecnologias avançadas, como automação Artificial Intelligence (AI) – Inteligência Artificial (IA), Robotic Process Automation (RPA) – Robótica de Processos e Business Process Management (BPM) – Gestão de Processos de Negócios. O termo foi popularizado pelo Gartner (2020), que o descreve como a “expansão da automação em organizações, combinando ferramentas e tecnologias para aumentar significativamente a eficiência, agilidade e inovação”.

A hiperautomação não se limita à automação de tarefas repetitivas e operacionais, como inserção de dados, a geração de relatórios e o preenchimento de formulários, mas busca automatizar processos complexos de ponta a ponta, promovendo ganhos substanciais de produtividade, redução de custos, melhoria da experiência do usuário e maior capacidade de adaptação frente às demandas do ambiente dinâmico. Um exemplo de automação de processos complexos na saúde pública seria a análise preditiva e gestão de campanha de vacinação, no qual a integração de tecnologias avançadas pode otimizar todas as etapas logísticas, mas também realizar a previsão de surtos.

A integração de múltiplas tecnologias digitais vem transformando a forma como os sistemas de saúde operam e tomam decisões. A combinação entre Inteligência Artificial (IA), automação robótica de processos (RPA), gerenciamento de processos de negócio (BPM) e análise avançada de dados permite automatizar tarefas que antes exigiam tempo e esforço humano, liberando profissionais para atividades mais estratégicas e voltadas ao cuidado com o paciente. Nesse sentido, as principais aplicabilidades da hiperautomação são:

- Nos serviços de saúde, essa automação estendida ultrapassa a simples execução de rotinas. Ela se aplica a processos como o agendamento inteligente de consultas, o processamento automático de laudos e exames laboratoriais, o monitoramento de pacientes em tempo real e a análise preditiva de risco clínico. Ao integrar dados de diversas fontes, como prontuários eletrônicos, sistemas de regulação e plataformas de vigilância epidemiológica, torna-se possível identificar padrões e antecipar situações críticas, como o aumento de casos de determinadas doenças em uma região.
- A adaptabilidade e a inteligência desses sistemas permitem o aprendizado contínuo a partir dos dados, possibilitando o aperfeiçoamento de protocolos assistenciais e o uso mais racional dos recursos públicos. Isso pode ser observado em soluções que ajustam automaticamente o fluxo de atendimentos de acordo com a demanda ou em algoritmos que recomendam condutas clínicas baseadas em evidências atualizadas e no histórico de cada paciente.

- Os princípios de transparência e governança garantem que cada processo automatizado seja rastreável e auditável, promovendo confiança e conformidade com as normas éticas e legais que regem a gestão pública. O acompanhamento das decisões automatizadas e a supervisão humana continuam sendo elementos essenciais para assegurar a responsabilidade e o uso ético das tecnologias digitais em saúde.
- Com a hiperautomação aplicada à saúde, as instituições tornam-se mais eficientes e analíticas, fortalecendo sua capacidade de resposta, prevenção e inovação. O resultado é um sistema de saúde mais integrado, sustentável e centrado no cidadão, capaz de oferecer serviços com maior qualidade, agilidade e equidade.

A hiperautomação é vista como um dos pilares da transformação digital na administração pública, permitindo aos governos inovarem na prestação de serviços, aumentar a transparência e melhorar o atendimento ao cidadão. Ao integrar tecnologias como IA e RPA, o setor público consegue automatizar desde atividades administrativas rotineiras até processos de análise preditiva e tomada de decisão baseada em dados (Pinheiro; Santos; Lima Neto, 2023).

Em resumo, hiperautomação é a abordagem estratégica que une tecnologias de automação e Inteligência Artificial para criar processos inteligentes, flexíveis e escaláveis, promovendo uma verdadeira revolução na forma como organizações públicas e privadas operam.

A hiperautomação, ao proporcionar processos cada vez mais inteligentes e integrados, se configura como um elemento fundamental dentro do contexto mais amplo da Transformação Digital (TD). Enquanto a hiperautomação potencializa a automação e o uso de Inteligência Artificial para aprimorar a eficiência e a tomada de decisão nas instituições, ela também serve de base tecnológica para a implementação de políticas públicas inovadoras. Dessa forma, a TD no setor público encontra na hiperautomação um caminho estratégico para promover serviços mais ágeis, transparentes e centrados no cidadão, consolidando um governo digital moderno e eficiente.

Segundo Pinheiro (2023), Transformação Digital (TD) consiste em soluções tecnológicas para resolver problemas. No setor público, a TD busca aplicar tecnologia para tornar as políticas públicas mais eficientes, eficazes e efetivas na solução dos desafios sociais. Tendo como objetivo gerar valor público ao usar tecnologia para melhorar serviços governamentais, tornando-os menos burocráticos e mais transparentes, além de aproximar governos dos cidadãos.

A Estratégia de Governo Digital (EGD) utiliza a Transformação Digital para orientar a administração e promover competitividade e produtividade em todos os setores econômicos, impulsionando processos produtivos e sociais por meio de um planejamento plurianual (Pinheiro; Santos, 2021).

Um exemplo de Transformação Digital no Setor Público é o país da Estônia, frequentemente citada devido à priorização da TD em praticamente todos os serviços públicos nos últimos anos. Para isso, foram realizados investimentos em segurança da informação e na formação dos servidores. Como resultado, o cidadão estoniano precisa comparecer presencialmente ao Estado apenas em três situações: casamento, divórcio e transferência de bens, sendo os demais serviços disponibilizados digitalmente (Pinheiro, 2022).

Para facilitar o entendimento da Transformação Digital, serão apresentados os principais recursos tecnológicos disponíveis. Existem duas tecnologias essenciais para a Transformação Digital: Business Intelligence (BI), que realiza análise descritiva dos dados, explicando problemas passados; e, Business Analytics (BA), que faz análise preditiva para identificar problemas antes que ocorram. BI e BA utilizam modelos estatísticos e de Inteligência Artificial para analisar dados. Isso ajuda tanto a entender problemas passados quanto a prever possíveis desafios futuros.

Além do BI e BA, o Business Process Management (BPM) possibilita o desenho dos fluxos de processos de negócio utilizando uma notação universal e recursos nocode (ou seja, sem a necessidade de codificação), com o objetivo de mapear e otimizar processos. Isso é viabilizado por meio da colaboração entre especialistas em tecnologia e negócios. E, a integração dessas tecnologias ao RPA permite automatizar parcial ou totalmente processos, proporcionando maior rapidez e precisão.

De acordo com Klaus Schwab (2018), a Quarta Revolução Industrial representa a fusão dos mundos físico e virtual por meio de novas tecnologias, como Inteligência Artificial e realidade aumentada, resultando na criação do metaverso. Segundo Schwab, a nova Revolução Industrial pode gerar valor público se houver responsabilidade coletiva na criação de um futuro no qual inovação e tecnologia estejam a serviço das pessoas, promovendo avanços na consciência moral da sociedade.

A Quarta Revolução Industrial, também chamada de Indústria 4.0, não possui definição única devido à sua abrangência. Em geral, refere-se à integração de tecnologias digitais, como Internet das Coisas e Computação em Nuvem, para automatizar processos e otimizar a troca de dados visando eficiência nos sistemas produtivos.

Uma revolução caracteriza-se por seu impacto significativo, distinguindo-se da evolução devido à sua natureza abrupta e não gradual. Trata-se de um processo que frequentemente desloca indivíduos de situações estáveis, resultando em transformações profundas na sociedade, economia e cultura.

## VOCÊ SABIA?

Os impactos das revoluções industriais marcaram profundamente a sociedade:

- **1ª Revolução Industrial:** Introduziu a mecanização por água e vapor, substituindo o transporte a cavalo e encurtando distâncias;
- **2ª Revolução Industrial:** A eletrificação permitiu produção em massa, tornando bens e serviços mais acessíveis e substituindo o trabalho artesanal pela manufatura;
- **3ª Revolução Industrial:** Com automação e eletrônica digital, iniciou-se a era da informação, eliminando cálculos e controles manuais;
- **4ª Revolução Industrial:** Tecnologias como Aprendizado de Máquina são usadas para otimizar processos de agrupamento, associação, classificação, predição, otimização e recomendação na atualidade (Schwab, 2018).

As mudanças trazidas pela revolução são inevitáveis, como em todas as anteriores. Quem for flexível sentirá menos impacto e se adaptará mais facilmente, quem resistir enfrentará mais dificuldades, mas também precisará se adequar.

Para o desenvolvimento do Governo 4.0, determinadas tecnologias foram essenciais para viabilizar a Era do Conhecimento: a presença de um grande volume de dados disponíveis em nuvem; a variedade de tipos de dados, incluindo dados estruturados, não estruturados e semiestruturados; e, a velocidade no acesso a esses dados, proporcionada pela tecnologia de Big Data. A análise adequada de todos esses dados disponíveis, especialmente os não estruturados como textos, imagens, áudios e vídeos, depende do uso de Inteligência Artificial.

Com algoritmos de Machine Learning (Aprendizado de Máquina) aplicados a esses dados, é possível identificar tendências a partir das informações rapidamente disseminadas pelas Big Techs, que armazenam e disponibilizam dados em grandes data centers por meio da Cloud Computing (Computação em Nuvem).

**Big Data** refere-se ao grande volume, variedade e velocidade de dados gerados continuamente por diferentes fontes digitais, exigindo ferramentas avançadas.

**Big Tech** são as grandes empresas de tecnologia, como Google, Amazon, Facebook e Microsoft, que dominam o mercado global e influenciam fortemente as soluções digitais.

**Data Center** é a infraestrutura física composta por servidores e equipamentos de rede utilizada para armazenar, processar e distribuir dados de forma segura e escalável.

Adaptar-se às mudanças envolve compreender novas tecnologias aplicadas ao Governo e à Indústria 4.0, destacando a importância da computação em nuvem associada a processos físicos. Essa integração permite gerenciar processos com sistemas distintos que se ajustam em tempo real, inspirando a denominação Era do Conhecimento na convenção de Davos em 2016. Em função do potencial da Cloud Computing, observa-se uma tendência crescente da sociedade em adotar a lógica de conectar dispositivos à Internet e interligá-los entre si por meio de uma vasta rede composta por coisas e pessoas. Essa rede possibilita a coleta e o compartilhamento de dados sobre o usuário e o ambiente com o qual interage, conceito este conhecido como Internet of Things (IoT).

O uso dessas tecnologias inovadoras permite que o Governo 4.0 (aplicação de tecnologias digitais avançadas, como computação em nuvem, Big Data, Inteligência Artificial e Internet das Coisas, para tornar a gestão pública mais eficiente) compartilhe conhecimento de forma eficiente e o aplique, por exemplo, em criações virtuais, que podem ser transformadas em objetos tridimensionais por meio da impressão 3D, como próteses e implantes personalizados. Esse processo torna a fabricação mais rápida, barata e flexível quanto ao tipo, material e tamanho dos produtos. O uso da tecnologia de impressão 3D na saúde pode ter seu uso desde uma ferramenta didática para o ensino com a utilização de modelos 3D até a impressão de tecidos e órgãos, diminuindo as filas para transplantes.

O Governo 4.0 representa uma nova era da gestão pública, marcada pela adoção estratégica de tecnologias digitais avançadas, como computação em nuvem, Big Data, Inteligência Artificial e Internet das Coisas. Essa abordagem visa tornar os serviços governamentais mais eficientes, transparentes e personalizados, permitindo o gerenciamento inteligente de grandes volumes de dados e a tomada de decisões em tempo real. A digitalização dos processos administrativos facilita a integração entre diferentes setores, promove a inovação e amplia o acesso dos cidadãos a serviços de qualidade, consolidando o papel do governo como agente ativo na transformação digital da sociedade.

O aumento da utilização de sensores inteligentes, a exemplo dos smartwatches, dispositivos projetados para captar estímulos externos do ambiente, possibilita o processamento eficiente dos dados coletados. Por meio da aplicação de algoritmos de Machine Learning e Deep Learning (conceitos serão detalhados à seguir) essas informações são convertidas em dados relevantes acerca do objeto monitorado.

Nesse contexto de transformação digital acelerada pela hiperautomação, torna-se fundamental compreender as bases conceituais e tecnológicas que sustentam essas mudanças. A evolução de sistemas inteligentes, impulsionada pela Automação Robótica (RPA), Inteligência Artificial (IA), pelo Aprendizado de Máquina (ML), pelo Deep Learning (DL) e pelos Modelos de Linguagem de Grande Escala (LLMs), está no centro dessa revolução, promovendo soluções inovadoras para desafios cada vez mais complexos.

A seguir, será detalhada cada uma dessas abordagens, suas principais características e aplicações, estabelecendo o alicerce necessário para aprofundar o entendimento dos algoritmos e práticas que impulsionam a nova era do conhecimento.

## AUTOMAÇÃO ROBÓTICA DE PROCESSOS (RPA)

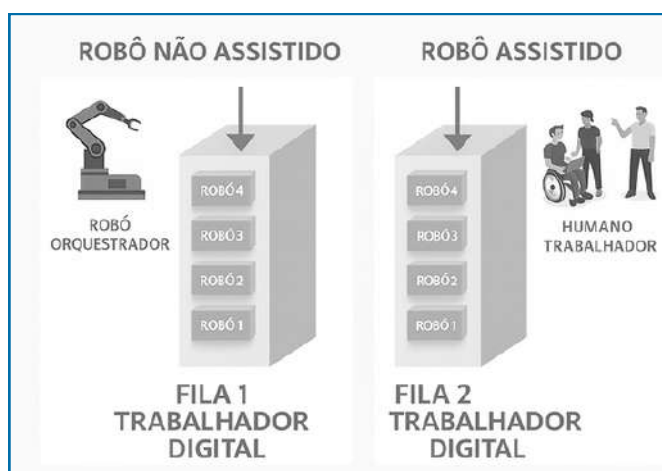
O tema da Automação Robótica de Processos (do inglês, Robotic Process Automation - RPA) é especialmente relevante para gestores, gerentes e profissionais de saúde, pois oferece ferramentas estratégicas para otimizar a rotina administrativa e assistencial dentro das instituições. Gestores podem utilizar o RPA para garantir maior eficiência e transparência na gestão de recursos, reduzindo falhas e custos operacionais ao automatizar tarefas burocráticas e repetitivas. Já gerentes e profissionais de saúde podem se beneficiar ao direcionar seu tempo para atividades de maior valor agregado, como análise de indicadores, tomada de decisão baseada em dados e atendimento ao paciente, enquanto os processos automatizados cuidam da integração de sistemas, atualização de cadastros e geração de relatórios em tempo real. Dessa forma, o RPA contribui diretamente para uma administração mais ágil, segura e centrada no cidadão, promovendo avanços significativos na qualidade dos serviços de saúde.

Embora o foco principal deste curso seja o uso da Inteligência Artificial (IA) aplicada à saúde, é fundamental compreender os conceitos básicos de Automação Robótica de Processos (RPA), pois eles servem como base para a evolução das tecnologias inteligentes. O entendimento de como os robôs digitais automatizam tarefas e integram sistemas permite compreender com maior profundidade os princípios e o funcionamento da IA, já que muitas soluções inteligentes partem da automatização de processos para, posteriormente, incorporar algoritmos avançados de aprendizado.

Assim, todas as informações abordadas neste tópico são essenciais para estabelecer uma conexão sólida com os conceitos de IA. Desse modo, compreender os fundamentos da Automação Robótica de Processos facilita a assimilação dos próximos conteúdos sobre robôs de software e suas aplicações práticas, permitindo visualizar como essas ferramentas evoluem e se integram ao contexto da transformação digital na saúde.

RPA (Robotic Process Automation), ou automação robótica, refere-se à automação de processos de negócio por meio de robôs de software, também chamados de trabalhadores digitais ou bots, conhecida como “Robótica de Software”. O RPA não deve ser confundido com o “Software de Robô”, utilizado para controlar robôs mecânicos, como os braços robóticos presentes na indústria automobilística (AIMDEK TECHNOLOGIES, 2022).

Figura 1: Tipos de Robôs.



Os RPA realizam tarefas repetitivas e volumosas nos processos de negócio, como o tratamento de grandes volumes de dados e documentos, permitindo que os trabalhadores foquem em atividades mais cognitivas. Assim, bots automatizam tarefas repetitivas e propensas a erro, gerenciando etapas dos processos de negócios que seguem regras e resultam em serviços ou produtos.

Fonte: Pinheiro, 2024.

Existem dois tipos principais de robôs utilizados na automação de processos: o assistido (ou atendido) e o não assistido (ou não atendido). O robô assistido requer intervenção humana para funcionar. Um usuário precisa clicar em um botão ou selecionar uma opção no menu para que o robô execute as atividades programadas. Já o não assistido é iniciado automaticamente por outro robô, chamado de orquestrador, controlador ou gatilho, que define a data e hora de início conforme determinado pelo administrador dos robôs.

A automação robótica de processos tem se mostrado uma das ferramentas mais poderosas no avanço da transformação digital na saúde. Seu princípio de funcionamento baseia-se na execução estruturada e sequencial de tarefas, garantindo que cada atividade ocorra na ordem correta e dentro de fluxos bem definidos. Esses robôs digitais seguem a lógica de filas de execução, processando as demandas conforme são recebidas e concluindo-as antes de iniciar as próximas.

Na prática, essa tecnologia permite automatizar atividades rotineiras e repetitivas que, no contexto da saúde, consomem tempo das equipes e estão sujeitas a erros humanos. Um exemplo é o processamento de autorizações de procedimentos médicos, que envolve análise de dados, verificação de cadastros e checagem de conformidade com normas. Com o uso de automação, essas etapas podem ser realizadas de forma contínua, reduzindo o tempo de resposta e liberando profissionais para funções mais analíticas e assistenciais.

A automação também tem papel importante na integração de sistemas clínicos e administrativos. Plataformas de regulação de leitos, faturamento hospitalar, controle de medicamentos e monitoramento epidemiológico podem operar de maneira sincronizada, utilizando robôs que transferem dados entre sistemas, atualizam planilhas e alimentam painéis de gestão. Isso aumenta a confiabilidade das informações e reduz a sobrecarga de trabalho operacional.

Ferramentas de automação podem ser implementadas de duas formas. Quando são assistidas, a execução depende da ação de um usuário, como em casos em que um profissional precisa validar dados sensíveis de pacientes antes de prosseguir. Já nas automações não assistidas, os robôs executam processos completos de forma autônoma, sem necessidade de intervenção humana, o que é útil em operações contínuas, como o envio diário de relatórios de vigilância ou a triagem automatizada de exames laboratoriais.

Essas soluções podem ser desenvolvidas tanto por meio de linguagens de programação, como Python, quanto com o uso de plataformas de baixa ou nenhuma codificação, conhecidas como Lowcode e Nocode. Essas ferramentas oferecem interfaces visuais intuitivas que permitem criar fluxos de automação por meio da seleção e conexão de componentes. Isso torna o desenvolvimento acessível a profissionais da saúde com conhecimento técnico básico, estimulando a inovação dentro das próprias instituições.

A aplicação da automação e da robótica na saúde não apenas potencializa a eficiência dos processos, mas também amplia e fortalece a capacidade analítica das equipes ao permitir que profissionais dediquem mais tempo à análise de dados e à tomada de decisões estratégicas. Ao automatizar tarefas rotineiras e repetitivas, os robôs liberam os colaboradores para se concentrarem em atividades de maior complexidade intelectual, como interpretação de informações clínicas, identificação de padrões e antecipação de necessidades dos pacientes. Dessa forma, as equipes podem utilizar ferramentas analíticas avançadas, explorar indicadores de desempenho e promover melhorias contínuas baseadas em dados concretos, elevando o nível de qualidade e precisão nos serviços prestados, amplia a segurança da informação e permite respostas mais rápidas em situações críticas. Ao combinar essas tecnologias com a Inteligência Artificial, hospitais e secretarias de saúde podem avançar para modelos de gestão mais preditivos, integrados e centrados no cuidado ao cidadão.

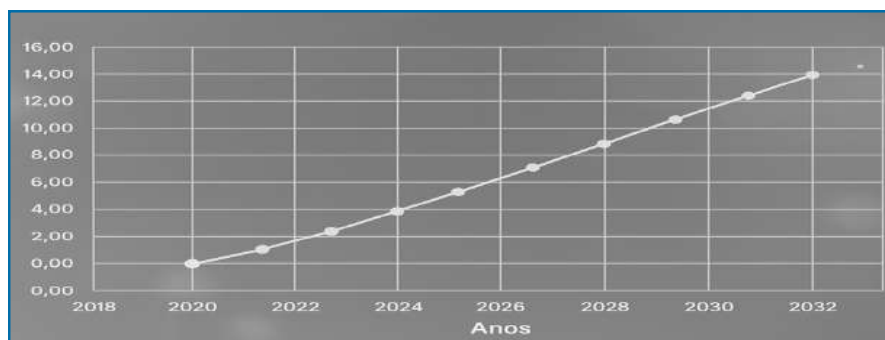
O trabalhador digital ou robô de software ou bot é, de forma simples, um script (um código de programação) que define o passo a passo de uma lógica que estava sendo executada por um trabalhador humano e que passa a ser atribuída para uma máquina, integrando sistemas ao atuar como uma camada intermediária que imita o uso humano de softwares, realizando processamentos e aproveitando recursos de TI já existentes.

A automação robótica conecta Application Programming Interface (API), isto é, tecnologia de interoperabilidade (conexão entre sistemas), serviços de IA e aplicativos, configurando regras nos próprios sistemas e permitindo que robôs transcrevam ou processem dados no back-end (é a parte do sistema responsável pelo armazenamento de dados) de forma a garantir o histórico da utilização do sistema.

Assim, se existe um processo repetitivo, é possível automatizá-lo desenvolvendo um script que o replique, deixando o trabalhador humano focado nas atividades cognitivas do processo. Nesse contexto, segue uma série de benefícios obtidos com o uso do RPA em uma organização como, por exemplo:

diminuição de custos, velocidade de execução, precisão das tarefas executadas, escalabilidade dos processos e aplicação de segurança. Nesse sentido, a automação permite que as tarefas sejam realizadas de acordo com a performance fornecida pela infraestrutura, sem variações causadas por fadiga. A precisão ocorre porque o trabalhador digital executa tarefas conforme programado. Já a escalabilidade permite aumentar o volume de trabalho sem elevar custos ou perder precisão.

**Figura 2:** Evolução no uso de RPAs.



Fonte: HBR, 2025.

O uso de robôs também adiciona camadas de segurança, tornando as atividades mais confiáveis. Devido a esses benefícios, o uso de RPA nas organizações tem aumentado, como mostra um estudo da Harvard Business Review (HBS) de 2015 apresentado no gráfico ao lado.

Apesar de muitas pessoas associarem o uso de RPA ao risco imediato de desemprego, imaginando que trabalhadores serão substituídos por robôs, essa visão não captura toda a complexidade do tema. No Brasil, onde grande parte da população ocupada tem baixa qualificação e enfrenta desigualdades históricas, é natural que esses receios apareçam com ainda mais força. No entanto, estudos mais recentes (Arntz et al., 2016), tanto nacionais quanto internacionais, mostram um cenário mais equilibrado: a automação costuma transformar funções, não simplesmente eliminá-las. Em muitos casos, ela libera os profissionais das tarefas repetitivas e abre espaço para que atuem em atividades mais analíticas, criativas ou de maior valor agregado, desde que haja investimento em qualificação. Pesquisas de instituições como OCDE, BID e IPEA (OCDE, 2019; BID 2020; IPEA, 2021) reforçam que o impacto da automação depende do setor, do nível de escolaridade e da maturidade tecnológica das organizações. Também apontam que tecnologias como RPA, embora representem um custo inicial, trazem ganhos expressivos de eficiência quando acompanhadas de programas de qualificação. Por isso, mais do que falar em substituição direta, a literatura atual destaca a importância de políticas de transição, requalificação e formação contínua, fatores essenciais para reduzir impactos negativos e ampliar os benefícios sociais e econômicos da automação.

O artigo de Lacity e Willcocks, "What knowledge Workers Stand to Gain from Automation", ("O que os trabalhadores do conhecimento podem ganhar com a automação", em português) revela que, ao contrário das expectativas de desemprego devido à automação, as empresas realocaram funcionários para funções cognitivas, deixando tarefas repetitivas para os sistemas automatizados.

Essas tarefas repetitivas e massivas, além de serem cansativas normalmente geram falhas, devido às características antes relatadas. Assim, atribuir essas atividades para as máquinas, deixando o trabalho de análise para os trabalhadores, gerou mais eficiência e eficácia.

Entretanto, um robô não se limita a executar tarefas repetitivas por meio de scripts; ele pode também incorporar algoritmos avançados para auxiliar na tomada de decisão utilizando técnicas de Aprendizado de Máquina (Machine Learning). Ao abordar a aplicação de Machine Learning (ML), é importante considerar o uso integrado de RPA, visto que dados são essenciais para viabilizar o aprendizado de máquina e, quanto maior o volume de dados obtido, mais eficiente será o processo de extração de características. Dessa forma, a utilização de robôs para realizar essas extrações representa uma vantagem significativa para o processo. Essa evolução marca a transição do RPA tradicional para o RPA Cognitivo, que integra capacidades de IA para processar dados de forma mais inteligente (AIMDEK TECHNOLOGIES, 2018).

