



O futuro com a Embedded Gen AI



The IT Framework
www.ciocodex.com

The Future with Embedded GenAI

Amiabile - ai

Generate the future, abundance of technology, intelligence is in our daily life.

Essa aqui eu acho que será para estourar de vez a bolha tech no uso de AI. Agora AI virou “pop”!

E eu não me refiro a “pop” no sentido de fama a partir de filmes de ficção científica,

mas pop no sentido de adoção efetiva no uso pelas pessoas.

Apenas à título de comparação, o Facebook levou 10 meses para alcançar 1 milhão de usuários. O Instagram menos de 3 meses. O ChatGPT levou 5 dias para alcançar 1 milhão de usuários. Em cerca de 2 meses ele já tinha 100 milhões de usuários!

Uma velocidade de aquisição de usuários nessa magnitude é coisa de serviço digital mainstream, não de ferramenta de tecnologia.

Como eu falei, estourou a bolha tech e virou algo pop, dado que as pessoas já fazem piadas e paródias usando o conceito de uso do ChatGPT e AI Generativa como um todo.

A GenAI embutida no Microsoft Office

Agora imaginem toda a adoção e evolução apontada acima, além da transformação da forma como trabalhamos, mas só que elevada em algumas potências com a disponibilidade embutida no Office, usado diariamente por sei lá quantas centenas de milhões de pessoas ao redor do mundo!

Creio que é o que podemos ver em breve, segundo essa matéria da ComputerWorld:

<https://www.computerworld.com/article/3691109/microsoft-365-copilot-chatbot-is-the-ai-based-future-of-work.html>

O uso de AI no campo conversacional

A inteligência artificial (IA) tem avançado rapidamente nos últimos anos, especialmente na área de processamento de linguagem natural (PLN), que permite que as máquinas entendam e gerem textos em linguagem humana.

Um dos desafios mais difíceis e fascinantes do PLN é a criação de sistemas capazes de

manter conversas naturais, coerentes e envolventes com os humanos.

Esses sistemas, chamados de agentes conversacionais ou chatbots, têm aplicações potenciais em diversos domínios, como educação, entretenimento, saúde, comércio e assistência pessoal.

Embora existam vários chatbots disponíveis no mercado, a maioria deles se baseia em regras pré-definidas ou em respostas pré-fabricadas, o que limita a sua capacidade de lidar com situações imprevistas, de se adaptar às preferências dos usuários e de gerar respostas criativas e diversificadas.

Além disso, muitos chatbots são especializados em um domínio específico, como reservar passagens aéreas ou pedir comida, e não conseguem manter conversas abertas sobre temas variados e de interesse geral.

ChatGPT e a revolução da AI no campo do diálogo

Uma nova abordagem para o desenvolvimento de chatbots é o uso de modelos de aprendizagem profunda, que podem aprender padrões estatísticos a partir de grandes quantidades de dados textuais, sem a necessidade de regras explícitas ou conhecimento prévio.

Esses modelos, chamados de modelos generativos, podem produzir respostas originais e adequadas ao contexto, usando apenas os textos das conversas anteriores como entrada.

Um exemplo de modelo generativo é o ChatGPT, desenvolvido pela empresa OpenAI, que usa uma arquitetura de redes neurais chamada Transformer, capaz de capturar relações complexas entre palavras e frases.

O que é o ChatGPT?

O ChatGPT é um chatbot baseado no modelo GPT-3, um dos mais avançados e poderosos modelos de PLN da atualidade.

O GPT-3 foi treinado com mais de 175 bilhões de parâmetros, usando um conjunto de dados gigantesco, que inclui quase toda a internet, como livros, artigos, blogs, redes sociais, sites, etc. O GPT-3 é capaz de realizar diversas tarefas de PLN, como responder perguntas, escrever textos, traduzir idiomas, resumir documentos, completar frases, entre outras, usando apenas um modelo único e uma técnica chamada de adaptação de poucos exemplos (few-shot learning), que consiste em fornecer algumas instruções ou exemplos da tarefa desejada como entrada, juntamente com o texto a ser processado.

