

IA

EXPLAINED

**EXPLORANDO O PODER DA AI, BATE-PAPO
SIMPLIFICADO**

Com Ricardo Duram

Junho 2024



CONTEÚD



Básicos da IA

ESTADO DA ARTE DA IA

O QUE É IA?

PORQUE AGORA?

MACHINE LEARNING

DIFERENTE TIPOS DE IA



Aplicações em IA

IA ENTRE INDÚSTRIAS

IA E MEDICINA

IA E INDÚSTRIA AUTOMATIVA

IA E EDUCAÇÃO

IA APLICADO PARA BOM OU RUIM?



Desafios da IA

ÉTICA

VIÉS

SEGURANÇA

PRIVACIDADE



O FUTURO DA IA

O PAPEL DOS DECISORES



COMO VAMOS **NAVEGAR** HOJE



NOSSA APRESENTAÇÃO É DIVIDIDA EM CINCO CODIFICAÇÕES POR CORES & CAPÍTULOS, PARA FÁCILITAR A NAVEGAÇÃO.

“I C O N E S”, ALGUNS PARA INDICAR MAIS DETALHES E INFORMAÇÃO OU RECURSO ESPECÍFICO SOBRE O TÓPICO.



PROFUNDO MERGULHO

lupa para pegar a visão mais profunda em um capítulo ou tema.

BÁSICOS DA IA

5 ESTADO DA ARTE DA IA

6 O QUE É IA?

10 PORQUE AGORA?

11 MACHINE LEARNING

16 DIFERENTE TIPOS DE IA



**Artificial
Inteligência (IA) é
provavelmente e
será a melhor ou a
pior coisa a
acontecer para
humanidade.**

Stephen Hawking



ESTADO DE ARTE DA IA

Tecnologia do PRESENTE OU FUTURO ?

Hoje, todo mundo está falando sobre IA e como ela vai mudar a maneira como vivemos nossas vidas. Mas a IA não trata apenas do futuro.

IA é, já está aqui, hoje, afetando pessoas em todo o mundo.
A IA está mudando a vida dos indivíduos. Usamos IA em nossa vida cotidiana: Google Maps, instruções de direção, filtros de e-mail de spam, sugestões de música, sugestões de amigos, pesquisas online e até filtros do Instagram, tudo isso é alimentado por inteligência artificial. Aplicações de IA começaram a transformar grandes indústrias, incluindo saúde, finanças, manufatura, varejo, educação e entretenimento. Portanto, a IA não está longe nem é estranha à nossa vida cotidiana. E tem imensas implicações para muitas esferas das nossas vidas moldadas por diversas políticas de uso.

Hoje falamos sobre Fundamentos de IA e como ela afeta nossa sociedade. Isto nos ajuda a pensar no atual e futuro do desenvolvimento de IA.

O QUE É IA ?



artificial intelligence noun

Definition of *artificial intelligence*

- 1 : a branch of computer science dealing with the simulation of intelligent behavior in computers
- 2 : the capability of a machine to imitate intelligent human behavior

Não existe um acordo único de definição aceita de inteligência artificial.

Geralmente é definido como a **capacidade de uma máquina desempenhar funções cognitivas e associamos a mente humana, como perceber, raciocinar, aprender e Solucionar problemas.**

A IA não é uma tecnologia única; é antes, uma coleção de diferentes tecnologias que pode permitir que a máquina atue em níveis semelhantes a inteligência dos humanos.

O estado atual utiliza machine learning, deep learning e reinforced learning aprendendo a produzir insights, e aprimorar o trabalho baseado em conhecimento e automatizar tarefas rotineiras.

Exemplos de tecnologias de IA que resolvem problemas hoje são robótica e veículos Autônomos, visão computacional, linguagem e agentes virtuais.

A pesquisa em IA usa ferramentas e insights de muitas áreas, incluindo ciência da computação, psicologia, filosofia, neurociência, ciência cognitiva, linguística, operações e pesquisa, economia, teoria de controle, probabilidade, otimização e lógica.



Explore a Biblioteca Digital EAAA para saber mais sobre como definir explicando termos relacionados à IA



John McCarthy, que cunhou o termo em 1956, e define isso: "E quando a Ciência e Engenharia tornar máquinas inteligentes?"



O Dicionário Oxford define-o como a teoria e desenvolvimento de sistemas computacionais capazes para executar tarefas que normalmente requerem recursos humanos e inteligência, como percepção visual, reconhecimento de fala, tomada de decisões e tradução entre idiomas.



Economist Intelligence Unit define IA como um conjunto de técnicas de ciência da computação que permite que os sistemas executem tarefas normalmente exigindo inteligência.



IA NÃO É UM CONCEITO NOVO

O termo “**inteligência artificial**” foi criado em 1956. Neste momento foi proposto o famoso [Turing teste](#) em 1950 e o termo “artificial inteligência” era usado na oficina de workshop na Faculdade Dartmouth em 1956. Aos atendentes eram disponibilizados milhões de dólares para criar uma máquina tão inteligente como um humano, mas não conseguiram alcançar a meta. Esse era o primeiro, e assim chamado “Inverno IA” (AI winter), da IA não ser capaz de completar as expectativas, seguido por outros “invernos” seguintes. **Até agora.**

POR QUE AGORA?

A IA já existe há muito tempo, mas apenas agora estamos vendo um aumento no avanço tecnológico da IA. **Há algumas razões principais para esta aceleração:**



PODER DE PROCESSAMENTO

O poder de processamento e capacidade de armazenamento de máquinas de hoje é muito maior do que qualquer coisa que vimos antes.



BAIXOS CUSTOS

Ao mesmo tempo, o custo do acesso à nuvem, armazenamento e processamento de dados caiu, e a disponibilidade para armazenar grandes conjuntos de dados baratos aumentaram.



DISPONIBILIDADE DE DADOS

Grande disponibilidade de “**big data**”, extremamente grandes conjuntos de dados que podem ser analisados computacionalmente para revelar padrões, tendências, e associações, especialmente relacionadas com comportamento e interações humanas.

Isto melhorou imensamente a nossa capacidade de construir sistemas inteligentes capazes de aprender.

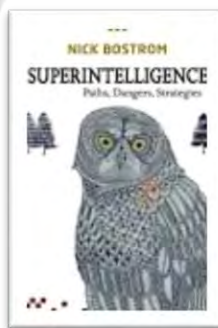
DIFERENTES TIPOS DE IA

IMPORTÂNCIA DE UM CLARO ENTENDIMENTO

Considerando os medos ao redor das tecnologias IA, é muito importante para entender a distinção entre diferentes níveis de IA.