Há várias formas de usar robôs para automatizar processos, e essas automações têm o objetivo de apoiar trabalhadores humanos, não os substituir, e, esses robôs (RPA) evoluíram. Os RPA passaram de aplicações desktop, que exigiam dados estruturados e intervenção humana, para sistemas em nuvem capazes de processar dados variados e operar de forma autônoma com Inteligência Artificial. Há quatro versões de RPA no mercado, a mais recente incorpora Inteligência Artificial e Natural Language Process (NLP) – Processamento de Linguagem Natural (PLN) para processar dados não estruturados, escrita e fala, possibilitando a hiperautomação. Essa tendência da hiperautomação, que combina RPA e IA, foi apontada como uma das principais forças tecnológicas que moldam o futuro das organizações (Bamberger, 2021; Maddox, 2019) e já demonstra aplicações práticas em setores como o contábil (Gül et al, 2021).

A versão 1.0 dos RPA visava aumentar a produtividade humana por meio de aplicações desktop instaladas na máquina do usuário, exigindo execução assistida e promovendo automações parciais apenas com dados estruturados.

Na versão 2.0, os RPA passaram a permitir que trabalhadores digitais fossem escalados por meio do uso de um robô orquestrador responsável pelo controle da execução dos demais robôs não assistidos, mantendo, entretanto, a limitação de processamento apenas de dados estruturados.

Na versão 3.0, os RPA passaram a oferecer workflows virtuais, isto é, fluxos de trabalhos digitais, com escalabilidade e flexibilidade, executando automações diretamente na nuvem, porém permanecendo restritos ao processamento de dados estruturados.

Os RPA na versão 4.0 realizam hiperautomação usando Processamento de Linguagem Natural para identificar dados não estruturados, como textos e voz, além de manipular imagens, textos e vídeos. Podem ser geridos diretamente na nuvem, de forma assistida ou não assistida.

Hiperautomação utiliza modelos de aprendizado de máquina, redes neurais e RPA para executar tarefas repetitivas em grande escala. Na hiperautomação, a mineração de dados (Data Mining), mediante aplicação de técnicas especializadas para extração de informações, constitui a base do aprendizado de máquina (Machine Learning). Esse processo de aprendizado ocorre por meio da generalização dos dados e representa uma das primeiras etapas a ser automatizada na implementação dos processos.

Essa evolução possibilitou a transição da automação para a hiperautomação, com a integração de Inteligência Artificial, convertendo o RPA em um modelo de “RPA Cognitivo”, sendo a hiperautomação uma abordagem fundamental para a transformação digital no setor público (Pinheiro; Guimaraes, 2023), especialmente no uso da engenharia de domínio para apoiar a tomada de decisão em governos (Pinheiro; Lima Neto, 2023).

Essa integração entre RPA e Inteligência Artificial evidencia como as fronteiras da automação tradicional foram ampliadas para além da execução de tarefas repetitivas, adentrando o universo da análise inteligente e adaptativa de dados. Dessa forma, compreender os fundamentos de Artificial Intelligence (AI), bem como seus subcampos como Machine Learning (ML), Deep Learning (DL) e Large Language Models (LLM), torna-se essencial para acompanhar e potencializar o desenvolvimento das soluções de hiperautomação.

### QUER SABER MAIS?

Acesse <https://alvarofpinheiro.net/hiperautomacao> e confira vídeo aulas com fundamentações e exemplos práticos que mostram, passo a passo, como aplicá-la.

## INTELIGÊNCIA ARTIFICIAL (IA)

A ascensão da hiperautomação e do RPA Cognitivo demonstrou como a Inteligência Artificial (do inglês, Artificial Intelligence) se tornou parte fundamental das soluções tecnológicas modernas, superando limites da automação convencional ao incorporar capacidades analíticas e adaptativas. Com a crescente integração de modelos de Machine Learning, Deep Learning e Large Language Model, é indispensável compreender os princípios que orientam a evolução da IA no contexto computacional. A seguir, explora-se os fundamentos da Inteligência Artificial, seus métodos de raciocínio e as bases que sustentam a transição dos sistemas tradicionais para arquiteturas inteligentes, capazes de interpretar, aprender e agir sobre dados complexos em ambientes dinâmicos.

De acordo com Russell e Norvig (2022), a Inteligência Artificial pode ser definida como o estudo de agentes que recebem percepções do ambiente e realizam ações para atingir objetivos. A IA busca construir sistemas capazes de agir de maneira racional diante de situações complexas, utilizando informações disponíveis para maximizar suas chances de sucesso. Essa abordagem considera tanto a tomada de decisão baseada em conhecimento quanto a capacidade de aprender e adaptar-se a novos contextos.

A Inteligência Artificial é um campo da ciência que busca desenvolver sistemas capazes de interpretar informações, aprender com dados e tomar decisões de maneira autônoma. No contexto da saúde, compreender os fundamentos dessa área é essencial para identificar como diferentes tipos de algoritmos podem apoiar o diagnóstico, o tratamento e a gestão de serviços de saúde de forma mais eficiente e segura.

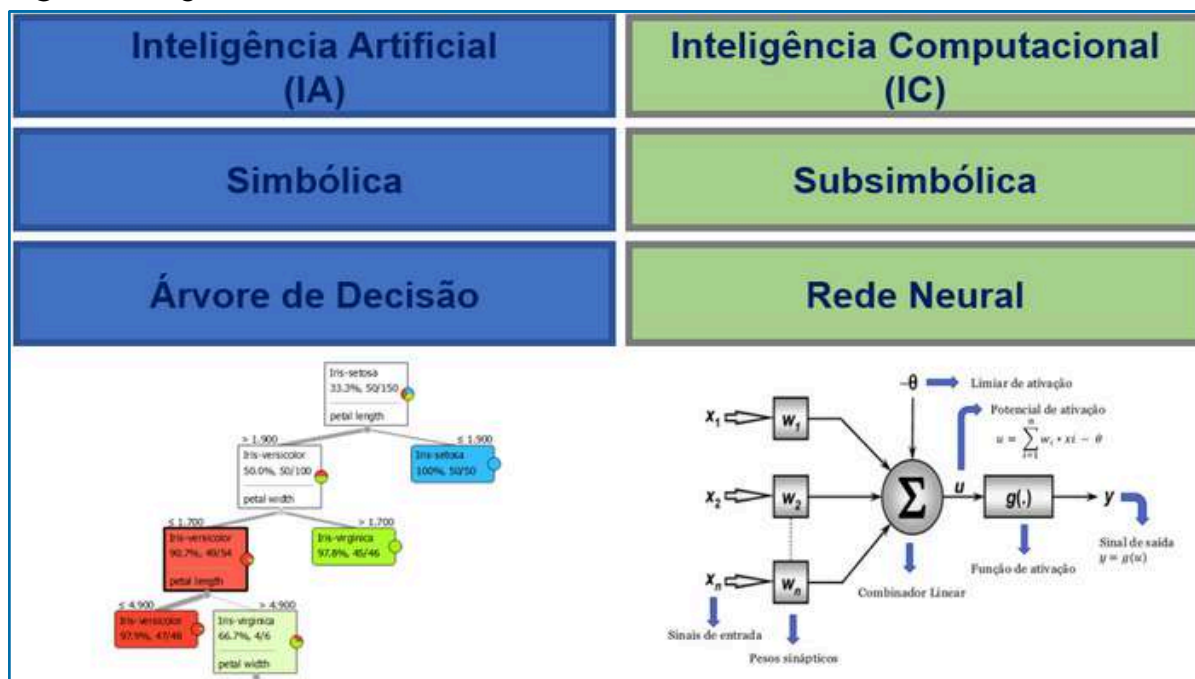
Existem duas perspectivas principais que ajudam a compreender a Inteligência Artificial: pensamento e comportamento humano, pensamento e comportamento racional. As definições que consideram o pensamento e o comportamento humano analisam até que ponto a tecnologia reproduz características cognitivas humanas, como o raciocínio e a tomada de decisão clínica. Já as abordagens focadas na racionalidade avaliam a capacidade dos sistemas de oferecer soluções mais eficazes e precisas, o que é especialmente relevante para aplicações que envolvem análise de exames, predição de surtos epidemiológicos e gestão hospitalar.

A partir dessas perspectivas, surgem quatro metodologias de estudo da Inteligência Artificial: a centrada no humano, que busca compreender o comportamento e as decisões médicas; a empírica, baseada na observação de dados clínicos e assistenciais; a racionalista, que utiliza modelos lógicos e matemáticos para apoiar diagnósticos; e a de engenharia, voltada ao desenvolvimento de sistemas inteligentes capazes de integrar informações de diferentes fontes da saúde pública.

Os algoritmos de Inteligência Artificial podem ser classificados como simbólicos e subsimbólicos. Os algoritmos simbólicos seguem regras lógicas explícitas e são mais transparentes e interpretáveis, o que os torna adequados para aplicações que exigem explicabilidade, como auditorias de prontuários e avaliação de conformidade em protocolos clínicos. Já os algoritmos subsimbólicos aprendem automaticamente a partir dos dados, reconhecendo padrões complexos em exames de imagem, sinais vitais ou registros eletrônicos, sendo amplamente utilizados em sistemas de apoio ao diagnóstico, reconhecimento de voz médica e análise de exames radiológicos.

Na prática, a combinação de abordagens simbólicas e subsimbólicas tem permitido avanços expressivos na saúde. Enquanto os modelos simbólicos ajudam na padronização de processos e no suporte à decisão administrativa, os modelos subsimbólicos alimentam soluções baseadas em aprendizado profundo, capazes de detectar precocemente doenças, analisar tendências epidemiológicas e personalizar tratamentos. Essa integração reforça o papel da Inteligência Artificial como ferramenta estratégica para fortalecer o cuidado, aumentar a eficiência operacional e aprimorar a tomada de decisão em todos os níveis do sistema de saúde.

Figura 3: Algoritmos Simbólicos e Não Simbólicos.

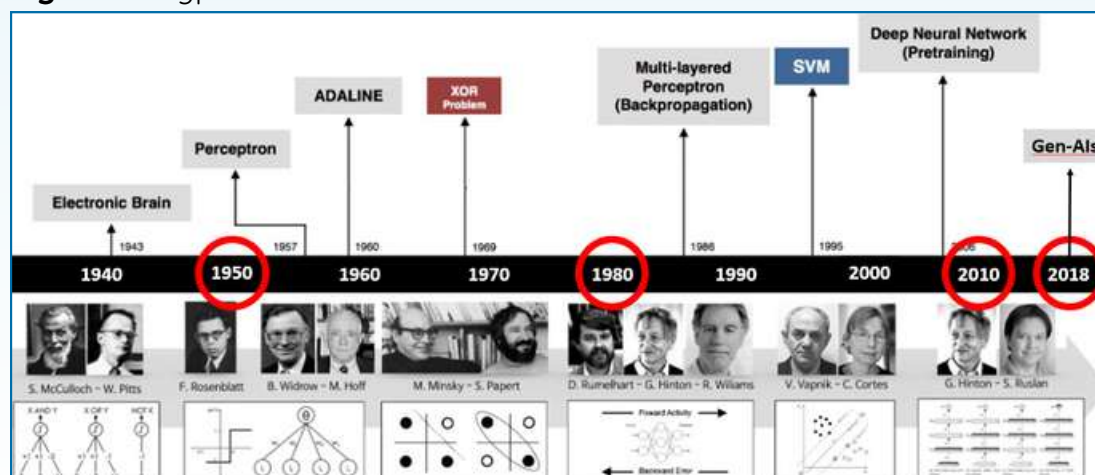


Fonte: Pinheiro, 2024.

### VOCÊ SABIA?

Ao longo das últimas décadas, a Inteligência Artificial (IA) passou por diferentes fases de desenvolvimento, cada uma responsável por avanços significativos na interação entre humanos e tecnologia. Desde sua origem nos anos 1940, período anterior ao advento da internet, até o cenário contemporâneo, a IA evoluiu de maneiras anteriormente consideradas impossíveis. Na primeira onda (1950), a ausência de internet era marcante. Na segunda onda (1980), as conexões disponíveis eram restritas. Já na terceira onda (2010), o acesso à rede permanecia concentrado em centros de pesquisa. Por fim, a quarta onda (2018) foi caracterizada pela ampliação do acesso aos algoritmos de IA, possibilitada pela OpenAI, por meio de interfaces baseadas em diálogos acessíveis ao público (Russell; Norvig, 2022).

Figura 4: Hypes da IA.

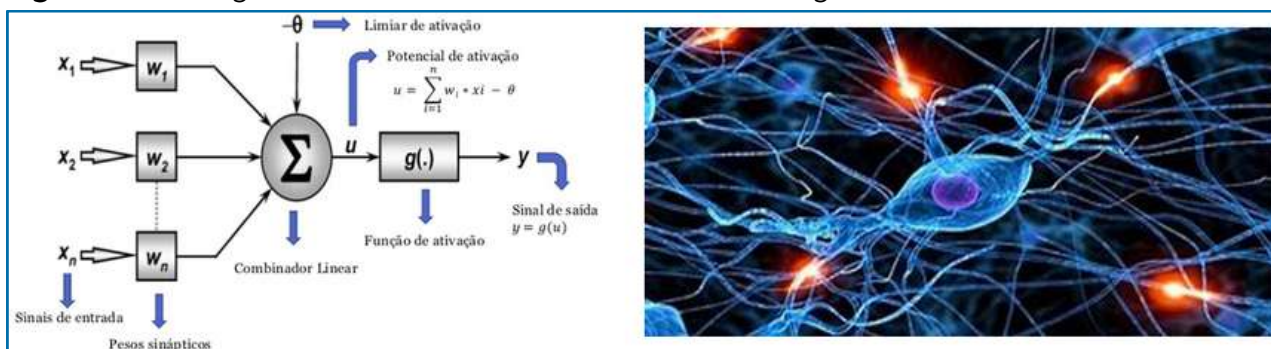


Fonte: Sefik Ilkin Serengil, 2018.

O advento do ChatGPT marcou o início da quarta onda da Inteligência Artificial, proporcionando um maior contato do público com a IA e gerando a percepção, em diversos setores, de que essa tecnologia representaria uma mudança substancial e acelerada no futuro da humanidade.

A possibilidade de interagir com IA pelo navegador transformou o cenário. Antes restrita a empresas e especialistas em plataformas como Amazon e Netflix, a IA passou a ser acessível ao público geral com a chegada do ChatGPT. A Inteligência Artificial utiliza programação e matemática para gerar conhecimento por meio de métodos não simbólicos.

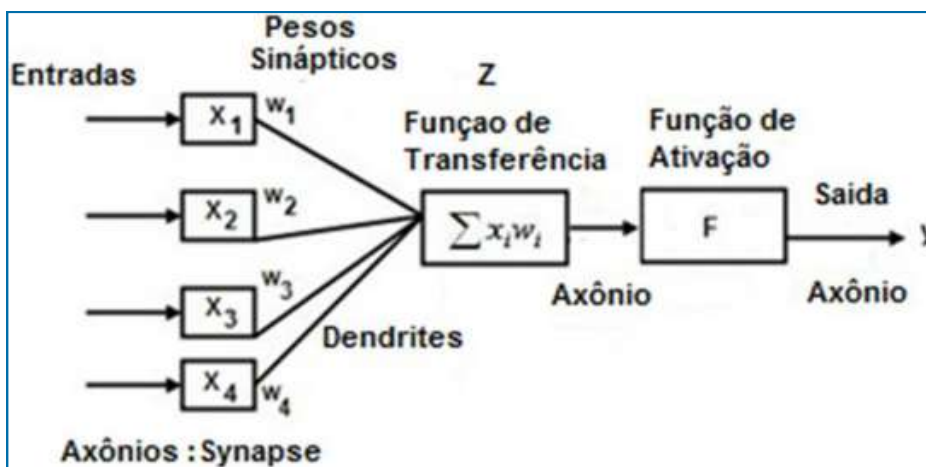
**Figura 5:** Analogia de uma rede neural artificial e biológica.



Fonte: Pinheiro, 2024.

IA funciona em máquinas, processa dados fornecidos por humanos e gera resultados por meio de algoritmos criados e controlados por pessoas.

**Figura 6:** Arquitetura de uma Rede Neural Artificial.



Fonte: Pinheiro, 2024.

As IAs não têm inteligência real; suas respostas são fruto de cálculos probabilísticos e não de compreensão emocional e sua tecnologia transformadora produz efeitos sociais e econômicos. Elas criaram oportunidades e resolveram desafios empresariais previamente considerados difíceis, com impactos para empresas e sociedade. Redes generativas (tipo de Inteligência Artificial capaz de criar conteúdo, como textos, imagens ou sons, a partir de exemplos que ela aprendeu) representam um exemplo dessas tecnologias. Abaixo segue uma pequena linha do tempo de marcos da sua evolução.

Warren McCulloch (neurocientista) e Walter Pitts (matemático), trabalharam juntos na elaboração de modelos computacionais baseados em algoritmos matemáticos conhecidos como lógica de limiar. Em 1943, eles publicaram um artigo sobre redes neurais artificiais, introduzindo o conceito do neurônio artificial, chamado neurônio de McCulloch-Pitts. Este modelo foi utilizado no desenvolvimento da IA por apresentar uma relação entre o funcionamento do neurônio e a lógica proposicional, o que facilitou sua integração com a computação digital.

A expressão Inteligência Artificial foi criada pelo cientista da computação John McCarthy durante uma conferência em Dartmouth, em 1956. Para evitar conflitos entre outros termos como "Teoria dos Autômatos" e "Cibernética", ele escolheu esse termo para ser neutro entre os termos existentes na época.

O Perceptron, criado por Frank Rosenblatt em 1957, foi fundamental para redes neurais artificiais ao introduzir o aprendizado supervisionado, base do Deep Learning. Inspirado nos neurônios biológicos, ele soma entradas e dispara sinais quando um limite é atingido, dando mais peso às conexões fortes.

Em 1959, Artur Samuel publicou um artigo pioneiro sobre aprendizado de máquina em damas, sendo o primeiro a usar um jogo para testar Inteligência Artificial. Seu método, baseado em árvores de decisão, permitia ao algoritmo avaliar possíveis jogadas e aprender autonomamente.

Em 1960, Bernard Widrow e Marcian Hoff, da Universidade de Stanford, criaram o algoritmo ADALINE, uma rede neural que aprende por repetição e se baseia nos conceitos dos neurônios McCulloch-Pitts.

Em 1986, David Rumelhart, Geoffrey Hinton e Ronald Williams publicaram o artigo seminal sobre o algoritmo de retropropagação (Backpropagation), revolucionando o campo das redes neurais artificiais e da IA. Esse método permitiu o treinamento eficiente de redes neurais multicamadas, tornando possível ajustar os pesos de milhares de conexões por meio do cálculo do gradiente, o que aumentou significativamente a capacidade de aprendizado das máquinas.

Em 2010, Geoffrey Hinton e Ruslan Salakhutdinov realizaram avanços significativos no campo do Deep Learning ao propor técnicas inovadoras para o treinamento de redes neurais profundas. Eles demonstraram que era possível superar limitações de redes profundas utilizando métodos de pré-treinamento não supervisionado, o que permitiu a extração de representações mais eficientes dos dados e impulsionou o desenvolvimento de aplicações modernas de Inteligência Artificial, como reconhecimento de voz e imagem.

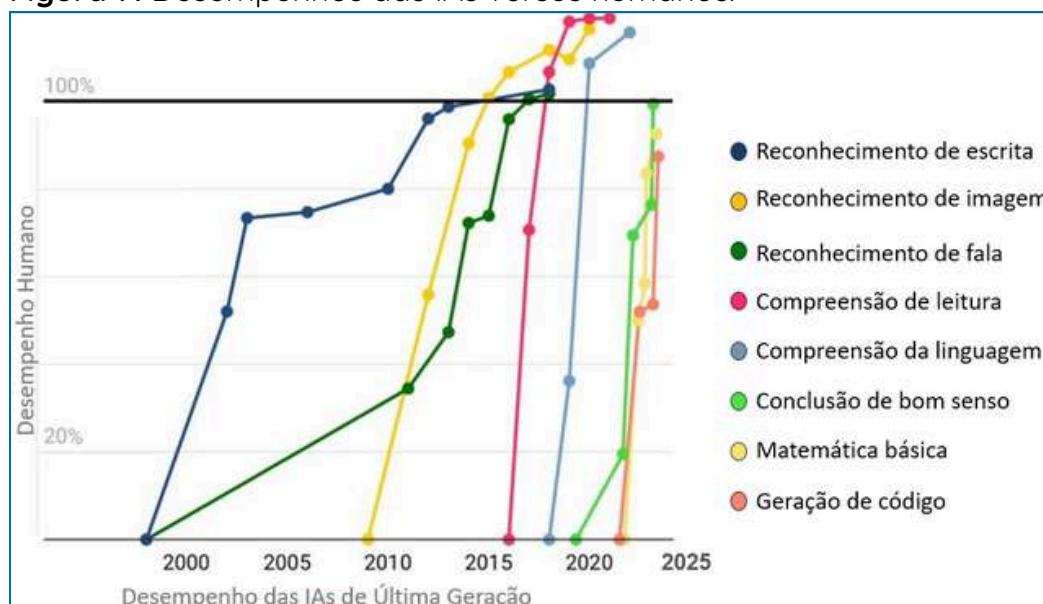
O algoritmo Transformer desempenha um papel relevante na IA generativa e contribuiu para o avanço do ChatGPT. Em 2017, os pesquisadores Ashish Vaswani e Jakob Uszkoreit buscavam aprimorar o Google Translate, enquanto Illia Polusukhin trabalhava no conceito de self-attention, que alterou a velocidade e o funcionamento dos algoritmos de processamento de linguagem natural.

Anteriormente, o Google Translate e outros sistemas de Processamento de Linguagem Natural realizavam a tradução palavra por palavra, seguindo uma ordem sequencial. Com a introdução do conceito de self-attention, que permite a análise não-linear dos elementos da frase, ocorreu um avanço relevante. Esse algoritmo processa toda a sentença simultaneamente, avaliando suas partes em conjunto ao invés de tratar cada palavra isoladamente, o que possibilita captar o contexto integral e realizar traduções paralelas.

Posteriormente, Vaswani juntamente com outros pesquisadores publicaram o artigo "Attention is All You Need" (2017), considerado uma referência para sistemas generativos. O estudo evidencia a distinção entre máquinas e seres humanos: enquanto a criatividade algorítmica está presente no desenvolvimento dos sistemas, a inteligência essencial reside na capacidade imaginativa dos programadores envolvidos.

Dessa evolução, (1940 - 1980) IA Fraca, (1980 - 2010) IA Forte, (2010 - 2017) IA de Aprendizado Profundo, (2018 - Aos Dias Atuais), temos a IA generativa, que funciona como um mecanismo probabilístico de preenchimento automático, desenvolvido a partir de grandes volumes de dados. Por meio desse processo matemático, esses sistemas são capazes de replicar funções cognitivas humanas baseadas em algoritmos, incluindo aprendizagem e resolução de problemas. Tornando o desempenho das IAs cada vez melhor, como pode ser observado a Figura 7.

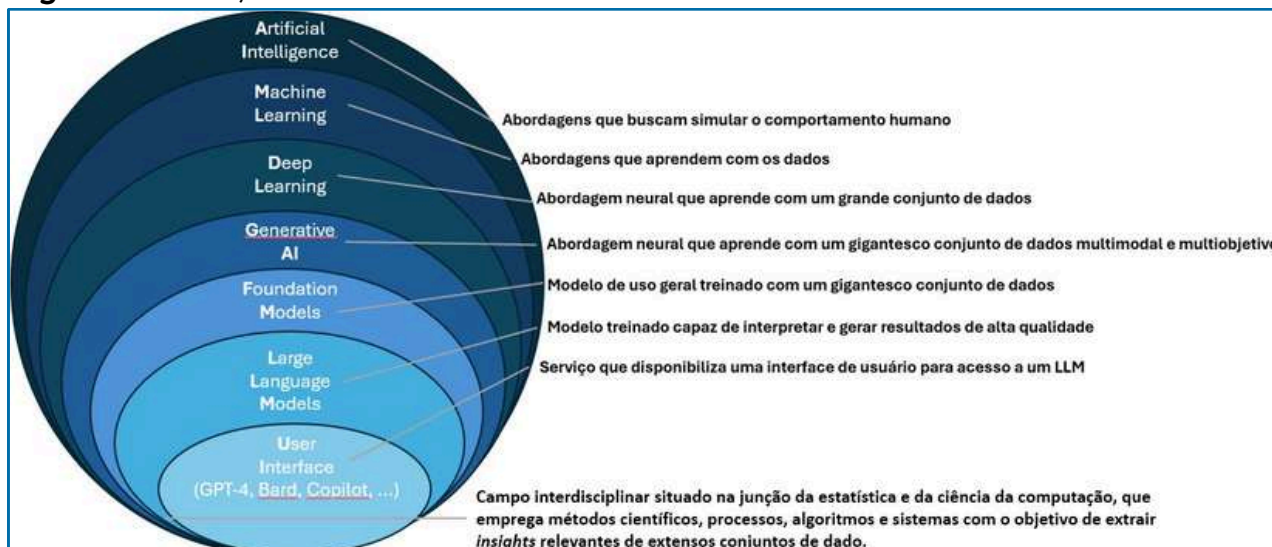
**Figura 7:** Desempenhos das IAs versus humanos.