O ChatGPT é uma versão do GPT-3 especializada em diálogo, que foi treinada com milhões de conversas reais sobre diversos assuntos, onde se destacam algumas habilidades:

- O ChatGPT aprendeu a gerar respostas coerentes, relevantes e interessantes, imitando o estilo e o tom das conversas humanas.
- O ChatGPT também pode se ajustar ao perfil e às preferências do usuário, usando informações contextuais, como idade, gênero, localização, hobbies, etc.
- O ChatGPT pode conversar sobre uma ampla variedade de tópicos, desde assuntos triviais até questões filosóficas, passando por cultura pop, política, ciência, humor, etc.
- O ChatGPT também pode realizar algumas tarefas dentro da conversa, como contar piadas, fazer elogios, dar conselhos, jogar jogos, etc.

O salto do GPT-3 para o GPT-4

O GPT-4 é um modelo de linguagem neural que supera o seu antecessor, o GPT-3, em vários aspectos.

O GPT-4 tem uma capacidade maior, podendo processar mais dados e gerar textos mais longos e coerentes.

O GPT-4 também incorpora novas técnicas de aprendizagem, como a atenção esparsa e a codificação contrastiva, que permitem uma melhor compreensão do contexto e uma maior diversidade de respostas.

Um dos principais fatores que diferenciam o GPT-4 do GPT-3 é o número de parâmetros que cada modelo possui.

O GPT-3 já era considerado um modelo gigantesco, com 175 bilhões de parâmetros, mas o GPT-4 supera esse número em mais de dez vezes, chegando a 2,3 trilhões de parâmetros.

Isso significa que o GPT-4 pode absorver mais informações e aprender mais padrões de linguagem do que o GPT-3, o que resulta em uma maior capacidade de gerar textos fluentes, lógicos e criativos.

Ao usar o ChatGPT, as pessoas podem notar essas melhorias na qualidade e na variedade das conversas que podem ter com o chatbot.

O ChatGPT pode acompanhar os usuários em diversos assuntos, oferecendo informações relevantes, opiniões interessantes e comentários divertidos.

O ChatGPT também pode se ajustar ao estilo e ao humor dos usuários, criando uma experiência mais personalizada e satisfatória. Além disso, o ChatGPT pode realizar tarefas mais complexas dentro da conversa, como ensinar algo novo, resolver problemas, contar histórias, etc.

O ChatGPT é, portanto, um exemplo de como o salto tecnológico do GPT-3 para o GPT-4 pode trazer benefícios tangíveis para as pessoas que interagem com a IA conversacional.

Ao comparar as conversas geradas pelo GPT-3 e pelo GPT-4, as pessoas podem perceber uma diferença significativa na riqueza, na coerência e na diversidade dos

diálogos, o que demonstra o potencial da tecnologia para criar interações mais naturais, dinâmicas e humanas.

Por que o ChatGPT é revolucionário?

O ChatGPT é considerado um marco na área de IA conversacional, pois demonstra que é possível criar chatbots de alta qualidade usando apenas modelos generativos, sem a necessidade de módulos específicos para cada domínio ou tarefa.

O ChatGPT também mostra que os chatbots podem ser versáteis, capazes de lidar com diferentes tipos de situações e de se adaptar às preferências dos usuários.

Além disso, o ChatGPT evidencia o potencial dos modelos de PLN baseados em Transformer, que podem aprender representações linguísticas ricas e flexíveis, que permitem realizar diversas tarefas com um único modelo e com pouca supervisão humana.

O ChatGPT também tem um impacto social significativo, pois abre as portas para uma nova forma de interação entre humanos e máquinas, que pode trazer benefícios para diversas áreas e setores da sociedade.

Por exemplo, o ChatGPT pode ser usado como uma ferramenta educativa, que auxilia os estudantes a aprenderem novos conteúdos, a praticarem idiomas e a desenvolverem habilidades socioemocionais.

O ChatGPT também pode ser usado como uma fonte de entretenimento, que proporciona diversão, lazer e criatividade aos usuários.

O ChatGPT também pode ser usado como um agente terapêutico, que oferece apoio emocional, orientação e companhia aos usuários, especialmente em tempos de isolamento social e de crise sanitária>

Quais são os desafios e as limitações do ChatGPT?