Então, que potenciais ameaças e benefícios podem ser racionalmente e realisticamente avaliados.

A IA evoluiu significativamente e hoje pode ser usado para preencher as lacunas nas habilidades do ser humano. Mas sistemas IA não são como os humanos em raciocínio e pensamento criativos e estratégico.



ANI INTELIGÊNCIA ARTIFICIAL ESTREITA

estreita ou IA fraca, focada em uma tarefa única e simples. Possui uma estreita gama de habilidades. IA estreita é algo que a maioria de nós interage diariamente (Google Assistente, Google Tradutor, Siri, ou Alexa).

AGI INTELIGÊNCIA ARTIFICIAL GERAL

Geral ou IA Forte está muito mais próxima da inteligência humana. No entanto, AGI ainda é um campo emergente. O cérebro humano é o modelo para a criação do General Intelligence, e ainda nos falta o conhecimento compreensivo se isso é funcional.

ASI INTELIGÊNCIA ARTIFICIAL SUPER INTELIGENTE

Inteligência artificial Super é um IA mais poderoso e sofisticado como inteligência humana. Para alcançar esse ponto e para ser chamada de ASI, uma IA precisará superar os humanos em absolutamente tudo. O tipo ASI é alcançado quando IA é mais capaz que o humano.



IBM Deep Blue é um exemplo de ANI, desenvolvido pela IBM em 1997. Ele é conhecido como o primeiro jogador de xadrez computacional a ganhar um jogo de xadrez em partida contra o campeão mundial Gary Kasparov.

Quatorze anos depois IBM aplicou IA ao desafio do **Jeopardy**, o quiz show que estava cobrindo questões sobre "apenas qualquer coisa".

IBM Watson incorporou aspectos de IA, aprendizado de máquina, profundidade em resposta a perguntas e processamento de linguagem natural para jogar e, finalmente derrotou o jogo o melhor campeão.

AGI EM 2040

Pesquisa de opinião de especialistas sobre o futuro Progresso em Inteligência Artificial, conduzido em 2011, sugeriu que a estimativa mediana foi de 50% de chance que a inteligência artificial geral (AGI) seria desenvolvida até 2040–2050.

Machine Learning

O aprendizado de máquina é um ramo da inteligência artificial, o núcleo da IA, com base na ideia de que **os sistemas podem aprender com os dados, identificar padrões e tomar decisões com o mínimo de recursos humanos na intervenção.**

O processo é alimentado por um algoritmo, uma função que é capaz de melhorar seu desempenho ao longo do tempo, treinando-se usando métodos de análise de dados e modelagem analítica. Algoritmos de aprendizado de máquina usam estatísticas para encontrar padrões em grandes quantidades de dados. Os dados incluem muitas coisas - números, palavras, imagens, cliques. Se puder ser armazenado digitalmente, poderá ser alimentado em um algoritmo de aprendizado de máquina.

O aprendizado de máquina é o processo que alimenta muitos dos serviços que usamos hoje - **sistemas de recomendação** como os do Netflix, YouTube e Spotify; **motores de busca** como Google e Bing; **alimenta mídia social** como Facebook e Twitter; **assistentes de voz** como Siri e Alexa.

Deep Learning

O aprendizado profundo é um subcampo do aprendizado de máquina, modelado a partir do sistema nervoso humano, conhecido como rede neural artificial. Ele usa uma técnica que dá às máquinas uma capacidade aprimorada de encontrar e amplificar até mesmo os menores padrões. Essa técnica é chamada de **rede neural profunda** - profunda porque tem muitas camadas de nós computacionais simples que trabalham juntos para trabalhar dados e entregar um resultado final na forma de **previsão.**

ALGORITMO é um conjunto de etapas para o programa de computador realizar um tarefa.

? Como os vídeos ao vivo podem ser transferidos entre a internet tão rapidamente?

»»» algoritmo de compressão áudio e visual

? Como o Google Maps nos dá instruções de direção?

»»» algoritmo de rotas e direção

? Como pode a NASA reorganizar e agendar movimentação de painéis solares na Estação Internacional Espacial ?

»»» algoritmo de Otimização e agendamento



Saiba mais sobre máquina aprendendo no curso gratuito Coursera [IA para todos](#).

O NÚMERO DE EMPRESAS COM IMPLEMENTAÇÃO DE IA CRESCERÁ

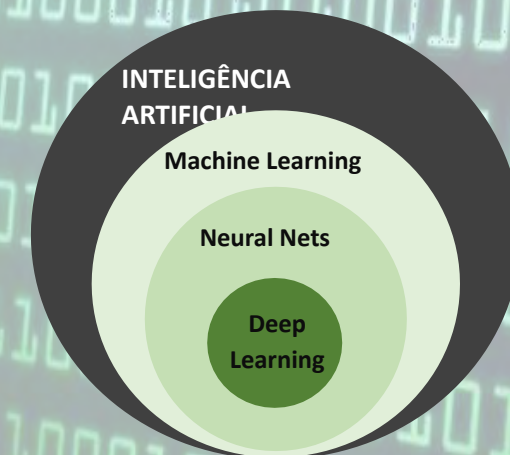
270%

EM QUATRO ANOS E

triplicado

NO ANO PASSADO

Gartner, 2023 Enquete



COMO PODE A MÁQUINA APRENDER ?

DADOS DE TREINO

Aprendizado Máquina profundo é confiar ao treinamento, dados para aprender os relacionamentos, aumentar a eficiência do modelo , e melhorar isso é a habilidade para alcançar a saída desejada.

Então, em-se a um conjunto de dados que foi coletado, preparado, e oferecido para o modelo com o propósito de ensinar antes da atividade de Implantação. A qualidade, quantidade, estrutura e conteúdo dos dados de treinamento são fundamentais e determinantes de como o aprendizado de máquina e modelos de aprendizagem profunda funcionarão em um ambiente real.

Há diferentes tipos de aprendizado de máquina:

SUPERVISIONADO

O dado é rotulado (organizado e descrito) para dizer a máquina exatamente quais padrões ele deve procurar. O algoritmo então deduz as características salientes que caracteriza cada rótulo e aprende para reconhecer aquelas características em novos dados.

NÃO SUPERVISIONADO

Os dados não possuem rótulos. A máquina apenas procura o quaisquer que seja padrões que pode encontrar. O Algoritmo deve estruturar dados, descobrir padrões, classificar entradas, aprender funções e produzir resultados sem recursos externos de validação ou apoio.

REFORÇADO

A IA aprende por **tentativa e erro** para alcançar um resultado objetivo. Ela experimenta muitas coisas diferentes e é recompensado ou penalizado dependendo se isso é, comportamentos que ajudam ou impedem que ele alcance o objetivo.