Fonte: OpenAI, 2025.

Os modelos de IA Generativa utilizam Inteligência Artificial para gerar conteúdo original (dados sintéticos) em resposta a comandos. A IA Generativa é um tipo de Inteligência Artificial que cria textos, imagens, vídeos, áudios ou códigos a partir de comandos do usuário com Large Language Model (LLM) – Grande Modelo de Linguagem, e com Multiagentes (Agentes Autônomos Inteligentes) que colaboram em um ambiente comum, e pode contribuir com melhor solução de problemas, fazendo que cada agente possa agir conforme suas próprias regras e objetivos, afetando o sistema como um todo, sendo a evolução das LLMs para as Large Action Model (LAM).

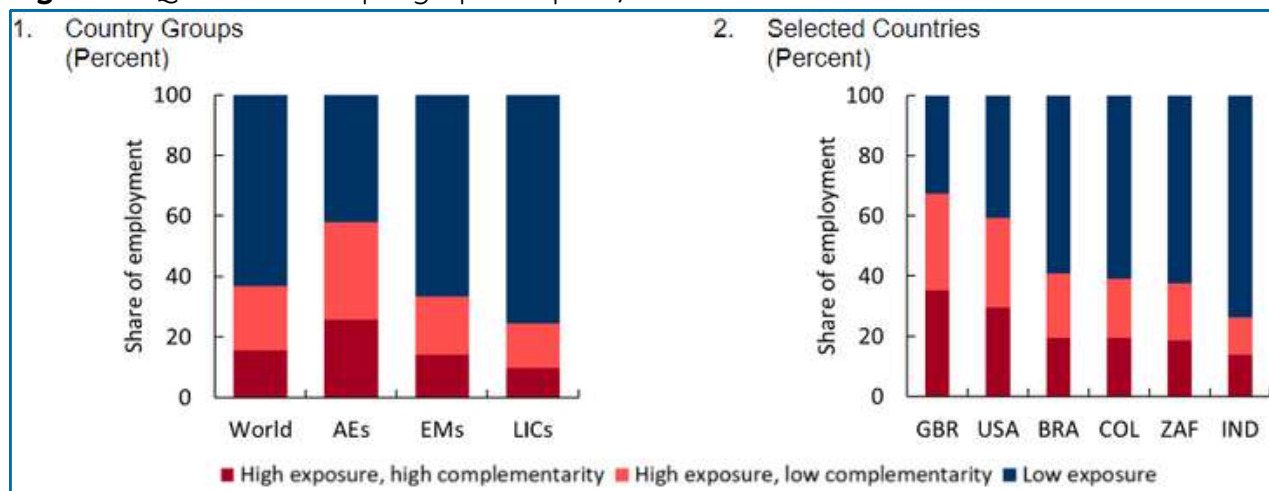
Figura 8: Evolução das Redes Neurais.



Fonte: Pinheiro, 2024.

Com a evolução da IA, vários estudos (FMI, 2024) mostram que a IA impactará 40% das profissões globais, afetando várias funções. Assim, é essencial criar políticas para aproveitar a IA e atender às demandas do mercado de trabalho.

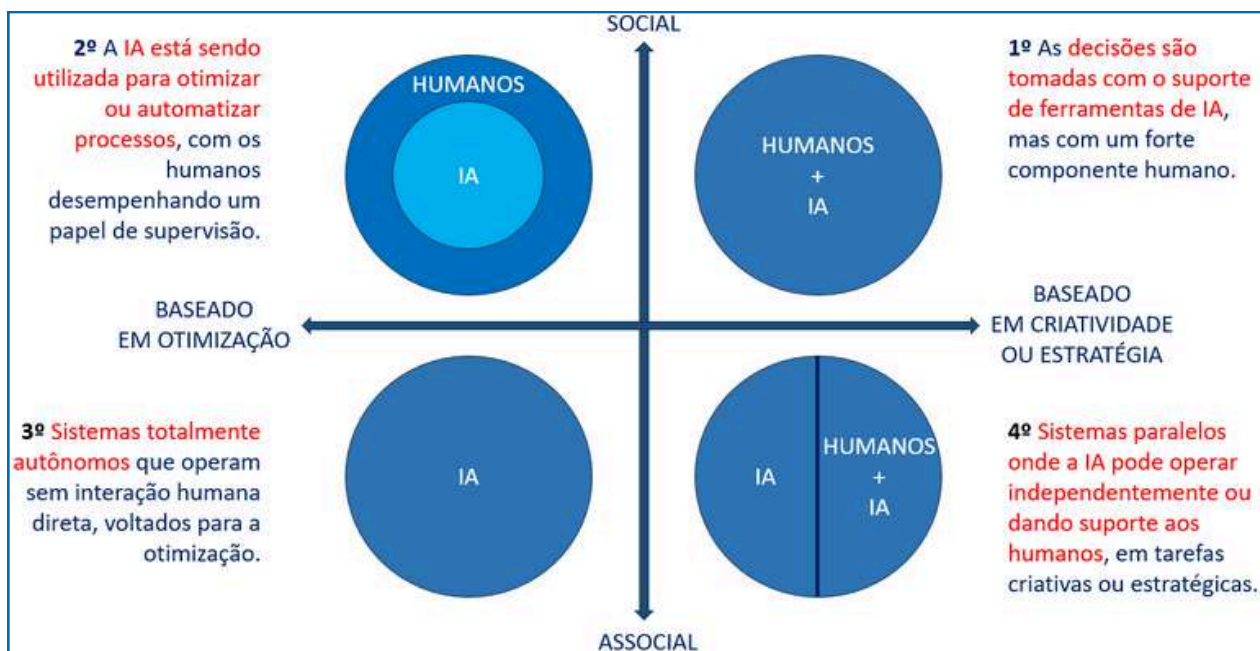
Figura 9: Quotas de emprego por exposição da IA.



Fonte: FMI, 2024.

A integração da IA aumenta a produtividade e os salários dos trabalhadores, enquanto a falta de adaptação resulta em estagnação (FMI, 2024). A IA já impacta diversas profissões globalmente. É fundamental que os países promovam a requalificação profissional para garantir uma transição justa, proteger empregos e reduzir desigualdades sociais. Visto que, o uso das IAs estar em todos os segmentos da vida cotidiana, como se pode observar na Figura 10.

Figura 10: Implementar IA alinhada aos objetivos.



Fonte: Kai-Fu Lee, 2019.

A Inteligência Artificial é uma área inovadora da ciência voltada para criar sistemas que realizam tarefas normalmente feitas por humanos, como entender linguagem e simular raciocínio, transformando diversos setores.

Dentre os principais subcampos da IA, destacam-se os modelos de linguagem generativa, desenvolvidos para interpretar, produzir e interagir com a linguagem humana de forma avançada. Esses sistemas são classificados em distintas famílias conforme suas arquiteturas e organizações responsáveis pelo desenvolvimento.

Dentro da Inteligência Artificial (IA), destacam-se subcampos como aprendizado de máquina, responsável por criar algoritmos que permitem às máquinas aprenderem com experiência, usando aprendizado não supervisionado, supervisionado, semisupervisionado e por reforço. Outros subcampos da IA, são:

- **A visão computacional:** Busca capacitar máquinas a interpretar o mundo visual, sendo aplicada em áreas como direção autônoma, vigilância e diagnóstico médico.
- **A robótica:** Desenvolve máquinas inteligentes capazes de executar tarefas físicas em setores diversos, desde a indústria, missões espaciais, como em cirurgias robóticas.
- **A IA explicável:** Visa criar transparência nos processos de decisão das máquinas, essencial para áreas como saúde, finanças e justiça.
- **Campo da interação humano-computador:** A IA foca em interfaces intuitivas, como assistentes virtuais e sistemas de reconhecimento de fala.

## QUER SABER MAIS?

Acesse <https://alvarofpinheiro.net/dados> e confira vídeo aulas com fundamentações e exemplos práticos que mostram, passo a passo, como aplicá-la.

## MACHINE LEARNING (ML)

Segundo Russell e Norvig (2022), aprendizado de máquina é um campo da Inteligência Artificial que desenvolve algoritmos capazes de aprender com dados e melhorar em tarefas ao longo do tempo, sem programação explícita. Esses sistemas extraem conhecimento de experiências, reconhecem padrões e ajustam seu comportamento para obter melhores resultados.

Conforme os autores, o objetivo do Aprendizado de Máquina é capacitar máquinas para executar tarefas complexas em cenários incertos, adaptando-se, generalizando por exemplos e tomando decisões a partir dos dados disponíveis, automatizando tarefas complexas, identificando padrões em grandes volumes de dados, adaptando sistemas a novos contextos, personalizando processos para diferentes usuários e permitindo que agentes inteligentes evoluam com a experiência.

O aprendizado de máquina permite que sistemas inteligentes evoluam e lidem com situações reais sem depender apenas de instruções explícitas, e, apresenta tipos distintos, os quais são: não supervisionado, supervisionado, semisupervisionado e por reforço, cada um com métodos, vantagens, desafios e aplicações específicos detalhados a seguir.

O aprendizado não supervisionado identifica padrões ou grupos em dados sem rótulos, usando técnicas como clusterização (agrupamento), associação de itens e redução de dimensionalidade, tudo sem intervenção humana.

O aprendizado de máquina supervisionado treina algoritmos com dados rotulados, permitindo que modelos generalizem resultados para novos casos. A escolha do modelo depende do problema e das características dos dados. Técnicas como classificação categorizam informações, enquanto predição estima comportamentos futuros com base em registros históricos.

O aprendizado semisupervisionado integra dados rotulados e não rotulados com o objetivo de aprimorar a performance dos modelos, posicionando-se entre as abordagens supervisionadas e não supervisionadas. Trata-se de uma solução eficaz quando o processo de rotulação é dispendioso. A combinação desses dois tipos de dados permite que os modelos semisupervisionado apresentem resultados superiores em relação às técnicas que utilizam apenas um tipo de dado.

O Aprendizado por Reforço é uma abordagem de Inteligência Artificial na qual sistemas automatizados aprendem a tomar decisões com base em tentativa e erro, visando otimizar recompensas estabelecidas por uma política. O modelo passa de ações iniciais aleatórias para estratégias mais sofisticadas conforme recebe feedback, sendo aplicado em áreas como robótica, jogos, sistemas autônomos e controle de processos.

A Inteligência Artificial traz soluções inovadoras a vários setores, aprendendo e se adaptando ao ambiente. Técnicas formais têm mostrado eficiência em robótica, controle de processos e sistemas autônomos, apontando grande potencial futuro.

Dessa forma, tanto o aprendizado de máquina quanto o aprendizado profundo (Deep Learning) representam marcos fundamentais no avanço da Inteligência Artificial, oferecendo estratégias distintas e complementares para que máquinas possam aprender, adaptar-se e resolver problemas complexos nos mais diversos domínios. Enquanto o Machine Learning destaca-se por abranger uma vasta gama de técnicas e abordagens, o Deep Learning potencializa a aprendizagem hierárquica e automatizada, revolucionando aplicações em áreas como visão computacional, processamento de linguagem e criatividade artificial (Pinheiro; Santos, 2020; Pinheiro; Lima Neto, 2022).

## APRENDIZADO PROFUNDO (DL)

O Aprendizado profundo (do inglês, Deep Learning), conforme discutido por Russell e Norvig (2022) em *Artificial Intelligence: A Modern Approach* (“Inteligência Artificial: uma abordagem moderna”, em português), constitui uma das vertentes mais marcantes da Inteligência Artificial contemporânea. Trata-se de um ramo do aprendizado de máquina baseado em redes neurais artificiais com múltiplas camadas, capazes de aprender representações hierárquicas a partir de grandes volumes de dados. A principal contribuição dessa abordagem está em permitir que sistemas construam, camada após camada, abstrações que vão desde características elementares até conceitos de alto nível.

No contexto mais amplo da Inteligência Artificial, o aprendizado profundo aparece como um desdobramento das redes neurais tradicionais (aprendizado de máquina). Enquanto as primeiras arquiteturas, com poucas camadas, apresentavam severas restrições em sua capacidade de generalização, as redes profundas superaram essa barreira ao explorar representações distribuídas e não lineares. Assim, máquinas passaram a organizar informações de maneira mais próxima ao funcionamento do cérebro humano, aprendendo desde padrões básicos até estruturas altamente complexas.

No plano das aplicações, o Deep Learning tornou-se o motor da chamada “Inteligência Artificial moderna”. Ele está presente em sistemas de reconhecimento de fala, assistentes virtuais, tradutores automáticos, plataformas de recomendação e algoritmos criativos que geram imagens, músicas e narrativas textuais. Sua utilização atravessa setores tão distintos quanto saúde, finanças, educação, entretenimento e logística, evidenciando seu caráter transversal e transformador.

Apesar dos êxitos, os desafios permanecem significativos. Redes neurais profundas são frequentemente criticadas por sua opacidade, funcionando como verdadeiras “caixas-pretas” de difícil interpretação. Sua dependência de grandes volumes de dados e recursos computacionais cria barreiras para organizações com infraestrutura limitada. Além disso, a capacidade de generalização fora do domínio de treinamento ainda é limitada, o que coloca em questão sua robustez. A isso somam-se preocupações éticas relacionadas a vieses algorítmicos, reprodução de desigualdades estruturais e riscos associados à tomada de decisões automatizadas.

Assim, para Russell e Norvig (2022), o Deep Learning deve ser compreendido como mais do que uma técnica avançada de aprendizado. Ele simboliza um marco histórico no desenvolvimento da Inteligência Artificial, pois consolidou a ideia de aprendizado hierárquico em múltiplas camadas e abriu caminho para inovações sem precedentes. Ao mesmo tempo, trouxe novos dilemas científicos, sociais e éticos que exigem atenção constante da comunidade acadêmica e profissional.

Em suma, o Deep Learning consolidou-se como um divisor de águas na trajetória da Inteligência Artificial, abrindo possibilidades inéditas para o processamento de informações em múltiplos níveis de abstração. Ao potencializar a autonomia dos sistemas na análise de dados complexos e no reconhecimento de padrões, transformou setores inteiros e promoveu avanços que repercutem desde o cotidiano até a fronteira da pesquisa científica. No entanto, as limitações quanto à interpretabilidade, demanda de recursos e questões éticas reforçam a necessidade de reflexão crítica e aprimoramento contínuo. O legado do Deep Learning reside não apenas em sua eficácia técnica, mas também nos desafios que lançou à comunidade acadêmica, industrial e à sociedade como um todo.

Neste cenário de inovação acelerada, um dos campos que mais se destaca na Inteligência Artificial contemporânea é o das Large Language Models (LLMs). Estas arquiteturas surgiram da evolução do processamento de linguagem natural, unindo avanços em redes neurais profundas a novas estratégias para compreensão, geração e interação textual. As LLMs representam um salto qualitativo ao permitir que máquinas dialoguem, interpretem e produzam linguagem humana com fluência e contexto, estendendo ainda mais as fronteiras da IA moderna. A seguir, exploraremos os princípios fundamentais das LLMs, seus marcos históricos e os desafios éticos e sociais associados a essa tecnologia transformadora.

## LARGE LANGUAGE MODEL (LLM)

Russell e Norvig (2022) definem a IA como o estudo de agentes capazes de perceber e agir racionalmente em ambientes complexos. A linguagem natural, nesse contexto, sempre foi um dos maiores desafios, pois representa o canal mais sofisticado de comunicação humana. Os primeiros esforços computacionais voltados para o processamento de linguagem natural baseavam-se em regras explícitas e em modelos estatísticos como os n-gramas, que estimavam a probabilidade condicional de palavras em sequências curtas. Ainda que úteis, esses sistemas eram incapazes de lidar com dependências linguísticas de longo alcance, o que limitava sua aplicação prática.

Entre as LLMs mais conhecidas e utilizadas atualmente estão o ChatGPT, Gemini e Claude, que se destacam por sua capacidade de compreender e gerar linguagem natural de maneira sofisticada. No contexto da saúde, essas ferramentas podem ser ajustadas para interpretar prontuários, auxiliar no atendimento ao paciente, automatizar a triagem de sintomas e aprimorar a comunicação entre profissionais e usuários do sistema. A personalização desses modelos, com bases de dados específicas do setor e rigorosos controles de privacidade, potencializa diagnósticos mais precisos e um suporte administrativo ágil, mantendo a segurança e a ética como prioridades.

Apesar de seus avanços, as LLMs levantam desafios críticos. Entre eles, destacam-se os vieses herdados dos dados de treinamento, a falta de transparência nos processos de inferência e o consumo energético elevado. Russell e Norvig (2022) enfatizam que a Inteligência Artificial deve ser concebida não apenas como racionalidade instrumental, mas como racionalidade prática, que leve em consideração valores humanos, responsabilidade ética e impactos sociais.

As Large Language Models representam a síntese de décadas de pesquisa em Inteligência Artificial, unindo estatística, aprendizado profundo e capacidade computacional. Ao mesmo tempo, reforçam o princípio de Russell e Norvig de que sistemas inteligentes devem atuar não apenas de maneira eficaz, mas de forma racional e alinhada aos objetivos humanos. Os marcos históricos, do GPT-2 ao GPT-5, evidenciam um avanço inegável em direção a agentes linguísticos generalistas, embora os dilemas éticos e sociais ainda representem um campo em aberto.

Em síntese, compreender o percurso histórico, os fundamentos teóricos e os desafios contemporâneos do Deep Learning e das Large Language Models é fundamental para qualquer pessoa que deseje atuar de forma crítica e inovadora no campo da Inteligência Artificial. A evolução das técnicas, aliada à reflexão ética e social, constitui a base para o desenvolvimento de sistemas cada vez mais alinhados aos interesses humanos e ao bem-estar coletivo.

## AtualizAI

Como podemos utilizar a hiperautomação no campo da saúde?

1. *Melhoria na eficiência operacional: Elimina tarefas manuais repetitivas e reduz filas de atendimento; Acelera processos administrativos como agendamentos, triagens e autorizações.*
2. *Redução de erros e retrabalho: Com robôs e IA atuando juntos, os dados são tratados com mais precisão; Informações médicas, prontuários e receitas são atualizados automaticamente, reduzindo falhas humanas.*
3. *Tomada de decisão estratégica: Algoritmos analisam grandes volumes de dados de saúde pública (Big Data) para apoiar políticas de prevenção, alocação de recursos e resposta a eventos em saúde.*
4. *Atendimento ao cidadão mais ágil e humanizado: Chatbots inteligentes orientam pacientes 24h, tiram dúvidas e até fazem triagens iniciais, e diminuindo filas de espera; e, Automatizações liberam os profissionais da burocracia para focarem no cuidado direto.*

*Imagine um cenário onde um distrito sanitário que precise organizar uma campanha de vacinação contra a gripe voltada para grupos prioritários. Os gestores já possuem os dados demográficos, cobertura vacinal do ano anterior e grupos de risco.*

Na IA, os LLMS podem contribuir para:

- *Planejar ações de mobilização de acordo com o território a partir dos dados fornecidos;*
- *Apontar regiões com baixa e alta cobertura;*
- *Estimar a quantidade de vacinas necessárias;*
- *Produzir materiais de apoio (panfletos, banners, informativos).*

Com esse panorama, encerramos o Módulo 1 e avançamos para o Módulo 2: Classes de Problemas e Tipos de Aprendizado. No próximo módulo, exploraremos como diferentes algoritmos são aplicados para resolver problemas variados, abordando desde métodos estatísticos até as vertentes evolutivas e conexionistas, além dos principais paradigmas de aprendizado supervisionado, não supervisionado e por reforço.

### QUER SABER MAIS?

Acesse <https://alvarofpinheiro.net/ml> e confira vídeo aulas com fundamentações e exemplos práticos que mostram, passo a passo, como aplicá-la.

## AULA VIRTUAL



Módulo 1 - Aula 1 - A transformação digital na Saúde



# MÓDULO 2

## CLASSES DE PROBLEMAS E TIPOS DE APRENDIZADO



## APRESENTAÇÃO

No campo da Inteligência Artificial, Russell e Norvig (2022) descrevem o aprendizado de máquina como um conjunto de métodos capazes de permitir que um agente aprenda a partir de dados ou experiências acumuladas, ajustando-se a contextos diversos de previsão, descoberta de padrões e tomada de decisão. Esses métodos podem ser diferenciados segundo o grau de supervisão disponível nos dados e a forma como a informação é explorada, o que dá origem a diferentes classes de problemas. Cada uma dessas classes possui fundamentos próprios e aplicações distintas, mas todas compartilham o mesmo objetivo de melhorar o desempenho do agente ao longo do tempo.

Os diferentes tipos de aprendizado de máquina se distinguem pela forma como os sistemas aprendem com os dados. No aprendizado não supervisionado, o agente identifica padrões e estruturas ocultas em conjuntos sem rótulos, aplicando técnicas como redução de dimensionalidade, agrupamento e descoberta de associações. No supervisionado, o modelo é treinado com exemplos rotulados para aprender relações entre entradas e saídas, permitindo classificação e regressão em tarefas como diagnóstico médico ou previsão de demanda. O semisupervisionado combina poucos dados rotulados com muitos não rotulados, propagando o conhecimento aprendido e reduzindo custos de rotulação. Já o aprendizado por reforço baseia-se na interação com o ambiente, ajustando ações a partir de recompensas e penalidades para maximizar resultados ao longo do tempo, sendo aplicado em áreas como robótica, controle adaptativo e sistemas de recomendação inteligentes.

Ao articular essas quatro classes de problemas não supervisionado, supervisionado, semisupervisionado e por reforço Russell e Norvig oferecem uma visão abrangente da aprendizagem de máquina como disciplina central da Inteligência Artificial. Essa classificação não apenas organiza as diferentes técnicas disponíveis, mas também fornece um quadro conceitual robusto para compreender como agentes inteligentes podem lidar com dados, experiência e interação, produzindo soluções que vão da identificação de padrões latentes até a tomada de decisão ótima em ambientes complexos.

## CLASSES DE PROBLEMAS DE APRENDIZADO DE MÁQUINA NA SAÚDE

O aprendizado de máquina constitui uma das áreas centrais da Inteligência Artificial e, segundo Russell e Norvig (2022), pode ser organizado em diferentes classes de problemas de acordo com o nível de supervisão dos dados e com os objetivos do aprendizado. Cada classe se aplica a contextos específicos, e a área da saúde oferece exemplos particularmente ilustrativos, dada a riqueza e a complexidade de seus dados, além do impacto direto sobre a vida humana.

Figura 11: Tipos de Aprendizado de Máquina.



Fonte: Pinheiro, 2025.

### Aprendizado Não Supervisionado

No aprendizado não supervisionado, os algoritmos não recebem rótulos prévios e precisam identificar estruturas ocultas nos dados. Um exemplo central é a redução de dimensionalidade em dados genômicos, nos quais milhares de genes são avaliados por paciente. Métodos como a análise de componentes principais (PCA) permitem comprimir essa informação, destacando padrões associados a mutações ligadas a cânceres específicos (Alizadeh et al., 2000).

O agrupamento também é amplamente utilizado em saúde, como na identificação de perfis de pacientes em unidades de terapia intensiva (UTI). Algoritmos de clustering, como k-means ou DBSCAN, podem agrupar pacientes com perfis clínicos semelhantes, auxiliando no ajuste de protocolos de tratamento (Zheng et al., 2017).

Já a associação se destaca em análises de saúde populacional. Regras do tipo “pacientes com hipertensão e obesidade possuem maior probabilidade de desenvolver diabetes tipo 2” podem ser extraídas de grandes bancos de dados hospitalares, apoiando estratégias de prevenção (Agrawal; Srikant, 1994).

### Aprendizado Supervisionado

O aprendizado supervisionado utiliza dados rotulados, nos quais entradas estão associadas a saídas conhecidas. Na classificação, modelos de Redes Neurais Convolucionais analisam imagens médicas para distinguir entre tumores benignos e malignos, como no caso do câncer de mama (Litjens et al., 2017).

#### QUER SABER MAIS?

Acesse [aqui](#) um artigo de revisão bibliográfica que explora as diversas aplicações da IA em imagens médicas, incluindo radiologia, patologia, cardiologia e outras áreas, na identificação de tumores.

A regressão é aplicada, por exemplo, na previsão de valores contínuos como o tempo de permanência hospitalar. Modelos lineares e não lineares podem prever quantos dias um paciente deve permanecer em uma UTI, otimizando a gestão de leitos (Rajkomar; Dean; Kohane, 2019).

Na predição, a saúde pública se beneficia de modelos capazes de antecipar surtos de doenças. Um caso relevante é a previsão de epidemias de dengue, baseada na combinação de dados climáticos, registros epidemiológicos e informações ambientais (Luz et al., 2008).

### Aprendizado Semisupervisionado

No aprendizado Semisupervisionado, apenas parte dos dados possui rótulos, o que reflete a realidade da saúde, onde a anotação manual por especialistas é cara e demorada. Em imagens médicas, como tomografias, poucos exames contam com diagnósticos validados, mas milhares estão armazenados sem rotulação. Algoritmos semisupervisionado utilizam uma amostra anotada para guiar a categorização do restante (Cheplygina et al., 2019).

Outro exemplo é a análise de prontuários eletrônicos, que frequentemente contêm relatórios clínicos em texto livre. Técnicas de propagação de rótulos em grafos podem inferir diagnósticos prováveis, permitindo padronizar informações e apoiar sistemas de decisão clínica.

### Aprendizado por Reforço

O aprendizado por reforço baseia-se na interação do agente com o ambiente, recebendo recompensas ou penalidades. Na busca, modelos são utilizados em simulações para encontrar protocolos de tratamento mais eficazes. Por exemplo, algoritmos testam combinações de drogas em pacientes virtuais até identificar políticas terapêuticas que maximizem a recuperação (Yu et al., 2019).

No campo da otimização, hospitais utilizam aprendizado por reforço para gerir recursos críticos, como a alocação de leitos e respiradores, equilibrando demanda e disponibilidade (Komorowski et al., 2018).

Já os sistemas de recomendação têm sido aplicados em terapias personalizadas. A radioterapia adaptativa é um exemplo: o modelo ajusta doses ao longo do tratamento de câncer com base nas respostas fisiológicas do paciente, buscando maximizar eficácia e minimizar efeitos colaterais (Tseng et al., 2017).

As **classes de aprendizado** de máquina descritas por Russell e Norvig encontram aplicações diretas e profundas na saúde. Enquanto o aprendizado não supervisionado revela estruturas ocultas, o supervisionado fortalece diagnósticos e previsões; o semisupervisionado equilibra custos de anotação e escala; e o aprendizado por reforço expande a personalização de tratamentos e a otimização de recursos. O impacto dessas abordagens converge para um futuro de medicina mais precisa, preventiva e personalizada, sustentada pela Inteligência Artificial.

Dessa forma, os diferentes paradigmas de aprendizado de máquina apresentam uma complementaridade essencial no cenário da saúde. Cada abordagem, seja não supervisionada, supervisionada, semisupervisionada ou por reforço, constitui um alicerce metodológico capaz de responder a desafios específicos, potencializando diagnósticos, prognósticos, gestão hospitalar e personalização terapêutica. O avanço dessas técnicas, aliado ao crescimento exponencial de dados biomédicos, aponta para uma transformação contínua e progressiva do cuidado em saúde, ancorada na precisão, eficiência e adaptabilidade proporcionadas pela IA.

Para que esses avanços se concretizem, é fundamental compreender as diferentes abordagens de desenvolvimento da Inteligência Artificial, cada uma oferecendo ferramentas, modelos e estratégias para resolver problemas complexos. Dentre as principais abordagens, destacam-se: a estatística, que fundamenta algoritmos probabilísticos; as árvores de decisão, que favorecem interpretações estruturadas; os métodos evolucionários e de enxame, inspirados em fenômenos naturais de adaptação coletiva; e as redes neurais, que simulam o funcionamento do cérebro humano. A seguir, serão exploradas as características centrais, aplicações e exemplos de cada uma dessas abordagens, evidenciando como estruturam o raciocínio computacional para a resolução de problemas em saúde.

## ALGORITMOS DA ABORDAGEM ESTATÍSTICA

Segundo Russell e Norvig no clássico "Artificial Intelligence: A Modern Approach" (4ª ed., 2022), a chamada abordagem estatística dentro da Inteligência Artificial emergiu como uma resposta à necessidade de lidar com incertezas, variabilidade nos dados e problemas de aprendizado em larga escala. Essa perspectiva busca modelar relações probabilísticas entre variáveis, empregando estatística e probabilidade como núcleo do raciocínio e da tomada de decisão.

Os autores defendem que a estatística se tornou o núcleo dominante da IA moderna, em especial após os anos 1990, porque fornece rigor matemático e capacidade de generalização em ambientes incertos (Russell; Norvig, 2022). Essa abordagem abrange desde algoritmos simples (Naive Bayes) até complexos modelos gráficos probabilísticos, constituindo a ponte entre IA simbólica e aprendizado de máquina baseado em dados.

Russell e Norvig enfatizam que os algoritmos estatísticos derivam de princípios de inferência probabilística. Os exemplos mais clássicos incluem:

- **Naive Bayes Classifier:** Baseado no Teorema de Bayes e na suposição de independência condicional entre atributos. É utilizado em classificação de textos, diagnósticos médicos e análise de spam. Exemplo em saúde: classificação de exames laboratoriais em "normal" ou "anômalo" com base em múltiplos biomarcadores.
- **Gaussian Mixture Model (GMM):** Algoritmos de agrupamento que representam dados como combinação de distribuições gaussianas. Associados ao algoritmo EM (Expectation-Maximization). Exemplo em saúde: identificação de subgrupos de pacientes com padrões similares em imagens de ressonância magnética.
- **Bayesian Networks:** Estruturas gráficas direcionadas que representam variáveis aleatórias e suas dependências condicionais. Exemplo em saúde: modelagem de riscos em UTI, como a probabilidade de falência renal dado histórico de hipertensão e diabetes.
- **Hidden Markov Model (HMM):** Usados para modelar sequências temporais com dependência probabilística entre estados. Exemplo em saúde: monitoramento de sinais vitais contínuos de pacientes, onde o "estado oculto" pode ser o estágio da doença.
- **Regressão Logística:** Embora simples, é um dos pilares da abordagem estatística para problemas de classificação binária. Exemplo em saúde: prever risco de infarto com base em idade, colesterol e pressão arterial.
- **Support Vector Machine (SVM):** Apesar de sua base geométrica, Russell e Norvig a colocam dentro da abordagem estatística, pois estima margens e distribuições de decisão. Exemplo em saúde: detecção automática de câncer em imagens de biópsia.

## ALGORITMOS DA ABORDAGEM ÁRVORE

A abordagem em árvore é apresentada como uma das mais clássicas e influentes dentro da IA e do aprendizado de máquina. Essa família de algoritmos tem como fundamento a construção de modelos hierárquicos de decisão, representados por estruturas em forma de árvore, em que cada nó interno corresponde a um teste sobre uma variável, cada ramo a um resultado desse teste e cada folha a uma decisão ou classificação final.

Russell e Norvig destacam que árvores de decisão são métodos simples, interpretáveis e eficientes, bastante utilizados em tarefas de classificação e regressão. Sua força está no fato de que: podem lidar com dados qualitativos e quantitativos; realizam partições recursivas do espaço de atributos; são de fácil interpretação para especialistas humanos (explicabilidade). As árvores de decisão utilizam medidas estatísticas como entropia, ganho de informação e índice de Gini para selecionar o atributo mais relevante em cada divisão. Principais Algoritmos de Árvore, são:

- **ID3 (Iterative Dichotomiser 3):** proposto por Quinlan (1986); usa ganho de informação para escolher atributos; é a base da evolução de algoritmos de árvore posteriores. Pode ser usado na classificação de sintomas clínicos para decidir se um paciente deve ser encaminhado a uma unidade de emergência.
- **C4.5: Evolução do ID3, também de Quinlan (1993):** introduz normalização do ganho de informação (razão de ganho), poda de árvores para evitar overfitting e suporte a atributos contínuos. Pode ser utilizado na análise de exames laboratoriais para identificar risco de diabetes, por exemplo.
- **Random Forests:** combina várias árvores treinadas em subconjuntos aleatórios dos dados, aumentando a robustez. Exemplo em saúde: sistema de apoio à decisão clínica para prever complicações pós-cirúrgicas.
- **Boosting (AdaBoost e Gradient Boosting):** combina árvores fracas em sequência, corrigindo os erros da anterior. Exemplo em saúde: modelos de previsão de mortalidade em UTI.

**Overfitting** é quando um modelo de aprendizado de máquina se adapta tão perfeitamente aos dados de treinamento que aprende até mesmo o ruído e os padrões aleatórios, perdendo a capacidade de generalizar e fazer previsões precisas em dados novos. É como um modelo que funciona perfeitamente em uma simulação, mas falha no "mundo real".

Segundo Russell e Norvig, os algoritmos baseados em árvore se tornaram populares não apenas por sua eficiência computacional e robustez em dados ruidosos, mas também porque oferecem explicações compreensíveis uma característica crítica em domínios como o da saúde, em que decisões precisam ser justificáveis diante de médicos, pacientes e órgãos regulatórios.

## AtualizAI

*Uma revisão integrativa consolidou o impacto significativo dos modelos de IA na predição de sepse neonatal, destacando a eficácia de algoritmos como Random Forest, SVM e Regressão Logística na identificação precoce de fatores de risco críticos, como hematócrito baixo, corioamnionite materna e variabilidade da frequência cardíaca. Esses modelos também vêm sendo aplicados em outras áreas da saúde, como na detecção antecipada de insuficiência respiratória em pacientes internados, na previsão de readmissões hospitalares e no monitoramento automatizado de sinais vitais em Unidades de Terapia Intensiva. Em sistemas de apoio à decisão clínica, técnicas semelhantes têm auxiliado médicos a reconhecer padrões sutis em exames laboratoriais e a ajustar condutas terapêuticas com base em dados históricos, contribuindo para diagnósticos mais precisos, intervenções rápidas e redução da mortalidade neonatal e hospitalar.*

*Essa precocidade na detecção permite que intervenções sejam realizadas em um estágio onde a eficácia do tratamento pode ser maximizada, potencialmente reduzindo a mortalidade e as complicações associadas à sepse (Bernardino et al., 2024).*

Clique [aqui](#) para saber mais!

## ALGORITMOS DA ABORDAGEM EVOLUCIONÁRIA

A abordagem evolucionária é discutida no contexto de algoritmos inspirados na biologia, em especial nos processos de seleção natural, reprodução e mutação. Esses métodos são agrupados sob o guarda-chuva da chamada computação evolucionária, que busca explorar espaços de busca complexos por meio de mecanismos análogos à evolução biológica.

Russell e Norvig apresentam a ideia de que, assim como organismos vivos evoluem em ambientes competitivos, algoritmos podem evoluir soluções para problemas difíceis em ambientes de busca. A metáfora biológica é aplicada de forma matemática: populações de soluções candidatas competem, se reproduzem e são submetidas a mutações. Ao longo das gerações, a qualidade média das soluções tende a melhorar, aproximando-se de um ótimo global. Principais Algoritmos da Abordagem Evolucionária:

- **Genetic Algorithms (GA):** Propostos por John Holland (1975); representam soluções como cadeias de genes (representações simbólicas); operadores principais: seleção, crossover (recombinação) e mutação. Exemplo em saúde: otimizar a dosagem de múltiplos medicamentos em protocolos de quimioterapia para maximizar eficácia e minimizar efeitos colaterais.
- **Genetic Programming (GP):** Variante dos GA em que os indivíduos são programas de computador; evoluem estruturas de árvore que representam expressões ou programas. Exemplo em saúde: descoberta automática de equações ou regras para prever progressão de doenças crônicas.
- **Evolution Strategies (ES):** Desenvolvidas por Rechenberg e Schwefel (década de 1960); focam em problemas contínuos de otimização; usam operadores de mutação adaptativa e seleção baseada em aptidão. Exemplo em saúde: ajuste fino de parâmetros em modelos de previsão de epidemias.
- **Evolutionary Programming (EP):** Proposta por Lawrence Fogel (1960); originalmente voltada para evolução de máquinas de estados finitos; se destaca pelo uso extensivo de mutação como operador central. Exemplo em saúde: modelagem preditiva de fluxos de pacientes em hospitais.

Russell e Norvig reforçam que os algoritmos evolucionários são especialmente úteis em espaços de busca de alta dimensionalidade; problemas de otimização combinatória (ex.: roteamento, agendamento); cenários de incerteza, nos quais métodos analíticos clássicos falham. Na área da saúde, exemplos notáveis incluem: Otimização de protocolos de tratamento personalizados; planejamento de radioterapia com ajuste de feixes de radiação; evolução de redes neurais (neuroevolução) para análise de imagens médicas.

A abordagem evolucionária, segundo Russell e Norvig, mostra como a biologia inspirou a IA a desenvolver algoritmos robustos para problemas complexos. Embora menos previsíveis e mais custosos computacionalmente que abordagens determinísticas, esses algoritmos são poderosos em cenários onde o espaço de busca é vasto e mal estruturado.

## ALGORITMOS DA ABORDAGEM ENXAMES

A chamada abordagem de Inteligência de Enxames (Swarm Intelligence) aparece como uma vertente dos métodos inspirados na natureza, especialmente no comportamento coletivo de organismos sociais como formigas, abelhas, pássaros e peixes. Essa linha de algoritmos busca resolver problemas de otimização e busca em espaços complexos a partir de interações locais entre múltiplos agentes simples, que em conjunto produzem comportamentos globais adaptativos.

Russell e Norvig destacam que a inteligência emergente pode surgir da cooperação descentralizada de agentes que seguem regras simples. Essa ideia rompe com a visão centralizada da IA simbólica e aproxima-se de conceitos biológicos e de sistemas complexos. Na prática, cada agente do “enxame” contribui com pequenas decisões baseadas em informação local, e o sistema como um todo converge para soluções de qualidade. Possuem como características fundamentais a Auto-organização, isto é, não há controlador central; Interações locais onde agentes ajustam comportamento com base nos vizinhos ou no ambiente; robustez, isto é, tolerância a falhas individuais; escalabilidade com eficácia em ambientes grandes e dinâmicos.

Os principais algoritmos de Inteligência de Enxames, são:

- **Ant Colony Optimization (ACO):** Proposto por Marco Dorigo nos anos 1990; baseia-se no comportamento de formigas reais que depositam feromônios para indicar rotas onde cada formiga-agente constrói soluções, reforçando as melhores trilhas com maior depósito de feromônio. Exemplo em saúde: Otimização de rotas de ambulâncias em tempo real para reduzir o tempo de resposta em emergências.
- **Particle Swarm Optimization (PSO):** Inspirado no movimento coletivo de bandos de aves e cardumes de peixes (Kennedy e Eberhart, 1995), onde cada partícula representa uma solução candidata que ajusta sua trajetória com base em duas informações: sua melhor posição histórica e a melhor posição do grupo. Exemplo em saúde: Utilização em ajuste de parâmetros em redes neurais aplicadas ao diagnóstico de imagens médicas.
- **Bee Colony Optimization (BCO):** Inspirados na forma como abelhas exploram fontes de alimento, equilibram exploração (busca de novas regiões) e exploração intensiva (refinamento de boas regiões). Exemplo em saúde: Pode ser utilizado na alocação de recursos hospitalares em situações de alta demanda, como surtos e epidemias.

Russell e Norvig enfatizam que algoritmos de enxames são aplicáveis a problemas de otimização combinatória e contínua, típicos da área médica: 1. Diagnóstico auxiliado por computador: ajuste de classificadores para detecção de câncer em exames de imagem; 2. Gestão hospitalar: otimização de escalas de médicos e enfermeiros, roteamento de ambulâncias, alocação de leitos; 3. Saúde pública: modelagem e previsão da propagação de epidemias, otimizando campanhas de vacinação.

A abordagem de inteligência de enxames, conforme sistematizada por Russell e Norvig, mostra como estratégias simples, inspiradas no comportamento coletivo da natureza, podem gerar soluções eficazes para problemas altamente complexos. Diferente dos métodos centralizados, essa abordagem enfatiza cooperação descentralizada, auto-organização e adaptabilidade, características cruciais para enfrentar ambientes incertos e dinâmicos.

## ALGORITMOS DA ABORDAGEM NEURAL OU CONEXIONISTA

A abordagem neural é apresentada como uma das mais influentes no campo do aprendizado de máquina. Inspirada no funcionamento do cérebro humano, essa abordagem busca modelar redes de neurônios artificiais que, conectados em camadas, aprendem representações complexas a partir de dados.

Russell e Norvig explicam que a ideia central das Redes Neurais Artificiais (RNAs) é capturar a noção de aprendizado distribuído e adaptativo: Cada neurônio artificial executa uma função de ativação simples; conexões entre neurônios são ajustadas por meio de pesos sinápticos; o aprendizado ocorre pela atualização iterativa desses pesos com base em exemplos de entrada e saída. Essa abordagem é particularmente eficaz quando o problema envolve padrões não lineares complexos, grande volume de dados e ruído.

Os principais Algoritmos Neurais:

- **Multi-Layer Perceptron (MLP):** Extensão do perceptron com uma ou mais camadas ocultas. Utiliza o algoritmo de retropropagação (Backpropagation) para ajustar pesos. Capaz de modelar funções não lineares complexas. Exemplo em saúde: diagnóstico automático de câncer em imagens de biópsia.
- **Radial Basis Function (RBF):** Cada neurônio da camada oculta utiliza uma função radial (geralmente gaussiana). Úteis em classificação e regressão com dados complexos. Exemplo em saúde: previsão de evolução de doenças cardíacas com base em múltiplos parâmetros clínicos.
- **Convolutional Neural Networks (CNN):** Inspiradas no córtex visual humano. Utilizam convoluções para extrair padrões espaciais em imagens. Exemplo em saúde: detecção de tumores em exames de ressonância magnética.

Russell e Norvig destacam alguns pontos fortes: 1. Generalização, conseguem aprender a partir de exemplos e se adaptar a novos dados; 2. Paralelismo, adequadas para processamento em larga escala; 3. Flexibilidade, aplicável em diversas áreas (visão, fala, texto, saúde, finanças). Mas também apontam limitações: Requerem grandes quantidades de dados, são computacionalmente intensivas, e frequentemente funcionam como “caixas-pretas”, dificultando a interpretabilidade.

A abordagem neural é particularmente poderosa em contextos médicos, tais como:

- **Visão Computacional:** detecção de câncer em imagens de biópsia (CNNs);
- **Previsão de Séries Temporais:** monitoramento de ECG e EEG (RNNs/LSTMs);
- **Modelos Preditivos:** prognóstico de doenças crônicas (MLPs, RBFs);
- **Redes Autoassociativas:** recuperação de padrões incompletos em exames (Hopfield).

Segundo Russell e Norvig, os algoritmos da abordagem neural representam um dos pilares centrais da IA moderna. Desde o perceptron de Rosenblatt até as CNNs, LSTMs e Transforms atuais, esses modelos mostraram como a inspiração biológica pode gerar sistemas computacionais robustos, capazes de aprender com dados e resolver problemas complexos em ambientes incertos.

## AtualizAI

Um bom exemplo de utilização da abordagem neural para o planejamento em saúde pode ser observado no artigo “Uso das redes neurais artificiais na aplicação de metodologia para alocação de recursos da saúde”, que teve como objetivo desenvolver um fator de alocação de recursos financeiros em saúde, considerando variáveis epidemiológicas, demográficas, socioeconômicas e educacionais dos municípios de Pernambuco.

A metodologia permitiu analisar 18 indicadores e gerar um fator de alocação (FA) que classificou os municípios em quartis de necessidade. O fator de alocação separou os municípios que possuem extremas necessidades de maior aporte financeiro daqueles que precisam com menor intensidade. O uso de metodologias como esta propõe que a distribuição de recursos deve ser contextualizada e, conseqüentemente, mais justa.

Clique [aqui](#) para saber mais!

## QUER SABER MAIS?

Acesse <https://alvarofpinheiro.net/ciencia-dados> e confira vídeo aulas com fundamentações e exemplos práticos que mostram, passo a passo, como aplicá-la.

## AULA VIRTUAL



### Módulo 2 - Aula 2 - Classes de Problemas e tipos de aprendizado



# MÓDULO 3

## ENGENHARIA DE PROMPTS



## APRESENTAÇÃO

A Engenharia de Prompt emerge como uma disciplina estratégica no campo da Inteligência Artificial Generativa (Gen-AIs), voltada para projetar, otimizar e estruturar comandos chamados prompts que orientam modelos de linguagem, imagem e sistemas multimodais a produzirem respostas mais úteis, precisas e contextualizadas.

Essa prática é essencial para aproveitar todo o potencial das redes neurais generativas, que são modelos de IA capazes de criar textos, imagens, sons e análises complexas a partir de instruções em linguagem natural. Com o avanço e a popularização dos LLMs (Large Language Models) a partir de 2018, impulsionados pela arquitetura Transformer e seus grandes progressos em Processamento de Linguagem Natural (NLP), dominar a Engenharia de Prompt tornou-se não apenas uma habilidade técnica, mas também criativa e analítica, cada vez mais valorizada em ambientes corporativos, acadêmicos e governamentais (Pinheiro, 2025).

No contexto atual, dominar prompts não é apenas uma habilidade técnica, mas também criativa e analítica, sendo cada vez mais requerida em ambientes corporativos, acadêmicos e governamentais.

Fundamentos históricos:

- **1940** – Neurônio artificial de McCulloch-Pitts;
- **1950** – Alan Turing e o artigo “O Jogo da Imitação”;
- **1960** – Surgimento do ELIZA, primeiro chatbot de Processamento de Linguagem Natural;
- **1980** – Redes Neurais Recorrentes (RNNs), aplicadas em tradução automática;
- **2017** – Arquitetura Transformer, marco para a compreensão de dependências longas;
- **2018** – Avanço dos LLMs, ampliando aplicações de automação inteligente.

**Chatbot** é um programa de computador que simula uma conversa humana, seja por texto ou voz, para realizar tarefas e responder a perguntas. Ele pode variar de sistemas simples com respostas pré-programadas a outros mais complexos que usam inteligência artificial e processamento de linguagem natural para entender e gerar respostas sofisticadas. Chatbots são usados em áreas como atendimento ao cliente, vendas e para fornecer assistência pessoal.

**Cuidado!** As Redes Neurais Generativas, não são IAs Gerais, essa categoria de Inteligência Artificial, as Artificial General Intelligence (AGI) ou Inteligência Artificial Geral (IAG) ainda não existem, apesar de serem o objetivo das Big Techs. Porém, apesar de serem surpreendentes as Generative Artificial Intelligence (GenAI) ou Inteligência Artificial Generativa (IA-Gen), “só” possuem capacidade de criar dados sintéticos com base no aprendizado de dados naturais, mas não são capazes de criar algo verdadeiramente novo.

Lembrando que passamos por diferentes etapas da evolução da Inteligência Artificial: IA Fraca (1940–1980), voltada a sistemas capazes de executar tarefas específicas, como cálculos e classificações simples; IA Forte (1980–2010), caracterizada pelo avanço da capacidade de raciocínio lógico e pelo surgimento dos primeiros agentes inteligentes; IA de Aprendizado Profundo (2010–2017), impulsionada por redes neurais com múltiplas camadas, que permitiram reconhecer padrões complexos em imagens e textos; e IA Generativa (2017–atual), responsável por criar conteúdos originais, como textos, imagens e sons, a partir de instruções em linguagem natural.

Atualmente, o foco das Big Techs, como a OpenAI, é desenvolver uma IA Geral (AGI), que seria capaz de compreender, aprender e executar qualquer tarefa intelectual humana. A etapa seguinte, ainda teórica, é a Superinteligência, na qual as máquinas ultrapassariam a capacidade cognitiva humana. Assim, podemos classificar seis categorias de IAs: IA Fraca, IA Forte, IA de Aprendizado Profundo, IA Generativa, IA Geral e Superinteligência, sendo importante destacar que as duas últimas ainda não existem, mas são alvo de intensa pesquisa e especulação científica.

O que é um **Prompt**? Um prompt é a solicitação enviada a um modelo de IA podendo assumir a forma de instrução, comando ou contexto. Seu papel é orientar o modelo para produzir respostas específicas e úteis. Prompts eficazes precisam de clareza, delimitação e objetivos bem definidos. São dez os elementos fundamentais da Engenharia de Prompt:

- Persona, define o estilo de resposta (ex.: professor, médico, advogado, especialista, ...);
- Tarefa, descreve o que deve ser feito (ex.: gerar roteiro, revisar texto, criar relatório, ...);
- Efeito, utiliza modificadores para intensificar a resposta (ex.: crítico, analista, criativo, ...);
- Limite, restrições do resultado (ex.: tempo, recursos, perfil de público, ...);
- Forma, formato esperado (ex.: tabela, artigo, relatório, ...);
- Fonte, informações a serem consideradas (ex.: documentos anexados, bases externas, ...);
- Tamanho, extensão da resposta (ex.: curta, detalhada, por etapas, ...);
- Análise, solicitação de justificativas, raciocínio passo a passo (Chain of Thought);
- Originalidade, espaço para criatividade e surpresa;
- Contexto, histórico de conversa e informações adicionais para consistência.

Esses elementos podem ser enriquecidos com a incorporação de aspectos estilísticos e comunicativos, divididos em:

- Personalidade (ex.: didática, crítica, empática, formal, casual, ...);
- Tom (ex.: amigável, respeitoso, irônico, inspirador, ...);
- Linguagem (ex.: acadêmica, técnica, jornalística, coloquial, ...);
- Estilo (ex.: descritivo, narrativo, expositivo, persuasivo, ...);
- Propósito (ex.: ensinar, informar, persuadir, criticar, motivar, ...).

A combinação desses aspectos permite precisão contextual e adaptabilidade a diferentes públicos e finalidades. A aplicabilidade de técnicas avançadas de Engenharia de Prompt também é importante:

- **Chain of Thought (CoT):** pedir raciocínio passo a passo;
- **Multiagentes:** dividir tarefas em diferentes agentes especialistas que interagem entre si;
- **Custom Instructions:** personalizar parâmetros globais do modelo;
- **Uso de dados anexados:** fornecer documentos ou datasets (conjunto de dados) para enriquecer respostas;
- **Biblioteca de Prompts:** criação de repositório temático (recrutamento, notas técnicas, revisões acadêmicas, artigos, aulas, ementas, ...).