? "Cachorro ou Muffin?"

Como ensinamos a um computador a diferença entre objetos visualmente quase idênticos em cor e forma? pode ser um desafio ver a diferença. O desafio foi aceito pela comunidade online e a IA foi colocada em ação. Imagens eram passadas através de APIs de reconhecimento e IAs foram capazes de ver a diferença com alta precisão (com algumas ocasiões podiam confundir muffin recheado com animal).



Deep Learning é a superpotência . Com isto você pode fazer um computador ver, sintetizar romance, arte, traduzir línguas, renderizar um diagnóstico médico, ou construir peças de carro que pode dirigir sozinho. Se isso não é a superpotência, Eu não sei que é.

André Ng , Fundador de deeplearning.ai



APLICAÇÕES EM IA

12

IA ENTRE INDÚSTRIAS

13

IA E MEDICINA

14

IA E INDÚSTRIA AUTOMATIVA

15

IA E EDUCAÇÃO

16

IA APLICADO PARA BOM OU RUIM?

”

**IA não é apenas
uma mudança
de jogo; isto irá
mudar o campo
de jogo inteiro.**

Neil Jacobstein

Singularity University Artificial
Intelligence and Robotics Chair

APLICAÇÕES IA ENTRE AS INDÚSTRIAS

APLICADA PARA BOM OU RUIM ?

A adoção da IA cresceu em uma ampla gama de áreas nos dois anos anteriores. A tecnologia por trás da IA está sendo implantada em fazer música e escrever livros, em medicina, farmacêutica e de saúde, em guerra e ação humanitária, na tomada de decisões sobre nosso crédito e seguros. Em resumo, em tomar decisões que afetam nossas vidas, quer queiramos ou não.

A IA tem sido usada para desenvolver e avançar em vários campos e indústrias, incluindo finanças, saúde, educação, transporte e muito mais.

Um algoritmo não pode escolher onde, quando ou como será usado, ou se será usado para boas ou más intenções. Esta decisão está em nossas mãos, bem como o desafio de garantir o uso ético da IA.

À medida que a IA se desenvolve e as suas aplicações crescem, **há uma grande oportunidade, mas também uma grande responsabilidade de garantir que contribua para o bem público e beneficie todos**, com justiça, confiabilidade, segurança e, quando apropriado, transparência e privacidade garantida.

IA E MEDICINA

IA era aplicado como o modelo que le e interpreta as descobertas de tomossíntese nos seios através de imagens digitais, tridimensionais nas mamografias que leva m múltiplas fotos de seios para detectar possível cânceres.

Os resultados mostraram que a ferramenta de aprendizagem profunda foi capaz de melhorar a precisão de detecção e reduziu os tempos de leitura pela metade. Sensibilidade, uma medida do teste da capacidade de identificar corretamente aqueles com a doença, aumentou de 77% para 85% por cento quando IA era empregado.

O tempo obrigatório por radiologistas para ler e interpretar as descobertas reduziu-ae de 2-3 minutos em media para 30 segundos com a assistência da IA. Tal a redução na leitura em tempo significa que os radiologistas serão capazes de interpretar mais imagens em menos tempo e diagnosticar mais indivíduos eficientemente.

»» melhorou eficiência e precisão

JUNTO COM O CÂNCER DE MAMA. OS PESQUISADORES ESTÃO OBSERVANDO COMO A IA PODE MELHORAR A EFICÁCIA E EFICIÊNCIA DO CUIDADO DO PULMÃO, CÉREBRO E CÂNCER DE PRÓSTATA, DE ENCONTRO COM AS EXIGÊNCIAS DE DIAGNÓSTICO.

*Câncer é o segundo líder na causa de morte globalmente (um em seis mortes é devido ao câncer). Câncer em Seios é a maioria de Câncer comum no mundo , junto com Câncer de pulmão, com **2.09 milhões de casos** .*

 **Mundo, Organização da Saúde, 2020**

 Google AI Blog

Até o Google está olhando em como a IA pode ser usada para diagnosticar o câncer. Eles estão treinado um algoritmo para detectar metástase de tumores

99%

UM SUCESSO PARA AVALIARMOS

Para um mergulho profundo em aplicações de IA na medicina, saúde e farmacêutica busque sobre o tema.



IA E INDÚSTRIA AUTOMATIVA

A AMEAÇA E A PROMESSA DE CONDUÇÃO

IA é usado para habilitar o carros para navegar através o tráfego e lidar com situações complexas . Programas de IA combinado com sensores de IoT, tal como câmeras, para garantir a propriedade e segurança ao dirigir. Veículos Autônomo são equipado com sensores, câmeras, e radares colhendo dados, trabalhando para perceber o ambiente em torno do veículo como a pista, árvores, tráfego, pedestres e luzes.

Existem diferentes níveis de automatização da condução: automatização para assistência ao condutor, direção parcial, direção totalmente automatizada e carros totalmente automatizados. Enquanto automação para assistência ao motorista e carros parcialmente automatizados estão em uso comercial, os tipos restantes ainda estão em condições de teste.

O impacto de IA é não limitado para automatização de carros. Fabricantes de carros e fábricas estão procurando caminhos para utilizar IA para melhorar a produção, qual traz medos sobre perda de empregos devido automação de trabalho.

INTERNET DE COISAS (IoT) descreve a rede física de objetos - “coisas” que são integrados com sensores, Programas, e outras tecnologias com o propósito de conectar e trocar dados com outros dispositivos e sistemas usando a Internet. Essa gama de dispositivos de cenário doméstico migram para indústrias sofisticadas e suas ferramentas.

APROXIMADAMENTE

1,25
milhão

PESSOAS MORREM EM
ACIDENTES DE ESTRADA
CADA ANO NA MÉDIA

3.287

MORTES
POR DIA



Associação para
segurança
internacional de
Estrada e Viagem

MERCADO
AUTOMOTIVO

6,5%

CRESCIMENTO
2019-2023

Global Automotive Motor
Market 2019 Industry Report



Nesta era de digitalização e tecnologias inteligentes, o uso de autonomia nos veículos como modo de transporte não é uma questão de 'se', mas de 'quando'.

Professor Chen Tsuhan
Nacional Universidade de
Cingapura



“CONSTRUINDO O
MOTORISTA MAIS
EXPERIENTE DO
MUNDO”

Certo agora, a minivan sem um motorista atrás do volante está dirigindo por um subúrbio de Fênix, Arizona. Waymo, o veículo autônomo da divisão de Alphabet, controladora do Google a empresa é a primeira a colocar carros nas estradas dos EUA sem um motorista de segurança.



Leia sobre os desafios de regulamentando dirigir sozinho carros.

IA E EDUCAÇÃO

PRÓXIMO GERAÇÃO DE APRENDIZADO PERSONALIZADO

Tecnologia Educacional é o processo integrando tecnologia em educação de uma maneira positiva que promove os mais diversos ambientes de aprendizado. Como IA torna-se mais proeminente nessa era de grandes dados, isso é mais adotado no campo da Educação também.

Os sistemas escolares existentes são concebidos sob o pressuposto de que todos os alunos progredem no mesmo ritmo. A IA tem o potencial de criar salas de aula altamente diferenciadas e individualizadas, não substituindo, mas ajudando professores para apoiar os alunos, compartilhando a carga de trabalho e recomendando o melhor curso de ação. Os aplicativos IA podem personalizar a experiência de aprendizagem, sugerindo objetivos de aprendizagem individuais, selecionando abordagens instrucionais e exibindo exercícios baseados nos interesses e no nível de habilidade de todos os estudantes.

Existem cinco áreas principais onde a tecnologia de IA está sendo desenvolvida e usada na educação: **hardware** – dispositivos inteligentes, **robótica e software** alimentados por IA; **visão e reconhecimento de emoção** para detectar confusão, envolvimento e para aprendizagem virtual; **interfaces de texto para voz** são usados em alfabetização e treinamento de linguagem; **processamento de linguagem natural** é usado na avaliação, opinião, e detecção de plágio; **algoritmos** estão sendo aplicados para criar aprendizado personalizado, adaptativo aos caminhos dos estudantes.



2019 Artificial Inteligência & Global Educação Relatório

VIÉS E PREOCUPAÇÕES DE PRIVACIDADE

Organizações devem agora considerar que tipo de dados está sendo coletado, como esse a informação é ser usada e o que controla são o lugar para salvaguardar a privacidade dos estudantes.

O Futuro de IA nas Escolas

PERSONALIZADO

IA pode coletar e analisar dados, ajudando educadores a criar planos de aprendizagem personalizados para cada estudante.

QUEBRA VIÉS

A IA pode ajudar a aproveitar soluções capazes de avaliar exames e avaliando artigos usando rubricas estabelecidas e benchmarks para automatizar conclusão e eliminar viés.

AGREGANDO ASSISTÊNCIA

Orientado por IA assistentes podem liberar tempo e ajudar os professores a dedicar mais horários para conectar com os alunos.

US\$ 1



US\$ 6

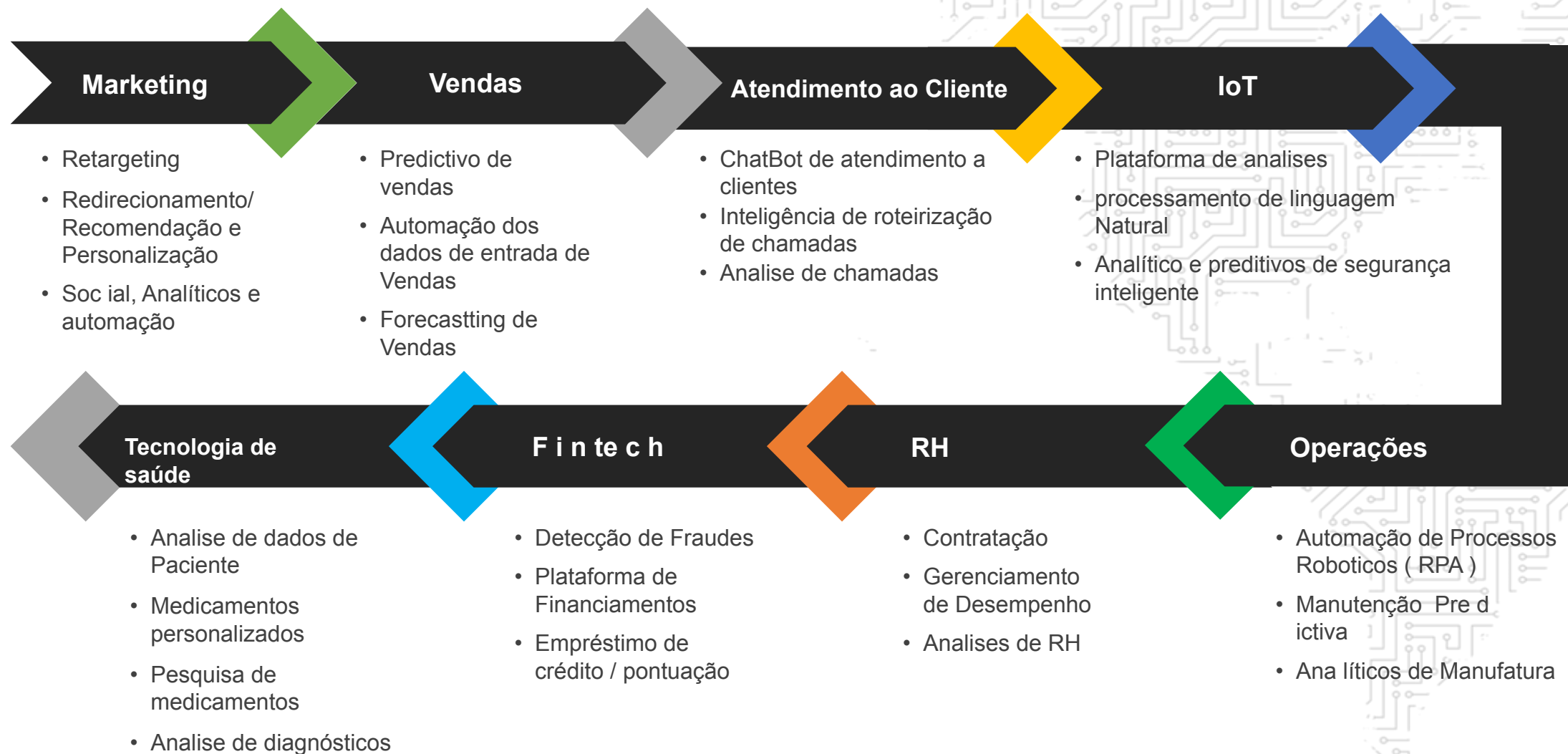


EM 2018, US\$ 1 BILHÃO NO MUNDO ERA GASTO EM IA NA EDUCAÇÃO, E QUE NÚMERO É ESPERADO PARA CRESCER ACIMA DE US\$ 6 BILHOES ATÉ 2025.



Verificar como Escolas K-12 adotaram IA em 2019.