Figura 12: Elementos e Aspectos para Engenharia de Prompt.



Fonte: Pinheiro, 2025.

Outro ponto importante de se observar é a privacidade e a ética. É essencial respeitar as diretrizes éticas e legais: evitar compartilhamento de dados pessoais e sensíveis. Atentar para a Lei Geral de Proteção de Dados Pessoais (LGPD) e regulamentações sobre uso de IA em ambientes institucionais. Compreender as limitações dos modelos, que podem produzir informações plausíveis, mas não necessariamente verdadeiras.

A Engenharia de Prompt já é aplicada em diversos contextos, como: Educação, com a criação de materiais didáticos personalizados; Governo, com a elaboração de despachos, notas técnicas e documentos oficiais; Empresas, com a geração de relatórios, análises de dados e planos de marketing; e Saúde, com o apoio na redação de pareceres, protocolos e integração de prontuários.

A Engenharia de Prompt 2.0 representa uma nova competência transversal, unindo ciência de dados, linguística computacional e habilidades criativas. Sua prática adequada potencializa produtividade, inovação e qualidade da interação entre humanos e sistemas de IA. Pois, “A IA é poderosa, mas é a imaginação humana, guiada por prompts estruturados, que define seus limites e alcances” (Pinheiro, 2024).

## AtualizAI

Aqui estão algumas recomendações listadas por KARAMI (2014) para a criação de Prompts a partir de LLM's:

- Seja o mais específico possível.

Exemplo pouco específico: “Me fale sobre doenças crônicas não transmissíveis”.

Exemplo mais específico: “Quais são os fatores de risco mais comuns para as doenças crônicas não transmissíveis?”

- Identifique o objetivo geral do Prompt primeiro.

Exemplo: Gostaria de receber uma lista de 5 ideias para uma apresentação em um evento científico para tornar as descobertas da minha pesquisa mais fáceis de entender.

- Tenha consciência temporal sobre o Prompt.

Exemplo sem referência de tempo: “Descreva o processo de recuperação de uma laqueadura tubária (esterilização feminina).”

Exemplo com referência de tempo: “O que uma paciente normalmente pode esperar durante as primeiras semanas de recuperação após uma Laqueadura tubária?”

- Use o método de prompt de uma tentativa/poucas tentativas.

Exemplo: Gere 10 nomes possíveis para um evento de promoção à vacinação para a Gripe. Nomes que eu gosto incluem as palavras: Saúde, Prevenção, Cuidado.

Você também pode: Definir expectativas realistas; Descrever o cenário e fornecer o contexto em torno da pergunta; Experimentar diferentes estilos de Prompt; Solicitar exemplos de Prompt ou sugerir Prompts.



## BIG TECHS

Quando se trata do desenvolvimento de Redes Neurais Generativas, um seleto grupo de grandes empresas de tecnologia, conhecidas como big techs, ocupa posição central no avanço e na consolidação dessas inovações. Entre as mais proeminentes está a OpenAI, responsável pela criação de modelos como Generative Pre-trained Transformer (GPT), que revolucionaram o processamento de linguagem natural e abriram caminho para assistentes virtuais cada vez mais sofisticados. A OpenAI se destaca não apenas pela excelência técnica, mas também pelo impacto social e ético que suas pesquisas e produtos suscitam no debate público e acadêmico.

Outro gigante nesse ecossistema é o Google, cuja divisão de pesquisa em IA desenvolveu a arquitetura Transformer, base fundadora da maioria dos Large Language Models (LLMs) modernos. O Google também lançou o modelo Gemini e investiu em aplicações avançadas, como o Bard (IA Generativa do Google, projetada para responder perguntas, criar textos e auxiliar usuários em tarefas cotidianas de forma interativa e eficiente) e sistemas integrados ao ambiente Workspace, levando a IA generativa a ferramentas cotidianas. A atuação do Google se estende à pesquisa aberta, colaborando com a comunidade científica global e fomentando o surgimento de soluções híbridas entre texto, imagem e código.

A Microsoft, reconhecida como parceira estratégica da OpenAI, tem desempenhado um papel fundamental no avanço da inteligência artificial generativa ao realizar investimentos robustos e contínuos nessa tecnologia. A empresa não apenas financia o desenvolvimento de modelos como o GPT, mas também incorpora essas soluções de IA diretamente em sua suíte de produtos e serviços, como o Copilot, presente no Microsoft 365. Essa integração permite que usuários de plataformas corporativas tenham acesso a assistentes inteligentes capazes de automatizar tarefas rotineiras, gerar conteúdos personalizados, analisar grandes volumes de dados e oferecer recomendações precisas em tempo real.

A Meta, anteriormente conhecida como Facebook, tem investido de maneira significativa no desenvolvimento de modelos de inteligência artificial generativa, sendo o LLaMA (Large Language Model Meta AI) um dos principais exemplos de sua atuação nesse campo. O LLaMA é especialmente destacado por sua arquitetura de código aberto, o que significa que pesquisadores, desenvolvedores e empresas de todo o mundo podem acessar, estudar e modificar o modelo conforme suas necessidades. Essa abordagem da Meta visa democratizar o acesso à tecnologia de IA permitindo que a comunidade global experimente, adapte e crie soluções inovadoras a partir da base fornecida pela empresa. A Meta também fomenta a colaboração entre diferentes setores e acelera o ritmo de inovação, estimulando a criação de novas aplicações, ferramentas e pesquisas na área de processamento de linguagem natural e inteligência artificial generativa.

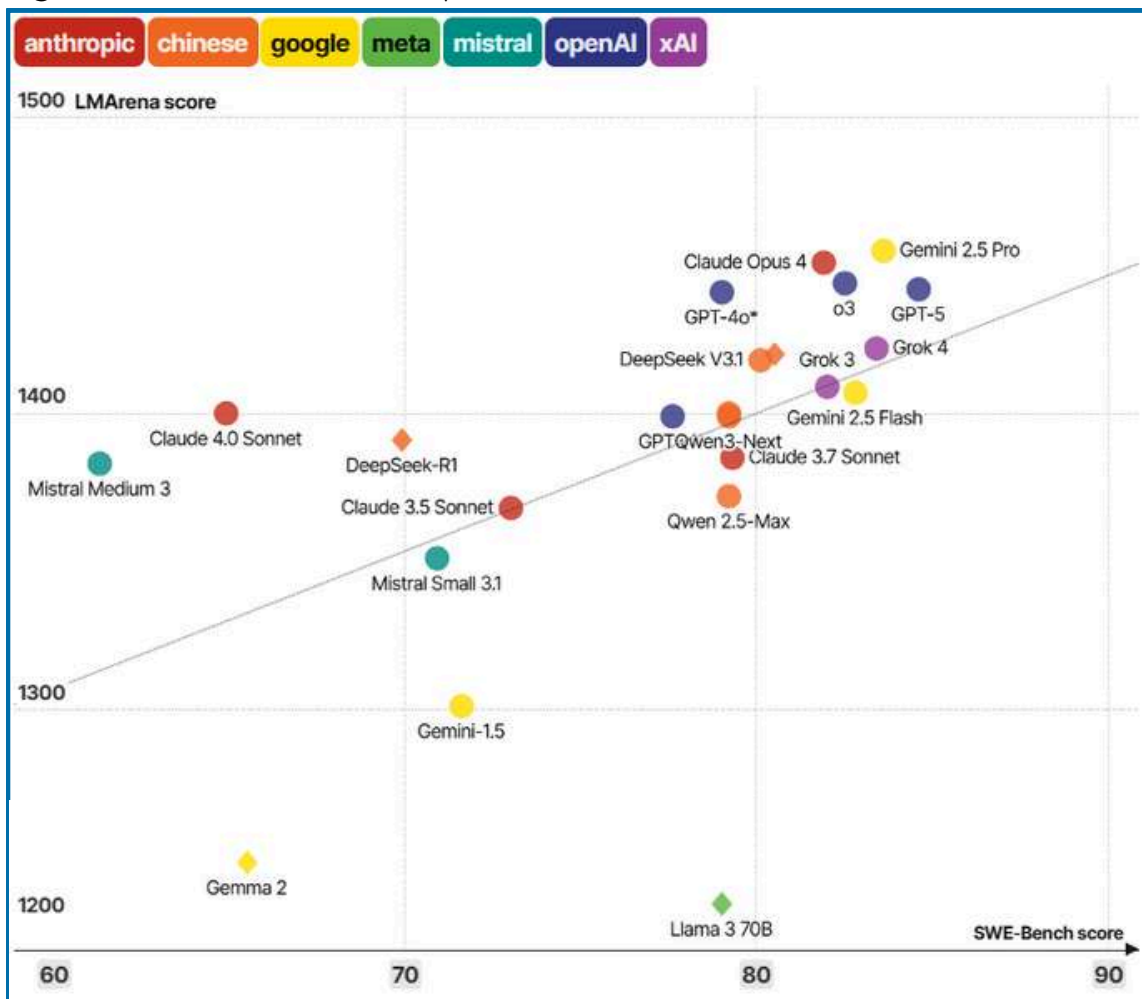
A Amazon também se destaca como uma das principais líderes mundiais no campo da inteligência artificial generativa. A empresa tem investido pesadamente no desenvolvimento e aplicação de redes neurais generativas, principalmente por meio da Amazon Web Services (AWS), sua plataforma de serviços em nuvem. Nessas soluções, a Amazon oferece ferramentas de IA que permitem desde a criação automatizada de conteúdos até o processamento avançado de dados em escala global.

Além disso, a empresa utiliza modelos generativos para aprimorar o atendimento ao cliente, automatizando respostas em seus canais de suporte e promovendo uma experiência mais ágil e personalizada. Outro aspecto relevante é a personalização em larga escala: algoritmos de IA são empregados para analisar o comportamento dos usuários, recomendando produtos e ofertas sob medida para cada perfil, o que potencializa a fidelização e o engajamento dos clientes. Essas iniciativas refletem o compromisso da Amazon em integrar tecnologias avançadas de IA em seus processos, consolidando sua posição de vanguarda na transformação digital de múltiplos setores.

Empresas chinesas de tecnologia, como Baidu, Alibaba e Tencent, vêm intensificando substancialmente seus investimentos na área de inteligência artificial generativa, desenvolvendo modelos próprios que rivalizam com as soluções ocidentais.

Essas Big Techs impulsionam a pesquisa e determinam o ritmo da transformação digital em múltiplos setores, combinando poder computacional, expertise em dados e abordagens colaborativas. Com isso, moldam o futuro das interações homem-máquina, criam paradigmas de produtividade e levantam discussões profundas sobre ética, governança e impacto social das tecnologias emergentes.

**Figura 13:** Maiores LLMs Ranqueadas.



Fonte: LLMarena & Bench Scores, 2025.

Entre os principais LLMs (Large Language Models) em destaque no cenário atual, o GPT (Generative Pre-trained Transformer), desenvolvido pela OpenAI, ocupa posição central. O GPT se destaca por sua capacidade de compreender e gerar textos com elevado grau de coerência, sendo utilizado em assistentes virtuais, geração automática de conteúdo e suporte à tomada de decisão em empresas e instituições públicas. Sua arquitetura Transformer, baseada em atenção contextual, permitiu saltos significativos de produtividade e inovação na área de processamento de linguagem natural.

Outro modelo de grande relevância é o Gemini, desenvolvido pelo Google. O Gemini representa a convergência de abordagens avançadas em IA unindo processamento textual, visual e até mesmo de código, o que amplia enormemente seu espectro de aplicações. Incorporado a ferramentas como o Bard e ao ecossistema Workspace, o Gemini facilita desde automações simples até tarefas complexas de análise de dados e geração de relatórios estruturados.

Já o Claude, criado pela Anthropic, é um LLM que enfatiza princípios de segurança, alinhamento e ética. O Claude foi projetado para minimizar vieses e riscos, oferecendo respostas mais seguras, controladas e transparentes, o que tem atraído interesse em áreas reguladas, como saúde e governo. Sua arquitetura é voltada para garantir robustez e previsibilidade, tornando-se referência em discussões sobre governança de IA.

Além desses, o Large Language Model Meta AI, mais conhecido com LLaMA, da Meta, fomenta a experimentação e o avanço coletivo da comunidade de desenvolvedores. Empresas chinesas também inovam com modelos próprios, como o ERNIE (Baidu), Tongyi Qianwen (Alibaba) e Hunyuan (Tencent), além dos DeepSeek e Manus, ampliando a diversidade e a competitividade global no campo de LLMs.

Essas soluções trazem não só avanços técnicos, mas também novos paradigmas de uso responsável, privacidade e transparência, sendo aplicadas desde a automação de tarefas administrativas até a criação de conteúdos artísticos e científicos. O cenário dos LLMs está em constante evolução, alimentando pesquisas e transformando a forma como interagimos, aprendemos e trabalhamos em diferentes setores.

Compreender as principais inteligências artificiais existentes, como GPT, LLaMA, Gemini, Claude e as soluções desenvolvidas pelas grandes empresas globais, é fundamental para o público-alvo que busca se manter atualizado e competitivo em um mercado cada vez mais digitalizado. Esse conhecimento permite a identificação de oportunidades de inovação, automação de processos e otimização de operações, além de favorecer a tomada de decisões estratégicas informadas sobre adoção de tecnologias, privacidade e ética. Ao dominar as características e aplicações desses modelos, profissionais e organizações podem aproveitar melhor as vantagens da IA adaptando-se rapidamente às transformações do cenário tecnológico e potencializando seus resultados.

## IA DA OPENAI

A OpenAI é uma das organizações mais influentes no desenvolvimento da Inteligência Artificial generativa, destacando-se especialmente pelos seus modelos de linguagem natural conhecidos como GPT (Generative Pre-trained Transformer). A trajetória evolutiva desses modelos não apenas impulsionou avanços técnicos, mas também redefiniu os limites da automação, criatividade e produtividade em diversos setores.

### GPT-1: O Ponto de Partida (2018)

O primeiro modelo da série, o GPT-1, foi lançado em 2018. Ele utilizava a arquitetura de rede neural artificial denominada de Transformer, baseada em atenção contextual, permitindo que o sistema compreendesse padrões de linguagem em grandes volumes de texto, ainda que com capacidade limitada. O GPT-1 funcionava como uma prova de conceito para o uso de pré-treinamento generalizado seguido de ajuste fino (fine-tuning) para tarefas específicas. Este modelo estabeleceu as bases para que modelos posteriores aprendessem com grandes volumes de texto e gerassem respostas contextualmente relevantes.

### GPT-2: Salto de Escala e Abertura (2019)

Em 2019, o lançamento do GPT-2 representou um grande salto em escala técnica e funcional. Com 1,5 bilhão de parâmetros, o GPT-2 conseguia gerar textos surpreendentemente coerentes, imitar estilos de escrita diversas, compor textos longos e realizar síntese de informações. O modelo chamou atenção pelo seu grau avançado de geração de linguagem, mas seu lançamento foi inicialmente restrito devido a preocupações éticas e sobre o potencial uso indevido. Posteriormente, o modelo foi liberado, fomentando a pesquisa aberta e o avanço coletivo da comunidade científica.

### GPT-3: Expansão e Versatilidade (2020)

O GPT-3, anunciado em 2020, trouxe uma expansão exponencial com seus 175 bilhões de parâmetros. O modelo passou a ser capaz de executar uma variedade ainda maior de tarefas, como tradução, resumo, geração de código, composição de poesias e respostas a perguntas complexas, tudo sem ajuste fino específico. O GPT-3 destacou-se pela impressionante capacidade de “few-shot learning”, isto é, técnica de aprendizado de máquina em que um modelo consegue aprender a realizar uma tarefa a partir de poucos exemplos rotulados, para realizar novas tarefas. Essa versão foi amplamente integrada em assistentes virtuais, produção de conteúdo automatizado, suporte à decisão, entre outros usos corporativos e institucionais.

### GPT-3.5: Aperfeiçoamento e Otimização (2022)

O GPT-3.5 representa uma fase de refinamento entre o GPT-3 e o GPT-4. Ele trouxe melhorias significativas em compreensão e geração de textos, tornando as respostas mais precisas, contextuais e seguras. Essa versão foi utilizada como base para produtos como o ChatGPT, que popularizaram a interação conversacional com IA oferecendo respostas mais naturais e relevantes.

### GPT-4: Contexto, Multimodalidade e Rigor (2023)

O GPT-4 marca um novo patamar evolutivo. Além de ampliar a escala de parâmetros (estimativas sugerem centenas de bilhões), ele introduziu capacidades multimodais, ou seja, permite processar e gerar não apenas textos, mas também imagens, códigos e outros formatos de dados. O GPT-4 apresenta avanços em segurança, mitigação de vieses, alinhamento ético e robustez, sendo utilizado em ambientes mais regulados, como saúde, educação e governo. Sua arquitetura permite uma compreensão profunda de contextos complexos e uma geração ainda mais fluida e precisa de informações.

### GPT-Omini

A série Omini surge em resposta à crescente demanda por modelos de IA mais flexíveis, capazes de operar em múltiplas modalidades e de se adaptar a contextos complexos, dinâmicos e imprevisíveis. Enquanto as gerações anteriores dos modelos GPT já haviam revolucionado o processamento de linguagem natural, a Omini incorpora avanços em arquitetura neural, treinamento distribuído e integração de diferentes tipos de dados (texto, imagem, áudio, código, entre outros). Estes modelos são projetados para oferecer um desempenho superior tanto em tarefas específicas quanto em aplicações generalistas, potencializando sua aplicabilidade em setores como saúde, educação, indústria criativa, governança e pesquisa científica.

### GPT Omini O1: A Fundação Multimodal

O Omini O1 inaugura a série com uma proposta inovadora: ser um modelo verdadeiramente multimodal desde a sua concepção. Diferente dos antecessores que adicionavam capacidades multimodais posteriormente, o O1 nasce com arquitetura projetada para processar texto, imagens e dados estruturados de forma integrada.

- **Arquitetura:** Utiliza núcleos de atenção expandidos, capazes de cruzar informações entre diferentes formatos de dados. A estrutura permite a assimilação de contexto textual e visual simultaneamente, favorecendo aplicações como análise de documentos, geração de relatórios ilustrados e sistemas de recomendação.
- **Capacidade:** Conta com dezenas de bilhões de parâmetros, sendo treinado em corpora multiformato, incluindo textos científicos, literatura, imagens públicas e bases de dados abertas.

- **Inovações:** Introduz mecanismos de adaptação dinâmica de contexto, permitindo que o modelo ajuste sua saída de acordo com mudanças abruptas no tipo de dado recebido. Isso é essencial para aplicações em ambientes reais, onde múltiplas fontes de informação convergem.
- **Impacto:** O O1 popularizou a IA multimodal em setores corporativos, sendo adotado em plataformas de gestão documental, educação remota e sistemas interativos de atendimento ao público.

### GPT Omini O3: A Expansão da Versatilidade e Eficiência

Com o lançamento do Omini O3, a série avança significativamente em termos de capacidade de processamento, desempenho e flexibilidade. O O3 é, em essência, uma evolução que reflete o amadurecimento das técnicas de treinamento federado e o aperfeiçoamento dos algoritmos de alinhamento ético.

- **Escalabilidade:** O O3 alcança centenas de bilhões de parâmetros, tornando-se um dos modelos mais robustos já desenvolvidos até então. Seu treinamento se dá em ambientes distribuídos, aproveitando clusters globais para acelerar o aprendizado e incorporar diversidade cultural e linguística.
- **Versatilidade:** Diferente de modelos anteriores, o O3 não apenas processa texto e imagens com alto grau de precisão também interpreta áudio, vídeo e código-fonte. Isto amplia sua aplicação para áreas como geração automática de conteúdo multimídia, suporte técnico, análise de segurança digital e tradução audiovisual.
- **Eficiência energética:** Incorpora algoritmos de compressão neural e técnicas de inferência otimizadas, reduzindo o consumo de energia durante a operação. Essa característica é fundamental para empresas preocupadas com sustentabilidade e custos operacionais.
- **Personalização:** Uma das grandes inovações é o módulo de personalização profunda, capaz de ajustar comportamentos e respostas para perfis organizacionais ou individuais, mantendo privacidade e segurança dos dados.
- **Impacto:** O O3 amplamente adotado em plataformas de comunicação empresarial, sistemas de saúde preditivos, ferramentas de design automatizado e sistemas de recomendação avançados. Sua receptividade foi elevada pela capacidade de gerar resultados precisos e contextualizados em múltiplos domínios.

*Cluster é um grupo de computadores ou elementos interconectados que trabalham juntos como se fossem um único sistema, com o objetivo de aumentar o desempenho e a confiabilidade.*

## GPT Omini O4: Inteligência Generativa Universal

O Omini O4 representa o ápice da série, sendo considerado por muitos especialistas um marco na inteligência artificial generativa. O O4 não se limita ao processamento multimodal, mas inaugura a era da inteligência generativa universal, capaz de integrar conhecimento tácito, raciocínio simbólico e inferência probabilística.

- **Arquitetura híbrida:** O O4 combina camadas profundas de transformers com módulos de raciocínio simbólico, permitindo a dedução lógica e resolução de problemas complexos.
- **Aprendizado contínuo:** Absorção de novas informações em tempo real sem a necessidade de re-treinamento completo, garantindo que o O4 permaneça atualizado e relevante.
- **Integridade ética:** Apresenta avanços sem precedentes em mitigação de vieses e alinhamento ético, adotando sistemas de monitoramento e correção automática para evitar respostas inadequadas.
- **Interoperabilidade:** O modelo é desenhado para interação direta com outros sistemas de IA, plataformas de dados e redes de sensores, promovendo ecossistemas integrados de inteligência artificial colaborativa.
- **Capacidade de raciocínio:** Supera limitações dos modelos anteriores ao demonstrar habilidades de raciocínio abstrato, síntese criativa e resolução de tarefas que exigem múltiplos passos lógicos, tornando-se referência em pesquisas avançadas e aplicações industriais.
- **Impacto:** O O4 está sendo aplicado em áreas de ponta como pesquisa científica automática, desenvolvimento de medicamentos, governança inteligente, educação personalizada e robótica autônoma.

**Benchmark** significa um ponto de referência, padrão ou critério de excelência usado para avaliar o desempenho de algo, como um processo, produto ou serviço, em relação aos melhores do mercado.

**Benchmarking** é o processo de fazer essa análise comparativa com concorrentes ou líderes de mercado para encontrar maneiras de melhorar.

## Impactos Sociais e Éticos dos Modelos Omini

A série Omini não apenas redefiniu o que é possível em termos de geração automática de conhecimento, mas também levantou discussões importantes sobre responsabilidade, privacidade e governança da inteligência artificial. Cada geração trouxe avanços em alinhamento ético, mitigação de vieses e transparência, buscando garantir que a IA seja benéfica e acessível para todos. No entanto, o aumento de poder e autonomia desses sistemas exige uma vigilância constante e esforços colaborativos entre desenvolvedores, reguladores e sociedade civil.

Com a evolução dos modelos Omini, vislumbra-se um futuro em que a IA será uma parceira ativa na solução de problemas globais, criação de conhecimento, educação personalizada e inovação científica. O desenvolvimento contínuo de arquiteturas híbridas, aprendizado dinâmico e integração multimodal promete expandir ainda mais o alcance e a utilidade da inteligência artificial. O desafio agora é assegurar que esses benefícios sejam distribuídos de forma equitativa, respeitando princípios éticos e promovendo um ambiente de colaboração entre humanos e máquinas.

## GPT 5.0: Uma Nova Fronteira na Inteligência Artificial

O GPT 5.0 surge como um marco tecnológico, combinando capacidades de raciocínio avançado, compreensão contextual profunda e integração multimodal, estabelecendo novos padrões de excelência em aplicações industriais, científicas e educacionais. O GPT 5.0 foi projetado sobre uma arquitetura aprimorada baseada em transformers, incorporando técnicas de aprendizado contínuo e mecanismos de ajuste dinâmico que permitem absorver informações em tempo real. Com trilhões de parâmetros e uma estrutura hierárquica que favorece o raciocínio abstrato, o modelo consegue realizar síntese criativa, resolução de problemas complexos e geração de texto com elevado grau de coerência e originalidade.

- **Compreensão de texto e contexto:** Supera versões anteriores ao interpretar nuances semânticas e pragmáticas, adaptando respostas conforme o contexto e as necessidades do usuário.
- **Interação multimodal:** Integra dados textuais, visuais e sonoros, viabilizando aplicações que vão desde análise de imagens até processamento de áudio, enriquecendo a experiência do usuário.
- **Aprendizado dinâmico:** Capacidade de incorporar conhecimentos recentes sem re-treinamentos extensivos.
- **Raciocínio lógico e abstrato:** As melhorias na arquitetura promovem desempenho superior em tarefas que exigem múltiplos passos lógicos, tornando o modelo apto para pesquisa científica automatizada, desenvolvimento de soluções inovadoras e educação personalizada.

A versatilidade do GPT 5.0 permite sua aplicação em diversos setores, transformando a maneira como organizações e indivíduos interagem com a inteligência artificial.