IA Principais Usos HOJE



Green IA E caso de uso real Brasil

PHYCUS NA NOVA GERAÇÃO DE APRENDIZADO PERSONALIZADO

EMPRESA BRASILEIRA DE GERAÇÃO DE CRÉDITOS DE CARBONO E SUAS INOVAÇÕES TECNOLÓGICAS.



DESEMPENHO PERSONALIZADO

Empresa usa IA nas suas plataformas de gestão, controle de ativos, controle de crescimento/Desenvolvimento e conversão de unidades em credito baseados em predição

Simple Blue Decarbonization



IA APLICADO PARA BOM OU RUIM?

AI FOR GOOD baseia-se em ideias anteriores de uso dos mais recentes avanços tecnológicos para resolver problemas sociais e desafios, nomeadamente a ideia Tech for Social Good de utilização de tecnologias digitais para resolver alguns dos os maiores desafios do mundo. Os maiores desafios mundo.



A Microsoft iniciou o programa “ **AI for Good** ”, fornecendo tecnologia, recursos e experiência para capacitar aqueles trabalhando para resolver problemas humanitários e criar as mais sustentável e acessível mundo.



Google iniciado “ **Google IA** ” IA para o programa social bom, trabalhando para usar IA endereçar social desafios através da pesquisa, engenharia e iniciativas para construir um ecossistema IA.



Parceria em IA reúne pesquisadores, acadêmicos, empresas e decisores políticos que invistam mais atenção e esforço em aproveitar a IA para contribuir com soluções para alguns dos problemas mais desafiadores da humanidade, incluindo saúde e bem-estar, transporte, educação e ciências.



AI for Good
Global Summit

An ITU experience

AI for Good é uma plataforma das Nações Unidas, centrada em torno das Cúpulas Globais anuais, que promove o diálogo sobre o uso benéfico da IA através do desenvolvimento de projetos concretos.

A série AI for Good conecta inovadores de IA com aqueles que procuram soluções para os maiores problemas do mundo, desafios para identificar aplicações práticas de IA que pode acelerar o progresso em direção ao UN Sustainable Development Goals (SDG's).

Em vez de perguntar que tipo de impacto IA terá sobre a humanidade, devemos começar por perguntando que tipo de impacto nós queremos que isso tenha. As pessoas trabalhando em IA, não em IA por si mesma, são, em última análise, responsável por como muito bem ou mal irá ser feito.

USO MALICIOSO DE IA

Como podemos previr, evitar, e mitigar o uso prejudicial e efeitos de uso malicioso de IA?

Esta é a questão colocada em 2018 num relatório escrito por 26 autores de 14 instituições, da academia, sociedade civil e indústria. Os autores, representantes do Future of Humanity Institute, do Centro para o Estudo do Risco Existencial, OpenAI e o Centro para uma Nova Segurança Americana argumentam que a IA não está apenas alterando a natureza e o alcance das ameaças existentes, mas também expandindo a gama de ameaças que iremos enfrentar. O objetivo do relatório é pesquisar o panorama das ameaças à segurança provenientes de ameaças intencionalmente maliciosas. usos de A.

Os autores sugerem que existem três domínios principais de segurança: segurança digital, que diz respeito em grande parte ataques cibernéticos; segurança física, que se refere à realização de ataques com drones e outros sistemas; e segurança política, que inclui exemplos como vigilância, persuasão através de propaganda e engano por meio de vídeo manipulado



Ler o inteiro relatório aqui : <https://maliciousaireport.com/>



DESAFIOS DA IA

19

IA E ÉTICA

20

IA E O VIÉS

21

IA E SEGURANÇA

22

IA E PRIVACIDADE



”

Como todas as tecnologias anteriores, a inteligência artificial refletirá o valores de seus criadores. Então inclusão importa - de quem o projeta para quem participa dos conselhos da empresa e quais perspectivas éticas estão incluídas.

Kate Crawford

Professor pesquisador da
NYU Cofundador de IA
AGORA

DESAFIOS DA IA E PREOCUPAÇÕES

O QUE DEVERIA SER PRIORIDADES E O QUE DEVERIA SER POLÍTICAS APLICADAS?

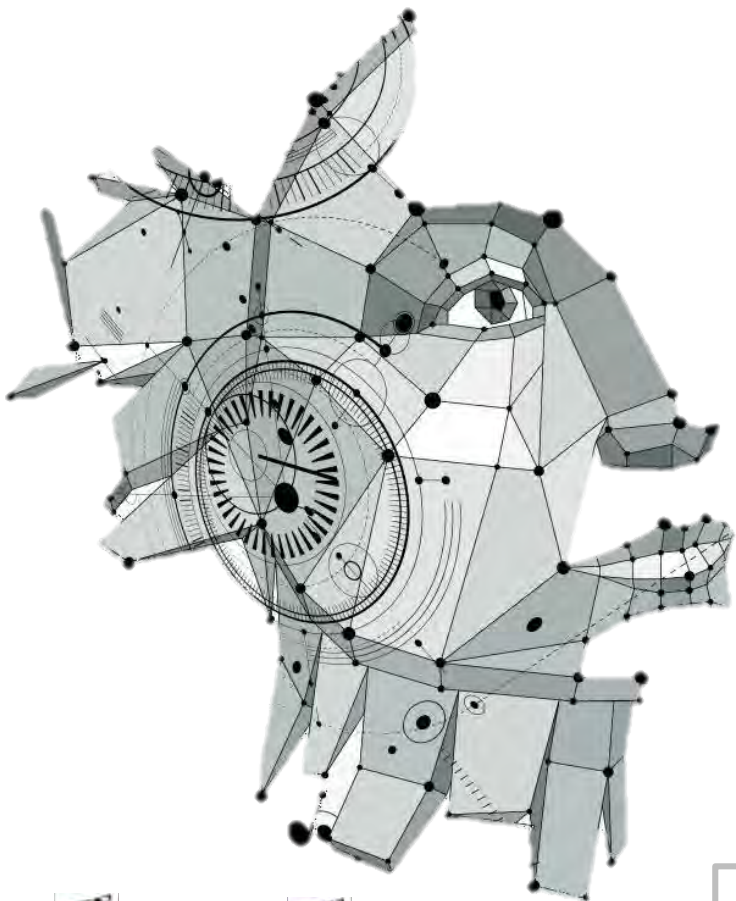
Como falamos, a IA já traz benefícios económicos, financeiros e desenvolvimento social, investigação médica e melhores cuidados de saúde, protege a segurança e os direitos humanos e ajuda a resolver problemas ambientais problemas. Tem um enorme potencial para ajudar a resolver os problemas mais difíceis do mundo. E melhorar inúmeras vidas.