- **Pesquisa científica:** Automatização de revisão bibliográfica, criação de hipóteses e geração de relatórios detalhados com rigor acadêmico.
- **Indústria farmacêutica:** Apoio ao desenvolvimento de medicamentos por meio de análise de dados clínicos, simulação de interações moleculares e sugestões de novos compostos.
- **Educação personalizada:** Produção de conteúdos adaptados ao perfil do estudante, acompanhamento do progresso e recomendação de estratégias de aprendizagem.
- **Assistência médica:** Gerenciamento de prontuários, auxílio no diagnóstico precoce e personalização do atendimento ao paciente.
- **Robótica autônoma:** Integração do modelo em sistemas de robôs para navegação, comunicação e tomada de decisão em ambientes dinâmicos.

O avanço proporcionado pelo GPT 5.0 levanta questões fundamentais sobre responsabilidade, privacidade e governança. O modelo foi desenvolvido com sistemas de monitoramento ético, mitigação de vieses e transparência algorítmica, buscando garantir que suas contribuições sejam benéficas e seguras para toda a sociedade.

- **Mitigação de vieses:** Processos automáticos identificam e corrigem potenciais respostas inadequadas, promovendo equidade e inclusão.
- **Transparência:** Relatórios e mecanismos de auditoria tornam o funcionamento do modelo acessível para especialistas e comunidades interessadas.
- **Colaboração humano-máquina:** O GPT 5.0 incentiva parcerias entre desenvolvedores, reguladores e sociedade civil para a criação de padrões éticos robustos e compartilhados.

Apesar dos avanços, a implementação do GPT 5.0 demanda atenção contínua aos riscos associados à automação excessiva, manipulação de informações e segurança dos dados. O futuro aponta para o desenvolvimento de arquiteturas híbridas, integração com sistemas sensoriais e aprimoramento constante em alinhamento ético.

É fundamental que os benefícios do GPT 5.0 sejam distribuídos de maneira equitativa, promovendo um ambiente de colaboração e respeito entre humanos e máquinas. O modelo representa não apenas um salto tecnológico, mas também uma oportunidade de redefinir o papel da inteligência artificial na resolução dos grandes desafios globais.

## IA da Google

A Google, ao longo dos últimos anos, consolidou-se como uma das líderes globais no desenvolvimento de modelos de Inteligência Artificial, com contribuições marcantes para o avanço da geração de linguagem natural, compreensão multimodal e aplicações em larga escala. A seguir, apresenta-se um panorama evolutivo dos principais modelos lançados pela empresa, em ordem cronológica, evidenciando suas características técnicas, inovações e impactos.

## BERT: Entendimento Profundo do Texto (2018)

Lançado em 2018, o Bidirectional Encoder Representations from Transformers (BERT) inaugurou uma nova era na compreensão de linguagem natural. Ao permitir a análise bidirecional do contexto das palavras em uma frase, o BERT revolucionou tarefas como busca, classificação de texto e perguntas e respostas. Sua arquitetura baseada em transformers possibilitou resultados sem precedentes em benchmarks de NLP, isto é, comparar diferentes modelos, tornando-se rapidamente padrão para aplicações que exigem entendimento semântico refinado.

## T5: Unificando Tarefas de Linguagem (2019)

Em 2019, surge o modelo Text-to-Text Transfer Transformer (T5), que propôs um paradigma inovador: tratar todas as tarefas de processamento de linguagem natural como problemas de transformação de texto para texto. Esse conceito tornou possível utilizar um único modelo para tradução, resumo, classificação, entre outras atividades, aumentando significativamente a flexibilidade e o poder dos sistemas de IA.

## LaMDA: Conversação Natural e Aberta (2021)

Em 2021, a Google apresentou o Language Model for Dialogue Applications (LaMDA), focado em conversas naturais e abertas. O LaMDA foi projetado para compreender nuances de diálogos, manter contexto e oferecer respostas relevantes em interações prolongadas. O modelo destacou-se pela capacidade de criar experiências conversacionais mais fluidas e naturais, aproximando-se do modo como os seres humanos interagem.

## PaLM: Escala, Multilinguismo e Raciocínio (2022)

O Pathways Language Model (PaLM), lançado em 2022, marcou uma nova etapa de escalabilidade, ao atingir centenas de bilhões de parâmetros. Com suporte a múltiplos idiomas e capacidades avançadas de raciocínio e geração de código, o PaLM ampliou a fronteira das aplicações corporativas e de pesquisa, trazendo melhorias tanto em desempenho quanto em segurança e mitigação de vieses.

## Gemini: Multimodalidade e Integração (2023/2024)

A linha Gemini representa um salto rumo à IA multimodal, capaz de processar texto, imagens, áudio e vídeo de forma integrada. Lançada em versões crescentes, Gemini reflete o investimento da Google na unificação de formatos de dados e na ampliação da robustez dos modelos generativos. Suas aplicações abrangem desde assistentes inteligentes até ferramentas avançadas de análise de dados, educação e criatividade digital.

A evolução dos modelos de IA da Google demonstra uma trajetória de crescimento acelerado em escala, capacidade técnica e impacto social. Cada nova geração contribuiu para superar desafios em compreensão contextual, multimodalidade e ética, aproximando a Inteligência Artificial do cotidiano das pessoas e dos processos organizacionais. Com o contínuo avanço desses sistemas, espera-se que o papel da IA se torne cada vez mais central na maneira como a sociedade produz, consome e interage com informações e tecnologias.

## IA da Meta

A Meta, impulsionada por uma visão de democratização da IA, estabeleceu uma trajetória marcada por lançamento de modelos abertos e altamente escaláveis. O ponto de partida foi o Large Language Model Meta AI (LLaMA), apresentado em 2023, que se destacou pela eficiência em tarefas de compreensão textual e geração de linguagem natural, com arquitetura pensada para facilitar pesquisas e aplicações experimentais.

Logo em seguida, vieram as versões aprimoradas: LLaMA 2, lançado ainda em 2023, trouxe avanços em segurança, alinhamento e capacidade colaborativa, ampliando o alcance para desenvolvedores e empresas interessadas em construir soluções personalizadas. Este modelo reforçou o compromisso da Meta com transparência e acessibilidade, oferecendo recursos robustos para integração em fluxos corporativos e acadêmicos.

Mais recentemente, LLaMA 3 elevou o patamar da Inteligência Artificial da Meta, abraçando multimodalidade e aprimorando habilidades de interação contextual. Essa evolução permitiu que o modelo fosse utilizado não só em aplicações de texto, mas também na análise e geração de imagens, consolidando a Meta como um dos principais agentes na convergência de IA aberta e responsável. A cada lançamento, a Meta buscou intensificar mecanismos de mitigação de vieses, aprimorar o desempenho energético dos modelos e expandir sua integração com ecossistemas digitais globais.

Dessa forma, a progressão dos modelos da Meta reflete um compromisso claro com escalabilidade, inovação técnica e responsabilidade social, posicionando-se como protagonista no cenário internacional de Inteligência Artificial.

## IA da Anthropic

Fundada com foco em segurança e alinhamento ético, a Anthropic rapidamente tornou-se referência no desenvolvimento de modelos de linguagem avançados. Seu progresso começou com o Claude 1, introduzido em 2023, já trazendo refinamentos em compreensão textual, geração de conteúdo e mitigação de respostas indesejadas. Em seguida, o Claude 2 ampliou tanto a capacidade de processamento quanto a habilidade de manter diálogos extensos e realizar tarefas complexas, apresentando avanços notáveis em análise de contexto e raciocínio lógico.

A terceira geração, Claude 3, representa o atual estado da arte da Anthropic, incorporando multimodalidade, maior eficiência energética e profundo compromisso com transparência e responsabilidade. Estes modelos evoluíram para se integrar a diferentes plataformas e fluxos de trabalho, atendendo demandas de empresas, pesquisadores e desenvolvedores. A abordagem da Anthropic privilegia o desenvolvimento seguro e orientado por princípios, destacando-se pelo uso de técnicas como Constitutional AI, que busca proteger os usuários contra vieses e promover interações benéficas.

Assim, a trajetória dos modelos da Anthropic evidencia um movimento contínuo em direção à robustez, à escalabilidade e à integração ética, consolidando seu papel de liderança frente aos desafios contemporâneos da Inteligência Artificial.

## IA da Microsoft

Assim como outras empresas líderes no segmento, a Microsoft consolidou-se no desenvolvimento de modelos de IA que integram robustez técnica, versatilidade e responsabilidade social. Entre os destaques, encontra-se a família de modelos Azure OpenAI, que oferece acesso a tecnologias avançadas como GPT e DALL-E, promovendo soluções personalizáveis para organizações de todos os portes.

A Microsoft também investiu em modelos proprietários otimizados para tarefas específicas, como o Turing, focado em compreensão de linguagem natural e análise de grandes volumes de dados, e o Florence, voltado para aplicações multimodais que unem texto, imagem e vídeo. Essas plataformas são integradas de maneira fluida ao ecossistema do Microsoft 365 e ao Power Platform, democratizando o acesso à Inteligência Artificial tanto para desenvolvedores quanto para usuários de negócios, através de interfaces intuitivas e ferramentas de automação.

O compromisso da Microsoft com ética e transparência manifesta-se no desenvolvimento de práticas responsáveis, visando mitigar riscos de vieses e garantir segurança nas aplicações. Seu investimento em parcerias estratégicas e pesquisa avançada potencializa a adoção de IA em setores como saúde, educação, finanças e indústria, promovendo inovação e produtividade em escala global.

## IA da Amazon

A Amazon, por sua vez, construiu um ecossistema de Inteligência Artificial alicerçado na escalabilidade e na oferta de soluções amplamente acessíveis. Seu portfólio é liderado pelos modelos da família Amazon SageMaker, que permitem criar, treinar e implementar aplicações de IA de ponta em ambientes cloud, democratizando o acesso a ferramentas sofisticadas para cientistas de dados e equipes de desenvolvimento. Complementando essa oferta, os modelos como Amazon Bedrock tornam possível a utilização de fundações generativas, integrando diferentes arquiteturas e fornecedores para criar soluções personalizadas e interoperáveis.

Além do processamento de linguagem natural e geração de conteúdo, a Amazon investe em IA aplicada a comércio eletrônico, logística e segurança, com destaque para o Amazon Rekognition (análise de imagens e vídeos), Polly (conversão de texto em fala) e Comprehend (extração de insights de texto). A integração dessas tecnologias aos serviços da AWS favorece automação, hiperpersonalização e análise preditiva em larga escala.

Preocupada com a responsabilidade e a governança, a Amazon tem desenvolvido práticas para mitigar riscos de vieses e assegurar privacidade dos dados, promovendo um ambiente seguro para clientes dos mais variados setores. Sua abordagem flexível e modular permite que empresas adaptem as soluções de IA às suas próprias necessidades, acelerando o processo de transformação digital e impulsionando a hiperautomação em ambientes corporativos dinâmicos.

## IAs Chinesas

O cenário chinês de Inteligência Artificial destaca-se pela rápida evolução tecnológica, forte apoio estatal e pela presença de gigantes como Baidu, Alibaba e Tencent. Essas empresas lideram o desenvolvimento de modelos proprietários de IA impulsionando pesquisas em processamento de linguagem natural, visão computacional, reconhecimento facial e machine learning em larga escala.

Entre os modelos notáveis, o ERNIE, da Baidu, sendo amplamente utilizado em buscas, tradução automática e assistentes virtuais; O Alibaba DAMO Academy, que investe em modelos orientados para aplicações em comércio eletrônico, logística inteligente e serviços financeiros; e o Tencent aposta no Hunyuan, com foco em multimodalidade e integração de dados provenientes de diferentes fontes, atendendo setores como saúde, entretenimento e cidades inteligentes.

O ambiente regulatório chinês enfatiza tanto o avanço quanto o controle ético, impondo diretrizes claras para uso responsável e privacidade. A integração de IA às plataformas digitais e industriais do país tem acelerado a automação de processos, o desenvolvimento de cidades inteligentes e a transformação digital de empresas de todos os portes.

## DeepSeek: Inovação Chinesa em Inteligência Artificial Generativa

Entre os protagonistas desse cenário inovador, a China tem se destacado pela celeridade na pesquisa e desenvolvimento de modelos proprietários. Nesse contexto, DeepSeek emerge como um dos projetos mais inovadores e relevantes do país, chamando atenção global por sua capacidade de criar modelos de linguagem de larga escala, eficiência multimodal e aplicações práticas robustas.

DeepSeek é uma iniciativa chinesa voltada para o desenvolvimento de grandes modelos de linguagem, com o objetivo de fortalecer a posição da China no campo da IA generativa. O projeto é liderado por uma equipe multidisciplinar composta por engenheiros, cientistas de dados e pesquisadores especializados em processamento de linguagem natural, aprendizado de máquina e arquitetura computacional de alto desempenho.

A abordagem do DeepSeek destaca-se por sua ênfase em eficiência, escalabilidade e integração multimodal, permitindo que o modelo compreenda e gere conteúdo a partir de diferentes fontes de texto, imagens, vídeos e dados estruturados. Esse foco multimodal posiciona o DeepSeek como uma solução promissora para aplicações em setores como educação, saúde, e-commerce, mídia digital e cidades inteligentes.

## IAs Chinesas

A arquitetura do DeepSeek baseia-se em avanços recentes em redes neurais profundas, que revolucionaram a maneira como máquinas processam a linguagem humana. O modelo é treinado em conjuntos massivos de dados, abrangendo línguas orientais e ocidentais, códigos-fonte, imagens e registros multimídia. A utilização de técnicas de treinamento distribuído e otimização de parâmetros permite que o DeepSeek alcance resultados comparáveis aos modelos ocidentais de ponta, como GPT (OpenAI) e LLaMA (Meta).

**Código-fonte** é o conjunto de instruções em uma linguagem de programação que define o funcionamento de um software, como um programa de computador ou aplicativo. Ele é a "receita" que o computador segue para executar as tarefas, sendo a base para o desenvolvimento de sistemas e aplicativos.

O desenvolvimento do DeepSeek tem despertado crescente interesse entre profissionais, gestores públicos, pesquisadores e organizações que buscam compreender o papel estratégico dos modelos avançados de inteligência artificial em suas áreas de atuação. Em um cenário marcado pela transformação digital acelerada e pela demanda por soluções eficientes, transparentes e sustentáveis, tecnologias como o DeepSeek oferecem novas possibilidades para análise de dados, automação inteligente e apoio à tomada de decisão. Assim, compreender a arquitetura, os princípios de funcionamento e os mecanismos de governança desse modelo torna-se essencial para quem atua nos setores acadêmico, empresarial ou governamental, ampliando a capacidade de aplicar IA de forma responsável e alinhada às necessidades contemporâneas. A escalabilidade é outro pilar do projeto: o DeepSeek foi desenvolvido para operar em infraestrutura de computação de alta performance, aproveitando GPUs avançadas e clusters de servidores. Isso garante não apenas rapidez na inferência, mas também capacidade de adaptação a diferentes demandas empresariais e acadêmicas. Além disso, a equipe por trás do DeepSeek tem investido em pesquisa para reduzir o consumo energético do modelo durante treinamento e execução, promovendo sustentabilidade no uso de recursos computacionais.

O treinamento do DeepSeek utiliza uma variedade de fontes, desde textos acadêmicos até bases de dados industriais. O processo é meticulosamente controlado para minimizar vieses, garantir precisão e proporcionar capacidade de generalização. A preocupação com segurança e ética é central, sendo adotados mecanismos de filtragem, monitoramento e atualização constantes para evitar geração de conteúdos inadequados ou desinformação.

Além disso, o DeepSeek incorpora protocolos de privacidade para proteção de dados dos usuários, atendendo às exigências regulatórias chinesas e internacionais. O modelo é desenhado para ser transparente em suas decisões, facilitando auditorias e rastreabilidade de resultados.

O DeepSeek tem sido aplicado em múltiplos cenários, destacando-se especialmente em:

- **Assistentes virtuais:** Capazes de entender e gerar respostas precisas e contextuais, suportando múltiplos idiomas;
- **Educação personalizada:** Ferramentas de tutoria e correção automática, adaptadas ao perfil do estudante;
- **Saúde digital:** Apoio em triagem de sintomas, análise preditiva de prontuários e geração de relatórios clínicos;
- **E-commerce:** Motores de recomendação hiperpersonalizados, atendimento ao cliente e geração de descrições de produtos;
- **Análise de mídia:** Indexação automática de conteúdos, síntese de informações e apoio à criação multimídia;
- **Cidades inteligentes:** Monitoramento em tempo real, automação de processos administrativos e interação com cidadãos.

A flexibilidade do DeepSeek permite sua adaptação desde soluções lowcode/nocode até aplicações Hardcode, facilitando a integração em plataformas empresariais e sistemas públicos.

O contexto competitivo do DeepSeek inclui gigantes tecnológicos como Baidu (ERNIE), Tencent (Hunyuan) e Alibaba (DAMO Academy), todos empenhados em avançar o estado da arte em IA generativa. O diferencial do DeepSeek reside em sua abordagem colaborativa, buscando parcerias com universidades, startups e órgãos governamentais para expandir sua base de dados, explorar novas técnicas de treinamento e acelerar a adoção em setores estratégicos.

Apesar dos avanços, DeepSeek enfrenta desafios comuns a projetos de IA de grande escala. Entre eles, destacam-se:

- Mitigação de vieses nos dados, especialmente em contextos multiculturais;
- Gerenciamento do consumo energético e sustentabilidade computacional;
- Garantia de privacidade e conformidade regulatória internacional;
- Adaptação a mudanças rápidas no panorama tecnológico global.

A equipe do DeepSeek está continuamente investindo em pesquisa para desenvolver modelos mais compactos, eficientes e interpretáveis, além de explorar a integração com hardware especializado (chips de IA) e plataformas Edge Computing (modelo de computação que leva o processamento de dados para mais perto da fonte onde eles são gerados). A chegada do DeepSeek ao mercado global reforça a posição da China como protagonista na revolução da IA generativa. O modelo tem potencial para transformar a maneira como empresas, governos e cidadãos interagem com tecnologia, tornando processos mais inteligentes, ágeis e personalizados. A influência do DeepSeek extrapola fronteiras, promovendo intercâmbio científico e colaborativo entre o Oriente e o Ocidente.

Ao democratizar o acesso à IA, DeepSeek contribui para a redução de desigualdades tecnológicas, facilita o desenvolvimento de soluções inovadoras e estimula o surgimento de novos negócios baseados em dados. A preocupação ética e o compromisso com boas práticas consolidam a reputação do projeto, inspirando outras iniciativas e reforçando o papel da Inteligência Artificial na construção de uma sociedade mais eficiente e inclusiva.

O DeepSeek representa um marco na evolução da Inteligência Artificial chinesa, reunindo inovação, eficiência e responsabilidade em um único modelo. Sua arquitetura avançada, capacidade multimodal e escopo de aplicações revelam o potencial transformador da IA generativa no século XXI. À medida que o projeto continua a evoluir, espera-se que o DeepSeek desempenhe papel central na transformação digital de organizações, governos e comunidades, consolidando-se como referência internacional em soluções de IA de alto desempenho.

### QUER SABER MAIS?

Acesse <https://alvarofpinheiro.net/ia> e confira vídeo aulas com fundamentações e exemplos práticos que mostram, passo a passo, como aplicá-la.

# AULA VIRTUAL



## Módulo 3 - Aula 1 - Engenharia de Prompts



# MÓULO 4

## FERRAMENTAS DIGITAIS NA SAÚDE



## APRESENTAÇÃO

No cenário atual de transformação digital e integração crescente entre profissionais de saúde e tecnologias inteligentes, torna-se essencial compreender as ferramentas que viabilizam o desenvolvimento de soluções automatizadas voltadas à melhoria da assistência, da gestão e da vigilância em saúde. As modalidades Hardcode, Lowcode e Nocode têm revolucionado a forma como sistemas de Inteligência Artificial, mineração de dados e automação de processos robóticos (RPA) são concebidos, implementados e mantidos no setor.

Essas abordagens têm permitido que equipes multiprofissionais desde analistas de sistemas e engenheiros clínicos até gestores e profissionais de enfermagem participem ativamente da criação de soluções inovadoras sem depender exclusivamente de conhecimentos avançados em programação. Essa democratização tecnológica favorece o desenvolvimento ágil de ferramentas que automatizam tarefas repetitivas, como o envio de relatórios de vigilância epidemiológica, o monitoramento de indicadores hospitalares, o controle de estoques de insumos e medicamentos e a integração de dados entre unidades de saúde.

A escolha entre Hardcode, Lowcode e Nocode impacta diretamente a agilidade dos projetos, a eficiência operacional, os custos de implementação e o potencial de escalabilidade das aplicações em saúde. Plataformas Lowcode e Nocode, por exemplo, permitem que secretarias de saúde criem painéis automatizados de notificação de doenças, sistemas de triagem inteligente de pacientes e rotinas automáticas de comunicação entre municípios e o Ministério da Saúde, reduzindo o tempo de resposta e melhorando a coordenação das ações de vigilância.

A seguir, serão apresentadas as particularidades e benefícios de cada modalidade, evidenciando como essas abordagens se complementam e contribuem para o avanço da automação inteligente e da Inteligência Artificial aplicada à saúde pública e hospitalar.

## HARDCODE

A modalidade Hardcode representa a forma mais técnica e tradicional de desenvolvimento de soluções digitais, baseada no domínio direto da programação e na criação de códigos personalizados para cada necessidade. Essa abordagem exige do profissional conhecimento aprofundado em linguagens de programação, lógica computacional e ferramentas especializadas, permitindo o controle total sobre cada etapa do sistema.

Na área da saúde, o uso de Hardcode é essencial em situações que demandam precisão, segurança e personalização de alto nível. Por exemplo, desenvolvedores podem criar sistemas de análise preditiva de surtos epidemiológicos ajustando algoritmos para processar dados do e-SUS, ou ainda construir rotinas que cruzam informações de diferentes bancos de dados para identificar riscos de reinternação de pacientes. Também é comum utilizar essa abordagem para integrar sistemas laboratoriais, gerenciar dados clínicos sigilosos com segurança e implementar modelos de Inteligência Artificial que apoiam diagnósticos médicos complexos.

A principal vantagem do Hardcode está na flexibilidade e no controle absoluto sobre o funcionamento do sistema, permitindo criar soluções inovadoras do zero e adaptar cada detalhe às necessidades específicas dos serviços de saúde. Essa autonomia é decisiva em projetos que envolvem grandes volumes de dados sensíveis, protocolos clínicos personalizados ou processos críticos de tomada de decisão, como os usados em hospitais e secretarias de saúde.

Entretanto, essa abordagem também apresenta desafios. O desenvolvimento tende a demandar mais tempo e investimento, além de exigir profissionais altamente especializados. A manutenção de sistemas Hardcoded requer cuidado constante, já que modificações no código podem afetar múltiplas funcionalidades.

Apesar dessas limitações, o Hardcode continua sendo fundamental para o avanço da automação inteligente em saúde, servindo como base para a experimentação, a prototipação e a criação de sistemas complexos que outras modalidades, como Lowcode e Nocode, não conseguem absorver integralmente. Exemplos práticos de técnicas Hardcode incluem o uso de Python para automatizar a extração de dados de portais de vigilância sanitária, a análise de laudos laboratoriais com bibliotecas como Pandas e NumPy e a automação de registros e notificações em plataformas públicas por meio de ferramentas como Selenium (Saurkar; Pathare; Gode, 2018) e (Simon; Jrhuggin, 2022).

*Python é uma linguagem de programação de alto nível, versátil e de fácil aprendizado, usada para uma grande variedade de aplicações, como desenvolvimento web, ciência de dados, inteligência artificial, automação de tarefas e desenvolvimento de software em geral.*

*NumPy e Pandas são ferramentas essenciais para análise de dados em Python.*

No contexto da Inteligência Artificial aplicada à saúde, diferentes plataformas e ambientes de desenvolvimento têm papel fundamental na criação, teste e implantação de soluções que otimizam a gestão de informações e o apoio à decisão clínica. Cada ferramenta oferece recursos específicos para o desenvolvimento Hardcode, ou seja, com codificação direta e controle total sobre os processos de automação e aprendizado de máquina.

Apesar dessas limitações, o Hardcode continua sendo fundamental para o avanço da automação inteligente em saúde, servindo como base para a experimentação, a prototipação e a criação de sistemas complexos que outras modalidades, como Lowcode e Nocode, não conseguem absorver integralmente. Exemplos práticos de técnicas Hardcode incluem o uso de Python para automatizar a extração de dados de portais de vigilância sanitária, a análise de laudos laboratoriais com bibliotecas como Pandas e NumPy e a automação de registros e notificações em plataformas públicas por meio de ferramentas como Selenium (Saurkar; Pathare; Gode, 2018) e (Simon; Jrhuggin, 2022).

No contexto da Inteligência Artificial aplicada à saúde, diferentes plataformas e ambientes de desenvolvimento têm papel fundamental na criação, teste e implantação de soluções que otimizam a gestão de informações e o apoio à decisão clínica. Cada ferramenta oferece recursos específicos para o desenvolvimento Hardcode, ou seja, com codificação direta e controle total sobre os processos de automação e aprendizado de máquina.

Em nível corporativo, plataformas como Amazon SageMaker e Microsoft Azure Machine Learning podem ser amplamente utilizadas por secretarias de saúde e instituições de pesquisa que precisam de escalabilidade e infraestrutura avançada. O SageMaker permite o treinamento distribuído de modelos preditivos em larga escala, como o monitoramento do avanço de doenças infecciosas, enquanto o Azure Machine Learning facilita a criação de pipelines de previsão de demanda hospitalar e análise automatizada de dados de vacinação.

**Selenium** é um framework de código aberto para automação de navegadores, utilizado principalmente para testar aplicações web e realizar tarefas repetitivas na internet, como extração de dados.

**Pipeline** no contexto de tecnologia e Inteligência Artificial, são conjuntos organizados de etapas ou processos automatizados que transformam, processam e analisam dados de forma sequencial, desde a coleta inicial até a entrega de resultados ou insights. Em aplicações de saúde, um pipeline pode incluir a extração de dados de diferentes sistemas, limpeza e preparação das informações, aplicação de algoritmos de aprendizado de máquina e geração de relatórios para suporte à decisão clínica ou administrativa.

**TensorFlow** é uma biblioteca de código aberto criada pelo Google para computação numérica, aprendizado de máquina (ML) e aprendizado profundo (deep learning).

**PyTorch** é uma biblioteca de aprendizado de máquina usada para aplicações como visão computacional e processamento de linguagem natural, originalmente desenvolvida pela Meta AI. Ela é reconhecida como uma das duas bibliotecas de aprendizado de máquina mais populares, juntamente com a TensorFlow, oferecendo software livre e de código aberto.

**Script Python** é um arquivo de texto com código Python que é executado por um interpretador para automatizar tarefas. Serve como um "roteiro" para automatizar tarefas que seriam realizadas manualmente.

A Google Colab é uma das opções mais acessíveis e amplamente utilizadas por profissionais que iniciam no desenvolvimento de IA. Ele permite programar em Python e utilizar bibliotecas como TensorFlow, PyTorch, Scikit-learn, NumPy e Pandas, possibilitando a criação de modelos capazes de analisar grandes conjuntos de dados de saúde. Em um contexto prático, pode ser utilizado para treinar modelos que identifiquem padrões de aumento de casos de dengue a partir de notificações municipais, ou para analisar exames laboratoriais e detectar alterações precoces em indicadores clínicos.

O Jupyter Notebook oferece uma estrutura semelhante, mas voltada à execução local ou em servidores institucionais. São amplamente usados em pesquisas que envolvem dados clínicos sensíveis, por permitirem maior controle de segurança e personalização. Profissionais de ciência de dados em hospitais e universidades utilizam esses ambientes para desenvolver algoritmos de previsão de leitos disponíveis, analisar séries temporais de internações e modelar fluxos de atendimento em unidades de emergência.

Já o PyCharm destaca-se como um ambiente robusto voltado a projetos de Inteligência Artificial de maior porte. Suas ferramentas de depuração e integração com frameworks como TensorFlow e PyTorch o tornam adequado para projetos hospitalares que envolvem redes neurais profundas, como sistemas de recomendação de condutas clínicas, identificação automatizada de lesões em imagens médicas e análise de risco de complicações pós-operatórias.

O VS Code, por sua vez, é um editor de código leve, mas poderoso, que integra notebooks e scripts Python de forma prática. Ele é bastante utilizado em projetos que envolvem automação de relatórios hospitalares, padronização de dados de vigilância epidemiológica e avaliação de desempenho de modelos de IA aplicados à gestão de estoques de medicamentos ou monitoramento de indicadores assistenciais.

Essas ferramentas exemplificam o potencial do desenvolvimento Hardcode no contexto da saúde, permitindo que equipes técnicas criem soluções personalizadas, seguras e eficazes para desafios reais, desde a vigilância epidemiológica até a automação de processos clínicos e administrativos.

### QUER SABER MAIS?

Acesse <https://alvarofpinheiro.net/python> e confira vídeo aulas com fundamentações e exemplos práticos que mostram, passo a passo, como aplicá-la.

## LOWCODE

No campo da Inteligência Artificial aplicada à saúde, nem todas as soluções exigem conhecimentos avançados em programação ou arquitetura de sistemas. Nos últimos anos, as plataformas Lowcode têm se destacado por democratizar o desenvolvimento de ferramentas inteligentes, permitindo que profissionais de diferentes formações, gestores, analistas, médicos, enfermeiros e técnicos participem diretamente da criação de sistemas e automações voltados à melhoria da assistência e da gestão em saúde.

Essas plataformas reduzem barreiras técnicas e permitem orquestrar automações, integrar APIs (Interface de Programação de Aplicativos) de IA e construir fluxos de trabalho complexos por meio de interfaces visuais e comandos simples. Na prática, isso significa que uma equipe de vigilância epidemiológica pode automatizar o cruzamento de dados de notificações de doenças, gerar painéis dinâmicos de casos suspeitos e confirmados, ou configurar alertas automáticos de aumento de incidência em determinados municípios, sem precisar escrever códigos extensos.

Ferramentas Lowcode também são úteis para otimizar tarefas administrativas e de monitoramento, como controlar a disponibilidade de leitos hospitalares em tempo real, emitir relatórios de atendimentos de urgência e integrar informações de sistemas estaduais e federais de saúde. Essas soluções aumentam a eficiência operacional e reduzem o tempo gasto com processos repetitivos, permitindo que as equipes técnicas concentrem esforços em análises e decisões estratégicas.

A seguir, serão apresentadas algumas das principais plataformas Lowcode utilizadas para automatizar processos e desenvolver aplicações inteligentes na saúde pública e hospitalar, destacando como elas facilitam a implementação de soluções baseadas em IA e dados:

- **Appian:** solução corporativa que combina automação e mineração de processos, muito útil para mapear gargalos na gestão hospitalar e analisar fluxos de autorizações e faturamento. Pode ser utilizada para melhorar a governança e a transparência nos processos administrativos de secretarias de saúde.
- **Automation Anywhere:** plataforma de automação empresarial com recursos de IA para interpretação de textos e documentos médicos, extração automática de informações de sistemas hospitalares e respostas automáticas a pacientes via bots. Pode ser usada, por exemplo, no monitoramento de solicitações de exames ou no encaminhamento de resultados laboratoriais para unidades básicas de saúde.
- **BigML:** plataforma voltada ao aprendizado de máquina automático (AutoML), com foco em modelagem de dados de forma simples e explicável. Pode ser aplicada à previsão de surtos sazonais, estimativa de demanda por leitos hospitalares ou análise de séries temporais de casos de doenças transmissíveis, fornecendo informações estratégicas para tomada de decisão.

- **KNIME:** ferramenta de ciência de dados Lowcode com interface totalmente visual. Pode ser usada para detectar anomalias em bases de dados hospitalares, identificar inconsistências em registros de vacinação ou analisar padrões de uso de medicamentos. Sua abordagem modular facilita a integração de diferentes fontes de dados sem necessidade de codificação.
- **Make:** plataforma visual para construir cenários de automação e análise de dados. Pode ser usada para monitorar menções a doenças em redes sociais, analisar sentimentos sobre serviços de saúde e gerar dashboards com indicadores em tempo real, auxiliando gestores na detecção precoce de surtos ou crises sanitárias.
- **Microsoft Power Automate:** integrada ao ecossistema Microsoft 365, é ideal para secretarias e hospitais que utilizam Outlook, SharePoint e Teams. Possibilita automatizar a triagem de documentos médicos com OCR, analisar relatórios de exames com IA Builder e enviar dados automaticamente para bancos de dados ou painéis de gestão, otimizando fluxos de trabalho e reduzindo burocracias.

**Dashboard:** Visualização consolidada de dados em gráficos e indicadores, permitindo acompanhar métricas e apoiar decisões rapidamente.

**OCR (Optical Character Recognition):** Tecnologia que converte textos presentes em imagens ou documentos digitalizados em texto editável e pesquisável.

**Open Source:** Modelo de software cujo código-fonte é aberto e pode ser usado, estudado, modificado e distribuído livremente.

**SaaS (Software as a Service):** Modelo de entrega de software em que a aplicação é acessada pela internet como um serviço, sem necessidade de instalação local, com atualização, manutenção e infraestrutura geridas pelo fornecedor.

- **Orange Canvas:** ferramenta open source voltada à análise de dados e machine learning por meio de uma interface gráfica. É possível construir modelos de classificação e regressão arrastando blocos visuais, o que facilita o uso por equipes de vigilância ou pesquisadores. Pode ser usada para identificar padrões clínicos em exames laboratoriais, analisar fatores de risco de doenças crônicas ou treinar modelos preditivos de diagnóstico supervisionado.
- **UiPath StudioX:** versão simplificada da plataforma de automação robótica UiPath, voltada a usuários não técnicos. Permite a criação de robôs para automatizar processos administrativos, como o processamento de notas fiscais hospitalares, agendamento de consultas e organização de autorizações médicas, integrando diferentes sistemas de forma segura e eficiente.
- **Zapier:** solução SaaS amplamente utilizada para integrações rápidas entre aplicativos. Na saúde, pode automatizar tarefas como enviar resumos de e-mails de vigilância epidemiológica para uma equipe de regulação ou alimentar chatbots de atendimento ao cidadão com informações de campanhas de vacinação, melhorando a comunicação entre unidades de saúde e a população.

Essas plataformas demonstram como as abordagens Lowcode ampliam a capacidade das equipes de saúde em criar, integrar e automatizar soluções inteligentes, tornando os processos mais ágeis, seguros e baseados em evidências. O uso dessas ferramentas tem contribuído para reduzir o tempo de resposta em emergências sanitárias, otimizar fluxos administrativos e melhorar a qualidade das informações utilizadas na gestão pública da saúde.

## NOCODE

A ascensão das plataformas Nocode representa um avanço importante na democratização do desenvolvimento de soluções digitais em saúde. Essas ferramentas permitem que profissionais sem formação em programação criem aplicações, automatizem processos e explorem recursos de Inteligência Artificial sem precisar escrever código. Essa acessibilidade amplia o alcance da inovação, acelera o desenvolvimento de soluções e libera as equipes para concentrar esforços em atividades estratégicas voltadas à melhoria da assistência e da gestão dos serviços de saúde.

No contexto do sistema público de saúde, as plataformas Nocode têm sido utilizadas para resolver desafios práticos, como o gerenciamento automatizado de planilhas de notificação de doenças, o envio de relatórios epidemiológicos e a criação de painéis interativos para o acompanhamento de indicadores. Por exemplo, um profissional de vigilância epidemiológica pode configurar um fluxo automatizado que compara semanalmente os casos de sarampo entre municípios, gera alertas em caso de aumento significativo e envia relatórios atualizados para as coordenações estaduais, tudo em um ambiente visual e intuitivo.

Essas ferramentas também auxiliam hospitais e unidades de saúde a monitorar estoques de medicamentos, gerar relatórios de atendimento ambulatorial e automatizar a coleta de dados sobre internações. A combinação entre automação e visualização de dados permite que gestores compreendam rapidamente o desempenho das unidades, identifiquem gargalos e tomem decisões baseadas em informações atualizadas.

A visualização de dados nas plataformas Nocode pode ser comparada a um modelo de blocos de montagem, semelhante a um conjunto de peças de LEGO, em que cada bloco representa uma etapa do processo: coleta, processamento e apresentação das informações. Essa metáfora ajuda a compreender como é possível construir soluções de forma simples e modular, conectando cada parte para formar um todo coerente e funcional.

No contexto da hiperautomação em saúde, o Nocode tem papel essencial ao eliminar barreiras técnicas e facilitar a criação de fluxos automatizados que integram Inteligência Artificial, análise de dados e automação de tarefas. Ao tornar as ferramentas digitais mais intuitivas e colaborativas, essas plataformas ampliam a eficiência dos serviços e fortalecem a capacidade de resposta e tomada de decisão em todos os níveis do sistema de saúde.

As plataformas Nocode têm ampliado o acesso ao uso da Inteligência Artificial na área da saúde, permitindo que profissionais sem formação em programação criem soluções inovadoras para problemas reais do dia a dia. Essas ferramentas possibilitam o treinamento de modelos, a automação de processos e a análise de dados de forma simples, visual e interativa. A seguir, são apresentadas algumas das principais plataformas e como elas podem ser aplicadas no contexto da saúde pública e hospitalar.

- **Azure AutoML:** permite gerar modelos preditivos de forma automática, sem exigir conhecimentos técnicos em programação. Ele pode ser utilizado para prever a demanda por leitos hospitalares, analisar a evolução de casos de doenças sazonais, identificar fatores de risco em pacientes crônicos e gerar análises explicáveis para apoio à decisão médica. Sua interface intuitiva e a integração com ferramentas do ecossistema Microsoft, como Power BI e Excel, tornam o uso mais acessível para equipes de gestão e pesquisa.
- **Google AutoML (Vertex):** é um serviço da plataforma Google Cloud que permite criar modelos de aprendizado de máquina para análise de texto, imagens e dados tabulares de maneira automatizada. Na saúde, pode ser aplicado na classificação automática de exames de imagem, na análise de registros eletrônicos de pacientes e na detecção de padrões em notificações epidemiológicas. Sua integração com as APIs do Google facilita o processamento de grandes volumes de dados hospitalares e a criação de relatórios preditivos.
- **Lovable:** é uma ferramenta que permite criar aplicativos e automatizar fluxos de trabalho de maneira colaborativa e intuitiva. No setor de saúde, pode ser utilizada para criar formulários inteligentes de triagem de pacientes, gerenciar fluxos de atendimento em unidades básicas de saúde e automatizar solicitações de relatórios entre secretarias e hospitais. Sua interface visual facilita o trabalho conjunto de profissionais de tecnologia, gestores e equipes assistenciais.
- **Teachable Machine (Google):** permite treinar modelos de reconhecimento de imagens, sons e movimentos sem precisar programar. Por meio de uma interface visual, o usuário pode capturar exemplos e ensinar o modelo a reconhecer padrões. Em ambientes de saúde, pode ser utilizada para criar sistemas que identificam gestos de pacientes durante fisioterapia, reconhecem a presença de equipamentos de proteção individual ou classificam imagens de exames em categorias visuais básicas, servindo como introdução prática à visão computacional.

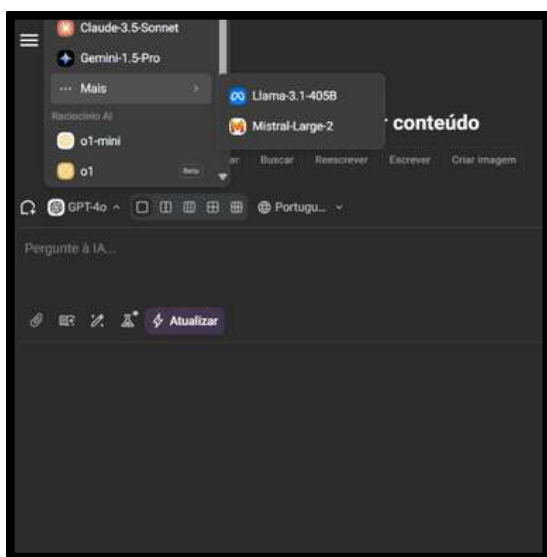
A visualização de dados em plataformas Nocode pode ser comparada a um modelo de blocos LEGO, em que cada peça representa uma etapa do processo: coleta, tratamento, análise e exibição das informações. Ao encaixar esses blocos, o profissional constrói soluções personalizadas, como painéis interativos de vigilância epidemiológica, mapas de calor com distribuição de casos ou gráficos dinâmicos de indicadores hospitalares, sem precisar escrever uma linha de código. Essa abordagem torna o aprendizado mais lúdico e facilita a compreensão dos conceitos de ciência de dados aplicados à saúde.

Essas plataformas demonstram como o uso de ferramentas Nocode pode tornar a Inteligência Artificial mais acessível, prática e transformadora para profissionais da saúde, contribuindo para uma gestão mais eficiente, diagnósticos mais precisos e decisões baseadas em evidências.

## QUER SABER MAIS?

Acesse <https://alvarofpinheiro.net/ferramentas> e confira vídeo aulas com fundamentações e exemplos práticos que mostram, passo a passo, como aplicá-la.

## EXEMPLOS DE FERRAMENTAS



<https://www.maxai.me/>



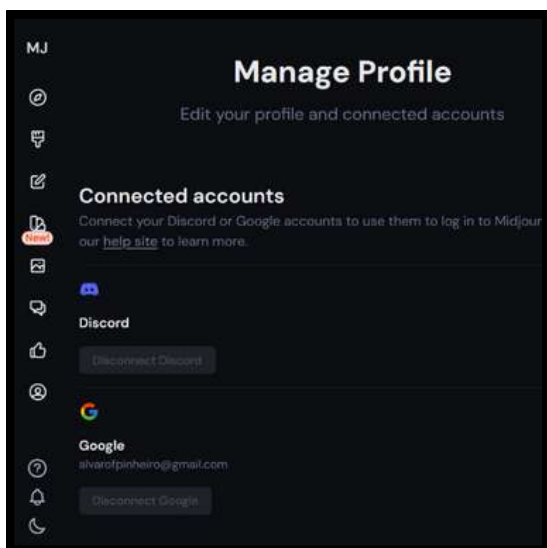
A plataforma MaxAI tem como objetivo principal centralizar e integrar diferentes tecnologias de Inteligência Artificial em um único ambiente, funcionando como um hub (plataforma de IAs) que facilita a automação de processos e o aumento da produtividade. No contexto da saúde, ela pode ser utilizada para integrar sistemas de gestão hospitalar, automatizar tarefas administrativas repetitivas, analisar grandes volumes de dados clínicos e epidemiológicos e gerar relatórios inteligentes que apoiam a tomada de decisão.

Por exemplo, uma secretaria de saúde pode utilizar a MaxAI para reunir dados de notificações de doenças, estoques de medicamentos e registros de atendimentos, transformando essas informações em painéis visuais e dinâmicos que permitem acompanhar a evolução de indicadores em tempo real. Em um hospital, a ferramenta pode auxiliar na análise de dados de internações, na previsão de demanda por leitos e na geração automática de relatórios gerenciais e clínicos, promovendo eficiência e reduzindo o tempo gasto com atividades manuais.

A MaxAI também pode ser aplicada para criar conteúdo automatizados, como relatórios semanais de vigilância epidemiológica, resumos executivos para gestores de saúde ou comunicações personalizadas para equipes e pacientes. Ao combinar a análise de dados estruturados com a geração de textos e visualizações, a ferramenta possibilita que profissionais compreendam melhor os resultados e tomem decisões baseadas em evidências.

A visualização de dados produzida pela MaxAI pode ser entendida de maneira lúdica por meio da analogia com blocos de LEGO, onde cada bloco representa uma etapa do processo de análise: coleta, integração, processamento e apresentação das informações. Assim como na montagem dos blocos, o profissional de saúde pode construir suas próprias soluções, encaixando as partes que formam um panorama completo da realidade assistencial e gerencial.

Ao tornar o uso da Inteligência Artificial mais acessível e visual, a MaxAI contribui para otimizar fluxos de trabalho, melhorar a comunicação entre setores e fortalecer a capacidade analítica das equipes de saúde, promovendo uma gestão mais inteligente, ágil e orientada por dados.



<https://www.midjourney.com/>

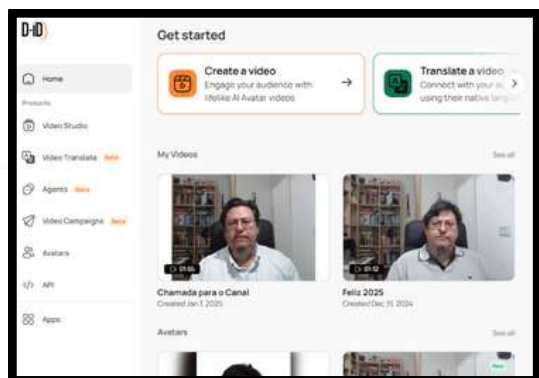


A ferramenta Midjourney é uma plataforma de geração de imagens que permite criar representações visuais de alta qualidade, com resolução de até 2048 por 2048 pixels, controle detalhado de variações e capacidade de ampliação sem perda de nitidez. Sua principal contribuição está em transformar informações textuais em imagens realistas e precisas, o que torna possível representar conceitos complexos de forma visual e intuitiva.

No contexto da saúde, o uso do Midjourney pode auxiliar na criação de materiais educativos, visualizações de dados clínicos e ilustrações de processos biomédicos e hospitalares. Por exemplo, pode ser utilizado para gerar representações gráficas de fluxos de atendimento em unidades de saúde, mapas visuais de distribuição de doenças ou diagramas explicativos sobre o funcionamento de modelos de Inteligência Artificial aplicados à triagem de pacientes. Essas imagens ajudam profissionais e estudantes a compreenderem informações técnicas de maneira mais clara e acessível.

A ferramenta também pode ser aplicada na produção de gráficos explicativos e dashboards visuais, integrando informações de vigilância epidemiológica, indicadores de desempenho hospitalar e dados de gestão pública. Essa visualização, quando combinada com técnicas de design inspiradas em modelos de blocos LEGO, permite que os dados sejam organizados de forma lúdica e modular, em que cada “bloco” representa uma categoria de informação, como casos confirmados, atendimentos realizados ou disponibilidade de recursos.

Essa abordagem visual facilita a interpretação de grandes volumes de dados e apoia a tomada de decisão em todos os níveis da gestão em saúde. Ao tornar os conceitos mais palpáveis e as informações mais acessíveis, o Midjourney contribui para aproximar a tecnologia da prática profissional, fortalecendo o uso da Inteligência Artificial como ferramenta de apoio à análise, comunicação e inovação no setor da saúde.



<https://studio.d-id.com/>

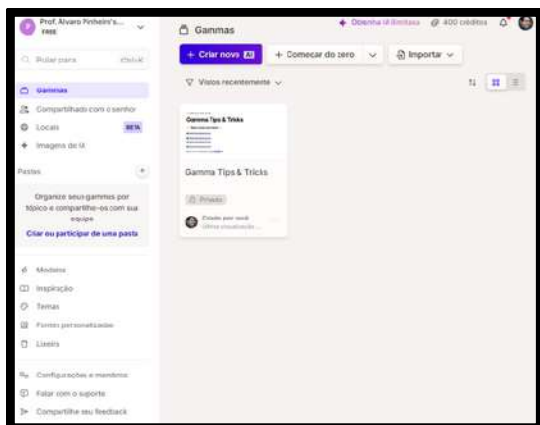


A plataforma D-ID permite transformar textos em vídeos de maneira simples e rápida, utilizando avatares realistas que falam em diferentes idiomas, inclusive português. Essa tecnologia converte informações escritas em apresentações visuais narradas, tornando a comunicação mais acessível, dinâmica e envolvente.

No campo da saúde, pode ser utilizada para produzir vídeos educativos sobre campanhas de vacinação, explicar protocolos de prevenção de doenças ou orientar a população sobre o uso correto de medicamentos. Em hospitais e universidades, pode ser empregada na formação de profissionais, criando vídeos explicativos sobre procedimentos clínicos, biossegurança ou uso de sistemas de Inteligência Artificial. Essa abordagem facilita o aprendizado e amplia o alcance das informações, especialmente em equipes distribuídas geograficamente.

Além disso, a D-ID pode auxiliar na comunicação de dados e indicadores de saúde de forma visual, substituindo relatórios extensos por vídeos que mostram resultados e gráficos animados. Essa forma de apresentação pode ser compreendida como uma “visualização viva de dados”, na qual cada componente do vídeo funciona como um bloco de LEGO, representando uma parte da informação: um bloco para os números de casos, outro para a cobertura vacinal e outro para a disponibilidade de recursos. Ao “montar” esses blocos, o profissional constrói uma narrativa visual que facilita a compreensão de cenários complexos.

O uso de ferramentas como a D-ID contribui para tornar a análise e a comunicação em saúde mais humanas, acessíveis e interativas. Ao transformar dados e textos em conteúdos visuais e sonoros, ela aproxima gestores, profissionais e cidadãos, promovendo uma cultura de aprendizado contínuo e comunicação eficaz apoiada pela Inteligência Artificial.



<https://gamma.app/>

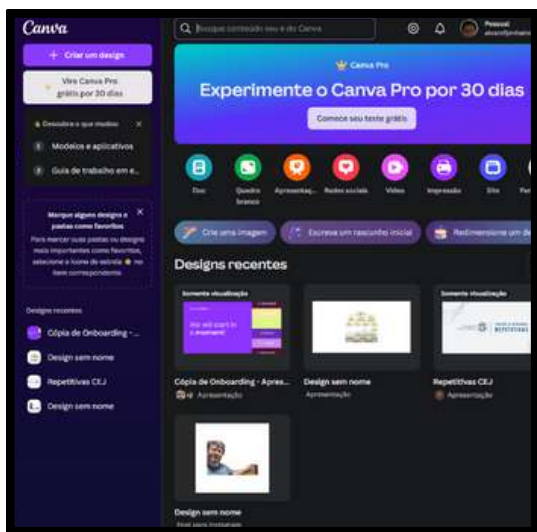


A plataforma Gamma é uma ferramenta que permite criar apresentações completas de forma automática a partir de textos, imagens ou documentos importados do Word e do Google Docs. Com poucos comandos, é possível transformar relatórios técnicos, dados brutos ou materiais acadêmicos em apresentações visuais dinâmicas e interativas.