No entanto, com os sucessos surgem novas preocupações. A fim de realizar de forma responsável todo o seu potencial, os desafios associados a ele têm que ser abordados. O Future of Life Institute está oferecendo um extensa lista de 14 tópicos de especial preocupação para a segurança e desenvolvimento no benéfico da IA, num futuro próximo e distante.

Algumas destas preocupações incluem os impactos económicos, as mudanças laborais e desigualdade, responsabilização, transparência e explicabilidade, vigilância, privacidade e liberdades civis, justiça, ética e direitos humanos, segurança da IA, segurança e cibersegurança, manipulação política e propaganda computacional, armas autônomas.

Estamos destacando quatro preocupações gerais: ética, Preconceito, segurança e privacidade, acreditando que são um desafio transversal para muitas aplicações de IA na indústria. Com uma breve visão geral, estamos oferecendo alguns exemplos, propostas e links para leituras adicionais.

IA DESAFIOS IA E ÉTICA



A ética da IA é um campo mais amplo, incluindo tópicos mais específicos de discriminação, justiça, preconceito e direitos humanos. Hoje, as ciências da computação e a IA carecem diversidade de gênero, racial e outras formas, que podem levar a falsas suposições, design distorcido e produtos discriminatórios e injustos.

Os algoritmos podem reproduzir e ampliar preconceitos sociais e discriminação de usar dados de treinamento que reflitam preconceitos existentes na sociedade ou que tenham uma representação distorcida. Se os programadores estiverem isolados dos problemas eles estão tentando resolver, as chances de consequências negativas não intencionais são aumentando. A IA ética tem que começar com diversas vozes e experiências incluído no início do processo de design.

A ética da IA reside, em última análise, na interseção da qualidade ética dos seus criadores e seus resultados e a qualidade ética do impacto que tem sobre as pessoas e as sociedades.

PRINCÍPIO ARTIFICIAL INTELIGÊNCIA

Mapeando o Consenso em Ética e Abordagens aos Princípios Baseadas em Direitos para IA



Berkman Klein Centro para Internet e Sociedade

O DESTINO d@ IA

MELHORES PRÁTICAS E PADRÕES EM JUSTIÇA, RESPONSABILIDADE, TRANSPARÊNCIA E ÉTICA IA.

Existem cinco maiores setores desenvolvendo no interseção de ética e IA:

SAÚDE

De possíveis erros de diagnóstico por falhas algoritmos para edição genética

AUTÔNOMOS

Fazendo decisões da vida e morte para passageiros humanos

SEGURO

algoritmos pode fazer decisões discriminatórias contra minorias

JORNALISMO

de “notícias falsas” para “deepfake” vídeos

MILITARES

usando IA tecnologia para drones mortais e outros

Muitos indivíduos e organizações estão trabalhando no desenvolvimento de ética, princípios e diretrizes para o criação e implantação de tecnologias IA. Esses padrões éticos só poderão ter efeito mensurável se vierem com responsabilidade.

EXEMPLOS DE IA PRINCÍPIOS DE ÉTICA



IEEE Eticamente Alinhado Projeto



ITI IA Política Princípios



Ética da DeepMind e Sociedade Princípios



Original Jeans.mp4



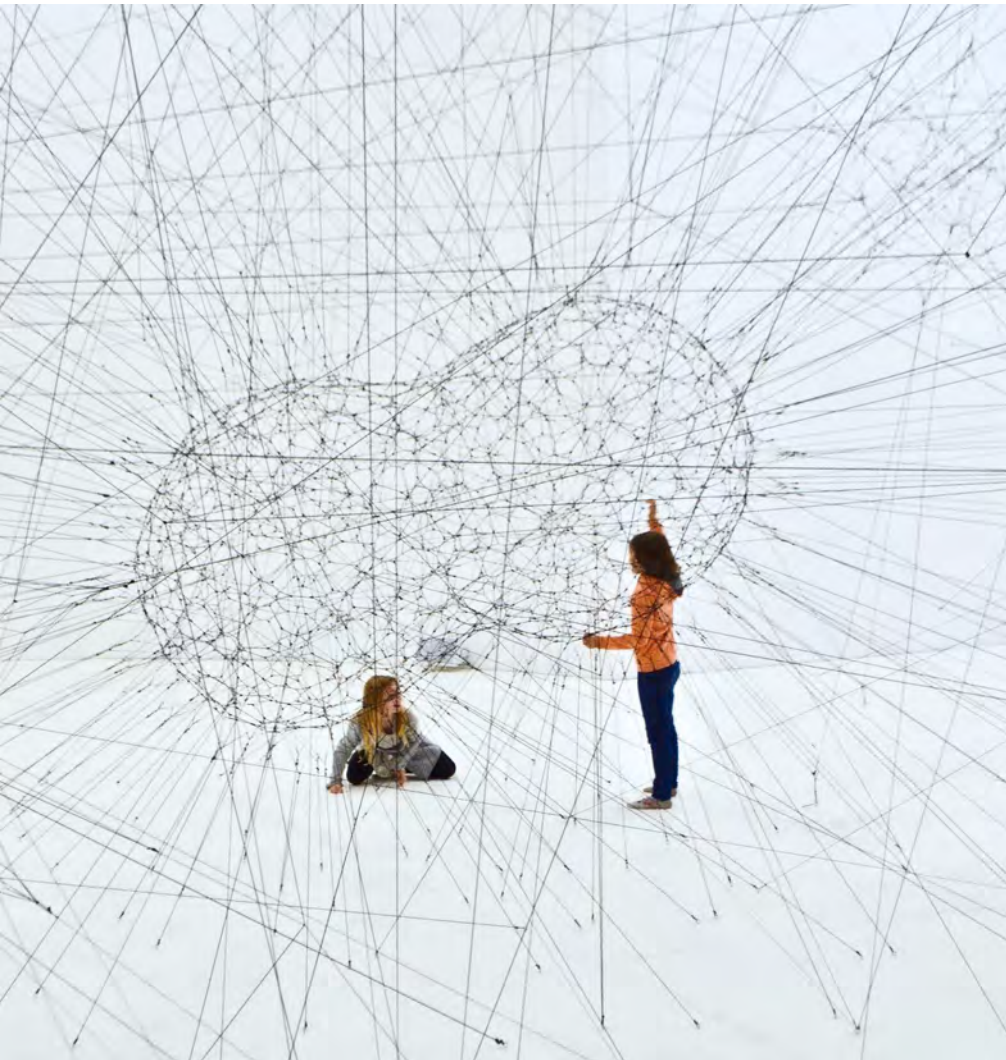
Original Vestido.mp4



swapped (1).mp4



swapped.mp4



Em consequência da fama de Rice, "Jim Crow" tornou-se, em 1838, uma forma pejorativa de se referir aos negros. No fim do século XIX, quando as legislaturas sulistas aprovaram leis de segregação racial dirigidas contra os negros, essas leis ficaram conhecidas como Jim Crow laws ('leis Jim Crow').