Na área da saúde, o Gamma pode ser utilizado para apresentar resultados de análises epidemiológicas, divulgar relatórios de desempenho hospitalar ou demonstrar comparativos de indicadores de gestão e vigilância. Por exemplo, um profissional de vigilância pode importar um relatório de casos de dengue e gerar uma apresentação que mostre a distribuição geográfica dos casos, as taxas de incidência por município e as tendências semanais, facilitando a interpretação dos dados por gestores e equipes técnicas.

A ferramenta também pode ser aplicada em processos formativos e campanhas, permitindo a criação de slides explicativos sobre temas como prevenção de infecções respiratórias, hábitos saudáveis ou protocolos de segurança hospitalar, tornando o conteúdo mais atrativo e acessível para diferentes públicos.

A visualização de dados no Gamma pode ser representada de forma lúdica, semelhante a um modelo de blocos LEGO, em que cada bloco representa uma parte da informação: um bloco para os dados de entrada, outro para a análise e outro para a visualização final. Ao combinar esses blocos, o profissional constrói apresentações coerentes e interativas, transformando números e relatórios em histórias visuais que facilitam a compreensão e a tomada de decisão. Logo, o Gamma contribui para o fortalecimento da cultura de dados, estimulando o uso da Inteligência Artificial como ferramenta de apoio à gestão, à pesquisa e à educação em saúde.



<https://www.canva.com/>

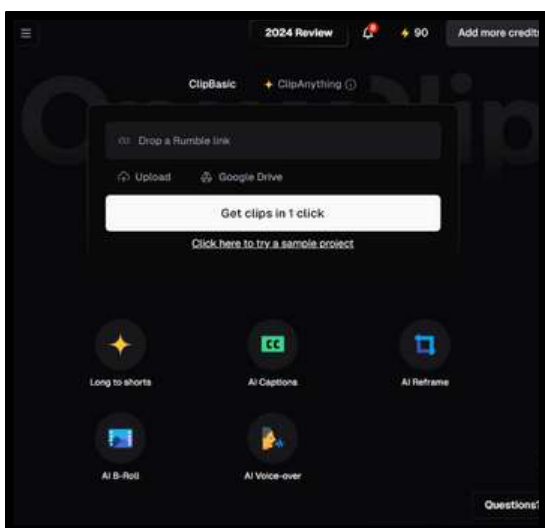


A ferramenta Canva é amplamente utilizada para criar apresentações, infográficos e materiais visuais de maneira simples e intuitiva. Por meio da opção “Docs” e do recurso “De Docs a Apresentações”, é possível transformar relatórios técnicos, textos ou documentos em apresentações visuais completas, sem necessidade de conhecimentos de design. Essa funcionalidade facilita a comunicação de informações complexas de forma clara, criativa e acessível.

Na área da saúde, o Canva pode ser utilizado para elaborar relatórios visuais de indicadores epidemiológicos, divulgar campanhas de prevenção, criar materiais educativos para pacientes e apresentar resultados de projetos de pesquisa. Por exemplo, um gestor pode importar um relatório com dados de casos de dengue e gerar uma apresentação que destaque os gráficos de incidência semanal, os mapas de distribuição dos casos por município e as ações de resposta adotadas, tornando o conteúdo mais compreensível para equipes técnicas e decisores.

O Canva também pode apoiar a visualização de dados clínicos e operacionais, permitindo a criação de painéis ilustrativos com informações sobre taxas de ocupação hospitalar, tempo médio de atendimento e desempenho de unidades de saúde. Essa abordagem estimula a criatividade dos profissionais de saúde e facilita a construção de narrativas visuais que transformam relatórios técnicos em materiais mais dinâmicos e estratégicos, facilitando a compreensão.

Ao promover a integração entre design, dados e Inteligência Artificial, o Canva se torna uma ferramenta essencial para profissionais de saúde que desejam comunicar resultados de forma eficiente, visual e acessível, fortalecendo a cultura de dados e a gestão estratégica.

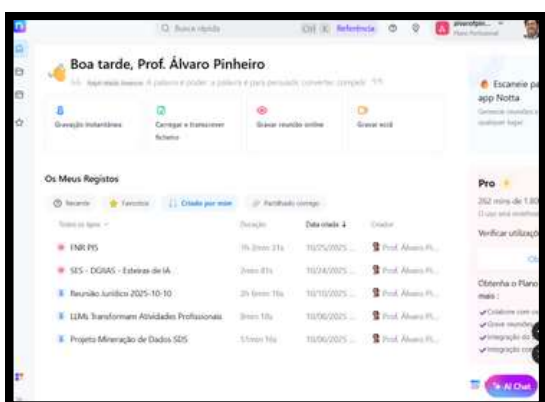



<https://www.opus.pro/> 

O OPUS permite analisar vídeos longos de plataformas como YouTube, Vimeo e Google Drive, gerando automaticamente a transcrição, selecionando os trechos mais relevantes e criando vídeos curtos com legendas para publicação em redes sociais como Instagram Reels, TikTok e YouTube Shorts.

No contexto da saúde, essa tecnologia pode ser usada para transformar longas gravações de videoaulas, palestras ou reuniões técnicas em vídeos curtos e informativos, tornando o conteúdo mais acessível e atrativo. Por exemplo, um gestor em saúde pode utilizá-la para resumir webinários sobre campanhas de prevenção, reuniões de monitoramento epidemiológico, gerando pequenos vídeos educativos.

Além disso, a ferramenta pode auxiliar na visualização de dados em formato audiovisual, convertendo relatórios e apresentações em narrativas visuais curtas. Um exemplo seria criar vídeos com gráficos animados mostrando a evolução de casos de doenças respiratórias, resultados de programas de prevenção ou indicadores de desempenho de profissionais de saúde, permitindo a tomada de decisões estratégicas entre gestores e profissionais, de acordo com as informações mais importantes. O uso dessa ferramenta contribui para fortalecer a comunicação em saúde, divulgar informações de maneira mais atraente e acessível e estimular a cultura de dados visuais, tornando o aprendizado e a gestão mais dinâmicos e baseados em evidências.



<https://app.notta.ai/> 

Notta.AI é uma ferramenta de transcrição automática baseada em inteligência artificial que converte áudio e vídeo em texto. A plataforma utiliza tecnologia de reconhecimento de voz para transcrever reuniões, entrevistas, palestras e outras conversas em tempo real ou a partir de arquivos gravados.

Principais Funcionalidades: Transcrição em tempo real de reuniões e conversas; Integração com plataformas como Zoom, Google Meet e Microsoft Teams; Suporte multilíngue (incluindo português); Resumos automáticos gerados por IA; Identificação de falantes diferentes; Exportação em diversos formatos (TXT, PDF, DOCX); Busca dentro das transcrições.

Aqui estão alguns exemplos de uso na Área da Saúde:

- Consultas Médicas, o médico pode gravar a consulta com o paciente e obter uma transcrição completa para documentação do prontuário, capturando detalhes importantes sobre sintomas, histórico e orientações fornecidas sem precisar digitar durante o atendimento;
- Telemedicina, durante consultas virtuais, o Notta pode integrar-se às plataformas de videoconferência para transcrever automaticamente a conversa, facilitando o registro preciso das informações compartilhadas remotamente;
- Entrevistas com Pacientes para Anamnese, captura detalhada do histórico médico do paciente, sintomas relatados e queixas principais, garantindo que nenhuma informação importante seja perdida durante a coleta de dados;
- Pesquisa Clínic, transcrição de entrevistas qualitativas com pacientes ou grupos focais, facilitando a análise de dados em estudos sobre experiência do paciente, adesão ao tratamento e satisfação com serviços de saúde; Reuniões Multidisciplinares, documentação de reuniões entre diferentes profissionais de saúde (médicos, enfermeiros, fisioterapeutas, psicólogos) para planejamento integrado do cuidado ao paciente.

Alguns benefícios para Profissionais de Saúde são: Economia no tempo de escrita de documentação; Maior atenção ao paciente durante a consulta; Redução de erros na documentação médica; Facilita auditorias e revisões de casos; Melhora a continuidade do cuidado através de registros mais completos.

## IMPORTANTE

*Ao utilizar ferramentas de transcrição automática em todos os cenários exemplificados como consultas médicas presenciais ou por telemedicina, entrevistas de anamnese, pesquisas clínicas, reuniões multidisciplinares e elaboração de vídeos educativos é imprescindível adotar rigorosos cuidados éticos e legais. É obrigatório obter o consentimento do paciente, garantir a conformidade com a LGPD, respeitar o sigilo médico e assegurar a confidencialidade dos dados em todas as etapas do processo. A atenção a esses princípios protege os direitos dos envolvidos e mantém a integridade das informações sensíveis no ambiente da saúde.*

## AULA VIRTUAL



### Módulo 4 - Aula 1 - Ferramentas Digitais na Saúde



## REFERÊNCIAS

AGRAWAL, R.; SRIKANT, R. **Fast algorithms for mining association rules**. In: INTERNATIONAL CONFERENCE ON VERY LARGE DATA BASES, 20., 1994, Santiago. Proceedings. Santiago, Chile: Morgan Kaufmann, 1994. p. 487-499.

AIMDEK TECHNOLOGIES. **Evolution of Robotic Process Automation (RPA): The Path to Cognitive RPA**. [S. l.]: Medium, 2018. Disponível em: <https://medium.com/@AIMDekTech/evolution-of-robotic-process-automation-the-path-to-cognitive-rpac3bd52c8b865>. Acesso em: 18 set. 2025.

ALIZADEH, A. A.; EISEN, M. B.; DAVIS, R. E.; MA, C.; LOSSOS, I. S.; ROSENWALD, A.; BOLDRICK, J. C.; SABET, H.; TRAN, T.; YU, X.; POWELL, J. I.; YANG, L.; MARTI, G. E.; MOORE, T.; HUDSON JR, J.; LU, L.; LEWIS, D. B.; TIBSHIRANI, R.; SHERLOCK, G.; CHAN, W. C.; GREINER, T. C.; WEISENBURGER, D. D.; ARMITAGE, J. O.; WARNKE, R.; LEVY, R.; WILSON, W.; GREVER, M. R.; BYRD, J. C.; BOTSTEIN, D.; BROWN, P. O.; STAUDT, L. M. **Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling**. Nature, London, v. 403, n. 6770, p. 503-511, fev. 2000.

ARNTZ, M.; GREGORY, T.; ZIERAHN, U. **The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis**. OECD Social, Employment and Migration Working Papers, n. 189, 2016. Disponível em: <https://doi.org/10.1787/5jlz9h56dvq7-en>. Acesso em: 3 dez. 2025.

BAMBERGER, P. How **Automation Is Shaping the Future**. [S. l.]: Medium, 2021. Disponível em: <https://phillipbamberger.medium.com/how-automation-is-changingour-world-8784ce159393>. Acesso em: 18 set. 2025.

BANCO INTERAMERICANO DE DESENVOLVIMENTO (BID). **El futuro del trabajo en América Latina y el Caribe: ¿Cómo será el mercado laboral para los trabajadores?** Washington, DC: BID, 2020. Disponível em: <https://publications.iadb.org>. Acesso em: 3 dez. 2025.

BISHOP, C. M. **Pattern Recognition and Machine Learning**. New York: Springer, 2006.

BROWN, T. B.; MANN, B.; RYDER, N.; SUBBIAH, M.; KAPLAN, J.; DHARIWAL, P.; NEELAKANTAN, A.; SHYAM, P.; SASTRY, G.; ASKELL, A.; AGARWAL, S.; HERBERT-VOSS, A.; KRUEGER, G.; HENIGHAN, T.; CHILD, R.; RAMESH, A.; ZIEGLER, D. M.; WU, J.; WINTER, C.; HESSE, C.; CHEN, M.; SIGLER, E.; LITWIN, M.; GRAY, S.; CHESS, B.; CLARK, J.; BERNER, C.; MCCANDLISH, S.; RADFORD, A.; SUTSKEVER, I.; AMODEI, D. **Language Models are Few-Shot Learners**. Advances in Neural Information Processing Systems, [S. l.], v. 33, p. 1877-1901, 2020.

CHEPLYGINA, V.; BRUIJNE, M. de; PLUIM, J. P. W. **Not-so-supervised: a survey of semi-supervised, multi-instance, and transfer learning in medical image analysis**. Medical Image Analysis, [S. l.], v. 54, p. 280-296, 2019.

DA SILVA, A. S.; PINHEIRO, Á. F.; MASSA, R. **Legal Decision-Making in Algorithmic Society: Observations on Techno-Animism and Communicactivation**. Springer Nature, 2025. Disponível em: [https://link.springer.com/chapter/10.1007/978-3-031-88209-8\\_7](https://link.springer.com/chapter/10.1007/978-3-031-88209-8_7). Acesso em: 18 set. 2025.

GARTNER. **Top Strategic Technology Trends for 2020**. [S. l.]: Gartner, 2020.

GÜL, O.; TUNA, I.; GÜLEÇ, D. **RPA IMPLEMENTATION in Accounting Firms**. Procedia Computer Science, [s. l.], v. 192, p. 4672-4680, 2021. Disponível em: <https://doi.org/10.1016/j.procs.2021.09.245>. Acesso em: 18 set. 2025.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA (IPEA). **Automação e o futuro do trabalho no Brasil**. Brasília: IPEA, 2021. Disponível em: <https://www.ipea.gov.br>. Acesso em: 3 dez. 2025.

KOMOROWSKI, M.; CELI, L. A.; BADAWI, O.; GORDON, A. C.; FAISAL, A. A. **The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care**. Nature Medicine, v. 24, n. 11, p. 1716-1720, 2018.

LACITY, M. C.; WILLCOCKS, L. **What Knowledge Workers Stand to Gain from Automation**. Harvard Business Review, 2015. Disponível em: <https://hbr.org/2015/06/what-knowledge-workers-stand-to-gain-from-automation>. Acesso em: 18 set. 2025.

LEE, K-F. **Inteligência Artificial: como os robôs irão mudar o mundo, a forma como trabalhamos e quem somos**. Rio de Janeiro: Globo Livros, 2019.

LITJENS, G.; KOOI, T.; BEJNORDI, B. E.; SETIO, A. A. A.; CIOMPI, F.; GHAFLOORIAN, M.; LAAN, J. A. W. M. van der; GINNEKEN, B. van; SÁNCHEZ, C. I. **A survey on deep learning in medical image analysis**. Medical Image Analysis, [S. l.], v. 42, p. 60-88, 2017.

LIU, S. **Size of the robotic process automation (RPA) market worldwide from 2020 to 2030**. [S. l.]: Statista, 2022. Disponível em: <https://www.statista.com/statistics/1089396/robotic-process-automation-market-size-worldwide/>. Acesso em: 18 set. 2025.

LUZ, P. M.; BEZERRA, V. C. C.; GABAGLIA, A. P. L. T.; ALBUQUERQUE, C. C. M. M.; STRUCHINER, C. J.; TEIXEIRA, M. G. L. C. **Time series analysis of dengue incidence in Rio de Janeiro, Brazil**. American Journal of Tropical Medicine and Hygiene, [S. l.], v. 79, n. 6, p. 933-939, 2008.

MADDOX, T. **Top 10 technology trends for 2020 include hyperautomation, human augmentation and distributed cloud**. TechRepublic, 2019. Disponível em: <https://www.techrepublic.com/article/hyperautomation-human-augmentation-and-distributed-cloudamong-top-10-technology-trends-for-2020/>. Acesso em: 18 set. 2025.

OPENAI. **GPT-4 Technical Report**. [S. l.]: arXiv, 2023. (arXiv:2303.08774). Disponível em: <https://arxiv.org/abs/2303.08774>. Acesso em: 15 set. 2025.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO (OCDE). **OECD Employment Outlook 2019: The Future of Work**. Paris: OECD Publishing, 2019. Disponível em: <https://doi.org/10.1787/9ee00155-en>. Acesso em: 3 dez. 2025.

PINHEIRO, Á. F. **Ambiente de Inteligência Artificial e Computacional para Camada de Serviços no Setor Público**. 2023. Tese (Doutorado em Engenharia da Computação) – Universidade de Pernambuco, Recife, 2023. Disponível em: <https://doi.org/10.5281/zenodo.8000503>. Acesso em: 18 set. 2025.

PINHEIRO, Á. F. **BRASIL**. Escola Nacional de Administração Pública. Escola Virtual de Governo. Brasília, DF: ENAP, 2022. Disponível em: <https://www.escolavirtual.gov.br/curso/797>. Acesso em: 18 set. 2025.

PINHEIRO, Á. F. **Guia da Engenharia de Prompt 2.0**. [S. l.]: Zenodo, 2024. DOI: 10.5281/zenodo.14757333. Disponível em: <https://doi.org/10.5281/zenodo.14757333>. Acesso em: 18 set. 2025.

PINHEIRO, Á. F. **Hiperautomação (IA + RPA): Automação de Processos para Transformação Digital no Setor Público**. [S. l.]: Zenodo, 2024. DOI: 10.5281/zenodo.10818860. Disponível em: <https://doi.org/10.5281/zenodo.10818860>. Acesso em: 18 set. 2025.

PINHEIRO, Á. F.; BECKER, R. M.; COSTA, A. L.; VENTIMIGLIA, A. **Global Perspectives on Animism and Autonomous Technologies**. Springer Nature, 2025. Disponível em: <https://link.springer.com/book/10.1007/978-3-031-88209-8>. Acesso em: 18 set. 2025.

PINHEIRO, Á. F.; GUIMARAES, R. **Hyperautomation in Government Digital Transformation**. Journal of Mathematical Techniques and Computational Mathematics, [S. l.], v. 2, n. 7, p. 330-336, 2023. Disponível em: <https://doi.org/10.33140/JMTCM>. Acesso em: 18 set. 2025.

PINHEIRO, Á. F.; LIMA NETO, F. B. de. **Use of Domain Engineering in Hyperautomation Applied to Decision Making in Government**. Journal of Advances in Artificial Intelligence, [S. l.], v. 1, p. 103-116, 2023. Disponível em: <https://doi.org/10.18178/JAAI>. Disponível em: [Hyper Automation in Government Digital Transformation | Opast Publishing Group](https://www.opastpublishing.com/Opast-Publishing-Group). Acesso em: 18 set. 2025.

PINHEIRO, Á. F.; LIMA NETO, F. B. de. **Use of Machine Learning for Active Public Debt Collection with Recommendation for the Method of Collection Via Protest**. In: INTERNATIONAL CONFERENCE ON DATA MINING AND APPLICATIONS, 8., 2022, Vancouver. Anais [...]. Vancouver: AIRCC Publishing Corporation, 2022. p. 99-108. Disponível em: <https://doi.org/10.5121/csit.2022.120909>. Disponível em: [Prediction of active debt in the State of Pernambuco, Brazil | Revista de Engenharia e Pesquisa Aplicada](https://www.revista.engenhariae pesquisa.com.br/active-debt-in-the-state-of-pernambuco-brazil). Acesso em: 18 set. 2025.

PINHEIRO, Á. F.; SANTOS, W. B. **Prediction of active debt in the State of Pernambuco, Brazil.** Revista de Engenharia da Universidade de Pernambuco, Recife, v. 5, n. 1, 2020. Disponível em: <https://doi.org/10.25286/repa.v5i1>. Acesso em: 18 set. 2025.

PINHEIRO, Á. F.; SANTOS, W. B. **Prioritization, and transparency in software development: action research in public administration.** In: BRAZILIAN WORKSHOP ON SOCIAL NETWORK ANALYSIS AND MINING, 10., 2021, Evento online. Anais [...]. Porto Alegre: SBC, 2021. Disponível em: <https://doi.org/10.5753/brasnam.2021.16147>. Acesso em: 18 set. 2025.

PINHEIRO, Á. F.; SANTOS, W. B.; LIMA NETO, F. B. de. **Intelligent Framework to Support Technology and Business Specialists in the Public Sector.** Institute of Electrical and Electronics Engineer - IEEE Xplore, [S. l.], v. 11, p. 15655-15679, 2023. Disponível em: <https://doi.org/10.1109/ACCESS.2023.3243195>. Acesso em: 18 set. 2025.

PINHEIRO, S. **Estatística Básica: a arte de trabalhar com dados.** 2. ed. São Paulo: GEN LTC, 2015.

RAJKOMAR, A.; DEAN, J.; KOHANE, I. **Machine learning in medicine.** New England Journal of Medicine, [S. l.], v. 380, p. 1347-1358, 2019.

RUSSELL, S.; NORVIG, P. **Artificial Intelligence: A Modern Approach.** 4. ed. Upper Saddle River: Pearson, 2022.

SAURKAR, A. V.; PATHARE, K. G.; GODE, S. A. **An Overview of Web Scraping Techniques and Tools.** International Journal on Future Revolution in Computer Science & Communication Engineering, Bikaner, v. 4, n. 4, p. 363-367, abr. 2018. Disponível em: <http://www.ijfrcsce.org/index.php/ijfrcsce/article/view/1529>. Acesso em: 18 set. 2025.

SCHWAB, K. **A Quarta Revolução Industrial.** São Paulo: Edipro, 2018.

SERENGIL, S. I. **A Visual History of Deep Learning.** Sefik.io, 2018. Disponível em: <https://medium.com/@shivasaichakradhar/the-history-of-deep-learning-a91a25f78398>. Acesso em: 18 set. 2025.

SIMON, M.; JRHUGGIN, S. **Selenium Client Driver.** [S. l.]: Selenium, 2022. In: Selenium 4.6 Documentation. Disponível em: <https://www.selenium.dev/selenium/docs/api/py/index.html>. Acesso em: 18 set. 2025.

TSENG, H. H.; LUO, Y.; CUI, S. B.; CHIEN, J. T.; TEN HAKEN, T. G.; NAQA, I. El. **Deep reinforcement learning for automated radiation adaptation in lung cancer.** Medical Physics, [S. l.], v. 44, p. 6690-6705, 2017.

VASWANI, A.; SHAZEER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; KAISER, L.; POLOSUKHIN, I. **Attention Is All You Need**. In: CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS, 31., 2017, Long Beach. Proceedings [...]. [S. l.]: Curran Associates, Inc., 2017.

WILLCOCKS, L. P.; LACITY, M. **Nine scenarios arising from the growing use of robots**. LSE Business Review, 2015. Disponível em: <https://blogs.lse.ac.uk/businessreview/2015/09/29/nine-likely-scenarios-arising-from-the-growing-use-of-robots/>. Acesso em: 18 set. 2025.

YU, C.; LIU, J.; NEMATY, S.; YIN, G. **Reinforcement learning in healthcare: a survey**. ACM Computing Surveys, [S. l.], v. 55, n. 1, p. 1-36, 2019.

ZHENG, T.; PEARCE, M.; RIZZO, S.; SHIE, S. J.; SCHLEIGH, W. C. **Predictive modeling of hospital readmissions using electronic health record-wide machine learning: a case-study**. PLoS ONE, [S. l.], v. 12, n. 7, p. e0181173, 2017.

## AUTOR: ÁLVARO FARIAS PINHEIRO



Pesquisador e Professor com Pós-Doutorado em Inteligência Artificial com ênfase em Tomada de Decisão Jurídica pela Universidade Federal de Pernambuco (UFPE-PPGD). Doutor em Engenharia da Computação com ênfase em Inteligência Computacional pela Escola Politécnica de Pernambuco da Universidade de Pernambuco (UPE-PPGEC). Mestre em Engenharia de Software pelo Centro de Estudos e Sistemas Avançados do Recife (CESAR).

MBA em Inteligência Artificial com ênfase em Marketing Digital pela Faculdade UNYLEYA. Especialização em Governo Digital pela Faculdade VERBO JURÍDICO e Metodologias de Desenvolvimento em Engenharia de Software pela União Brasileira de Tecnologia (UNIBRATEC). Extensão universitária em Gestão Governamental pela Faculdade de Administração de Pernambuco (FCAP). Bacharel em Sistemas de Informação com ênfase em Engenharia de Software pela Faculdade Integrada do Recife (FIR). Professor Colaborador no Programa de Pós-Graduação em Direito da Universidade Federal de Pernambuco (UFPE-PPGD). Servidor Público Estatutário da Agência Estadual de Tecnologia da Informação de Pernambuco (ATI/PE) no cargo de Analista em Gestão de Tecnologia da Informação e Comunicação (AGTIC) na função de Diretor-Geral de Inovação e Inteligência Artificial em Saúde (DGIAS) da Secretaria de Saúde do Estado de Pernambuco (SES/PE).

Fonte: <http://lattes.cnpq.br/3430763912816425>

View on mobile



PARA ACESSO AS MINHAS  
REDES, SEGUE LINK:

[HTTPS://LINKTR.EE/ALVAROFFPINHEIRO.](https://linktr.ee/alvarofpinheiro)



**CURSO**

# INTELIGÊNCIA ARTIFICIAL NA SAÚDE

## CONTATO

Rua Quarenta e Oito, 224, Espinheiro - Recife - PE

<https://esppe.saude.pe.gov.br/>

(81) 3184-4105 / 3184-4106



**@ESPPE\_SES\_PE**

APONTE A CÂMERA DO SEU CELULAR  
PARA O QR CODE AO LADO E FIQUE POR  
DENTRO DE TUDO O QUE ACONTECE NA  
ESPPE.



**CURSO**  
**INTELIGÊNCIA**  
**ARTIFICIAL NA**  
**SAÚDE**



Secretaria  
da Saúde



**GOVERNO**  
**PER**  
**NAM**  
**BUCO**  
ESTADO DE PERNAMBUCO