IA DESAFIOS **IA É VIÉS**

A IA pode ajudar a reduzir o preconceito, mas também pode reproduzi-lo e escalá-lo. IA nos sistemas 'aprendem' com base nos dados que recebem, e os dados refletem condições sociais, históricas e políticas em que foi criado, incluindo de forma implícita, gênero ou preconceitos raciais.

Como resultado, o viés algorítmico pode impactar o acesso dos indivíduos a recursos e serviços, o nível de vigilância que sofrem, seu tratamento pela polícia e pelo governo, acesso à educação, seguro ou empréstimo. Desta forma, se não for implementada adequadamente, a IA pode introduzir ou reforçar disparidades nas sociedades, levando a informações imprecisas e resultados injustos para populações diversas.

O julgamento Humano está parado, e é necessário para garantir o suporte correto da IA na decisão de fazer o justo.

POLICIA PREDITIVA

Nos Estados Unidos, algoritmos de IA estão decidindo sobre a detenção ou libertação dos réus, fornecendo recomendações sobre penas de prisão. O Relatório ProPublica mostrou evidências de que o sistema IA era tendencioso contra os réus negros, pois constantemente recomendava sentenças mais longas em comparação com brancos para o mesmo crime.

"NOVO JIM CODE"

"Apresento o conceito de o Novo Código Jim" para explorar um faixa de projetos discriminatórios que codifica desigualdade: por amplificar explicitamente hierarquias raciais, ignorando, mas sim replicando divisões sociais, ou mirando para corrigir o preconceito racial, mas em última análise fazendo bastante o oposto."

Ruha Benjamin,
Universidade Princeton, Autor de
"Corrida pela Tecnologia"

Um estudo de 2018 realizado na UC Berkeley descobriu que o consumidor de empréstimo de tecnologia não era discriminado contra outros candidatos minoritários.



Discriminação ao empréstimos ao consumidor na ERA das FinTechs

IA E SEGURANÇA

Segurança de IA refere-se ao projeto técnico de sistemas de IA que visa evitar impactos potencialmente negativos, acidentes, e consequências não intencionais e prejudiciais de um projeto.

Os pesquisadores da Open AI definem três áreas de segurança da IA: **ESPECIFICAÇÃO**, que garante que o comportamento de um sistema de IA esteja alinhado com as verdadeiras intenções dos operadores; **ROBUSTEZ**, que garante que um sistema de IA continue a operar dentro de condições seguras e em limites às perturbações; **GARANTIA**, que garante que podemos entender e controlar sistemas de IA durante as operações.

A IA aberta define uma meta de segurança de IA de longo prazo como garantir que sistemas IA avançadas estão alinhados com os valores humanos, que fazem coisas que as pessoas querem que eles façam.



MICROSOFT CHATBOT

O chatbot da Microsoft, Tay, é um recente exemplo de como uma IA pode aprender o comportamento negativo de seu ambiente, produzindo resultados bastante no oposto de o que seus criadores tinham em mente.



TESLA ACCIDENT

Em 2016, o carro Tesla em piloto automático teve um acidente, no qual o veículo confundiu um caminhão branco com um céu claro. O acidente levou a vida do motorista. Isto é um exemplo de uma IA entendendo mal o seu ambiente e tomando ação errada ou mortal como resultado.

" Queremos a solução para o problema de segurança antes que alguém descubra a solução para o problema de IA."

Nick Bostrom,
filósofo, autor de
"Superinteligência"

PROBLEMA DE ALINHAMENTO DO VALOR

Qualquer consideração de IA a segurança também deve incluir alinhamento de valores :
como podemos projetar inteligência artificiais que pode alinhar com a diversidade global de valores HUMANOS?

RISCOS EXISTENCIAIS GLOBAIS

Risco existencial de inteligência artificial geral (AGI) é a hipótese de que seu substancial progresso da AGI poderia algum dia resultar na extinção humana ou algum outra catástrofe mundial irreversível.

UMA PESQUISA REALIZADA
EM 2008 PELO FUTURO DA
HUMANIDADE ESTIMOU

5%

A PROBABILIDADE DE
EXTINÇÃO POR
SUPERINTELIGENCIA

Por volta de 2100



Instituto Futuro de Humanidade

IA DESAFIOS IA E PRIVACIDADE

Num mundo de dispositivos inteligentes e conectados em nossas casas, locais de trabalho e espaços públicos, há uma grande quantidade de dados que esses dispositivos, redes e plataformas geram, processam ou compartilham. Diante do fato de que esses dados podem ser explorados torna-se cada vez mais comuns. A possibilidade de uso indevido dos dados para criação de perfis falsos, rastreamento ou identificação de pessoas torna-se um aumento a preocupação dos humanos e leva a maiores esforços em proteger a privacidade dos dados.

As preocupações com a privacidade dos consumidores e dos cidadãos atingiram o pico nos últimos anos, em linha com, entre outras coisas, o ascender de tecnologias avançadas como IA. A IA depende das grandes quantidades de dados para fazer inferência precisa, previsões e decisões, o que traz um desafio adicional para as sociedades, indústrias e governos para **equilibrar a necessidade de mais dados com a proteção dos consumidores e cidadãos.**

RECONHECIMENTO FACIAL

Reconhecimento facial potencialmente apresenta riscos aos direitos humanos e liberdades civis em muitos países em todo o mundo. Pode reforçar o discriminatório em práticas da justiça criminal e para a educação.

AFFECT RECOGNITION

É uma tecnologia que tenta ler as emoções internas de análise perto do rosto e é conectado ao fazer reivindicações sobre o humor das pessoas, saúde mental, nível de envolvimento e culpa ou inocência. Já é sendo usado para fins discriminatórios e propósitos antiéticos, muitas vezes sem o conhecimento das pessoas.

QUESTÃO DE VIGILÂNCIA

Um número crescente de estados está implantando ferramentas avançadas de IA para monitorar, rastrear e vigiar cidadãos. Redes de sensores, rastreamento de mídia social, reconhecimento facial, o reconhecimento expressões faciais, são algumas das maneiras pelas quais a IA tem sido usado em vigilância nos EUA, China e muitos outros países

AO MENOS 75 DE 176
PAÍSES GLOBALMENTE
ESTÃO ATIVAMENTE
USANDO IA TECNOLOGIAS
PARA FINS DE VIGILANCIA



51%
DE AVANÇADO
NA
DEMOCRACIA

37%
FECHADO
ESTADOS
AUTOCRÁTICO



Os dados acima são
de AIGS - Índice
Vigilância IA Global

O índice não distingue entre IA de vigilância usado para propósitos legítimos e ilegalidade de vigilância digital.



O FUTURO DE IA

25

O PAPEL DE DECISORES





”

Quando chegarmos à década de 2040, seremos capazes de multiplicar nossa inteligência de humanos em mais de um bilhão de vezes. Que será uma mudança profunda que é singular na natureza. Os computadores vão continuar transformando-se em tamanhos menores e menores.

Em última análise, eles serão tão pequenos que estarão dentro dos nossos corpos e cérebros e nos farão mais saudáveis, e mais espertos.

Ray Kurzweil

Diretor de Engenharia no Google,
Cofundador da Universidade de
Singularidade

O FUTURO DE IA

O QUE É O PAPEL DE DECISORIOS ?

Prevê-se que as máquinas sejam melhores que os humanos na tradução de línguas, em 2024, escrevendo redações para o ensino médio até 2026, dirigindo um caminhão até 2027, trabalhando em varejo em 2031, escrevendo um livro best-seller em 2049 e executando cirurgias sem acompanhamento em 2053. Na verdade há 50% de chance de que as máquinas superem os humanos em todas as tarefas dentro de 45 anos, dizem os entrevistados da pesquisa realizada pela Universidade de Oxford e Universidade de Yale em 2020. **A pesquisa** foi distribuída para 352 especialistas em IA, que compartilharam as suas opiniões sobre o futuro da IA.

Nem todos os especialistas concordam quando e se a superinteligência de IA será desenvolvida e como ele coexistirá com os humanos. Os pesquisadores colocaram apenas 5 por cento chance em computadores trazendo resultados próximos ao nível de extinção humana. Um quarto dos pesquisadores de IA pesquisados por Etzioni em 2016 disseram que pensavam inclusive que a superinteligência nunca se materializaria.

[Um estudo de Oxford afirmou que 47% dos empregos corriam o risco de se tornarem automatizados nos próximos 20 anos, o que aumentou a preocupação com a perda de empregos. Ainda assim, outros especialistas afirmam que a IA e a automação impactarão principalmente indústrias específicas, enquanto potencialmente criando mais empregos no processo. Todos estes temas e outros mencionados neste aqui precisam estar na agenda dos decisores.](#)



O PAPEL DOS DECISORES

Os decisores têm a responsabilidade de se envolverem no desenvolvimento da IA, planejamento as discussões. A IA é profundamente difundida e influencia muitas esferas de nossas vidas, impactando a segurança, a democracia, liberdade, direitos humanos.

Parar a IA não é uma **opção** e **precisamos ser mais rápidos, compreender esse desenvolvimento e impacto, bem como moldar e fazê-lo de maneira ética, segura e responsável por meio de elaboração de políticas**. O futuro da IA não é aquele com que sonhamos ou o u nós tememos, mas um que planejamos para e um que nós cuidadosamente projetamos para. IA pode ajudar a criar um mundo muito melhor, se nós construirmos de forma certa.

IA Curiosidad & Verdades



É preciso haver mais cruzamentos e falas entre indústria, sociedade Civil e organizações de meio acadêmico que trabalham para o avanço dessas tecnologias e as instituições governamentais que vão representar eles.

Terá Lyon , Diretor Executivo,
Segura IA

Aprendizado de Máquina: A maior parte do avanço recente em IA é devido ao aprendizado de máquina, uma técnica que permite aos sistemas aprenderem e melhorarem automaticamente a partir da experiência, sem serem explicitamente programados para cada tarefa.

Redes Neurais: As redes neurais artificiais, inspiradas na estrutura do cérebro humano, são uma tecnologia chave por trás da IA moderna. Elas são compostas por camadas de neurônios artificiais que processam dados e aprendem padrões complexos.

IA e Saúde: A IA está revolucionando a área da saúde, ajudando no diagnóstico precoce de doenças, personalização de tratamentos e descoberta de novos medicamentos. Algoritmos de IA podem analisar imagens médicas com uma precisão muitas vezes superior à dos humanos.

Deep Learning: Um subsetor do aprendizado de máquina, o deep learning utiliza redes neurais profundas com muitas camadas, permitindo a IA reconhecer padrões em grandes volumes de dados, como imagens e texto, com precisão impressionante.

Processamento de Linguagem Natural (PLN): Esta área da IA se concentra em permitir que as máquinas compreendam e respondam à linguagem humana. Aplicações incluem assistentes virtuais como Siri e Alexa, tradução automática e chatbots.

IA em Jogos: A IA é usada extensivamente em jogos para criar adversários inteligentes, ajustar a dificuldade com base no desempenho do jogador e até para gerar conteúdo proceduralmente, como níveis e mundos de jogo.

Ética e IA: O desenvolvimento e implementação de IA levantam questões éticas significativas, incluindo o viés algorítmico, a privacidade dos dados, o desemprego tecnológico e a necessidade de regulamentação para garantir que a IA seja usada de maneira justa e segura.

Impacto Econômico: A IA tem o potencial de transformar diversas indústrias, desde manufatura até serviços financeiros, aumentando a eficiência e possibilitando novas formas de inovação.

IA Curiosidad & Verdades

Médicos robóticos serão capazes de atender cerca de 3 mil pacientes por dia a partir do segundo semestre deste ano

China revela que terá o primeiro hospital de inteligência artificial do mundo.

O "Hospital Agente", desenvolvido pela Universidade Tsinghua, em Pequim, terá a capacidade de economizar "milhões" com a sua interação automatizada. Liu Yang, líder da equipe que desenvolve a instalação, declarou que o hospital de IA trará imensos benefícios tanto para os profissionais médicos quanto para os pacientes, segundo reportagem no "Global Times". De acordo com estudos, seriam necessários cerca de dois anos para que os médicos humanos tivesse a mesma performance de atendimento.

Testes conduzidos por pesquisadores chineses já mostraram que os agentes médicos de IA, alimentados por grandes modelos de linguagem da medicina, alcançam uma impressionante taxa de precisão de 93,06%. Yang acredita que a unidade esteja operacional no segundo semestre de 2024.

**BEM-VINDOS A
PERGUNTAS E
COMENTÁRIOS SOBRE
O BATE-PAPO**

**VOCÊ PODE ENTRAR EM CONTATO
DIRETAMENTE NO**