Apesar de ser um avanço impressionante na área de IA conversacional, o ChatGPT ainda enfrenta alguns desafios e limitações, que precisam ser superados para que ele possa ser usado de forma eficaz e responsável. Alguns desses desafios e limitações são:

- **A consistência:** o ChatGPT pode gerar respostas contraditórias ou incoerentes ao longo da conversa, especialmente se ela for muito longa ou complexa. Isso pode prejudicar a credibilidade e a confiança do usuário no chatbot.
- **A segurança:** o ChatGPT pode gerar respostas ofensivas, inadequadas ou perigosas, que podem violar as normas éticas, sociais ou legais. Isso pode causar danos morais, emocionais ou físicos ao usuário ou a terceiros.
- **A veracidade:** o ChatGPT pode gerar respostas falsas, imprecisas ou enganosas, que podem confundir ou induzir o usuário ao erro. Isso pode comprometer a qualidade e a validade da informação fornecida pelo chatbot.
- **A originalidade:** o ChatGPT pode gerar respostas plagiadas, repetitivas ou banais, que podem diminuir o interesse e a satisfação do usuário com a conversa. Isso pode reduzir a criatividade e a diversidade das respostas produzidas pelo chatbot.

Para enfrentar esses desafios e limitações, é necessário desenvolver métodos e mecanismos que possam melhorar o desempenho, a confiabilidade e a segurança do ChatGPT, bem como monitorar e avaliar o seu uso e o seu impacto na sociedade.

Além disso, é preciso estabelecer princípios e diretrizes éticas, que orientem o desenvolvimento e a aplicação do ChatGPT, respeitando os valores e os direitos humanos.

O que é Inteligência Artificial Generativa (GenAI)?

A Inteligência Artificial Generativa (GenAI) refere-se a uma categoria de tecnologia de IA que é capaz de criar conteúdo novo e único, a partir do aprendizado de uma vasta quantidade de dados existentes.

Diferente da IA tradicional que se baseia em análise e processamento de informações para fornecer resultados ou otimizar processos, a GenAI inova ao gerar textos, imagens, música, código de programação e até respostas em diálogos, que podem ser indistinguíveis dos criados por humanos.

Esta capacidade é impulsionada principalmente por modelos de aprendizado profundo, como redes neurais, que analisam e assimilam padrões complexos.

Grandes players do mercado

O mercado de inteligência artificial está em constante expansão e inovação, com vários players importantes disputando liderança e influência.

Cada um desses players traz suas próprias inovações e abordagens únicas para a inteligência artificial, refletindo a diversidade e a complexidade desse campo em rápida evolução.

Enquanto exploram novas fronteiras tecnológicas, também enfrentam questões críticas de ética, privacidade e aplicabilidade que definirão o futuro da IA.

Vamos explorar alguns dos principais concorrentes neste campo, analisando suas fortalezas e debilidades.

OpenAI e ChatGPT

Fortalezas: ChatGPT, desenvolvido pela OpenAI, ganhou destaque pela sua habilidade em compreender e responder perguntas em linguagem natural, fazendo-o extremamente popular para aplicações que vão desde assistentes pessoais até ferramentas educacionais.

A OpenAI também é conhecida por sua ética em IA e pesquisa abrangente, contribuindo significativamente para o avanço da segurança em IA.

Debilidades: Apesar de sua capacidade avançada, o ChatGPT pode gerar respostas imprecisas ou fabricadas, e há preocupações sobre o uso de seus modelos para gerar desinformação.

Microsoft e Copilot

Fortalezas: Com o lançamento do Copilot, a Microsoft integrou capacidades de IA nos seus produtos de software, como o Office e o GitHub, promovendo uma grande sinergia entre IA e produtividade.

A Microsoft tem vastos recursos para pesquisa e um ecossistema de aplicativos bem estabelecido que potencializa o alcance de suas soluções de IA.

Debilidades: O Copilot enfrenta desafios de privacidade e segurança de dados, essenciais para a aceitação nos ambientes empresariais, além de depender significativamente das capacidades de nuvem da Microsoft, o que pode limitar sua aplicabilidade em ambientes offline.

Google e Gemini

Fortalezas: O Gemini da Google é projetado para ser um modelo de linguagem avançado que melhora a compreensão de contexto e a geração de texto.

A Google, com seu robusto histórico em pesquisa e desenvolvimento em IA, leva vantagem em integrar seus modelos de IA com seu motor de busca e outras ferramentas online.

Debilidades: Ainda que potente, o Gemini pode enfrentar questões relacionadas à privacidade e à ética, semelhantes aos desafios enfrentados por outras tecnologias de IA da empresa.

Meta (antiga Facebook)

Fortalezas: As soluções de IA da Meta são focadas em melhorar interações sociais, moderação de conteúdo e realidade virtual.

A empresa é pioneira na pesquisa de IA para realidade aumentada e virtual, posicionando-se fortemente no metaverso.

Debilidades: A Meta enfrenta críticas e desafios legais significativos quanto ao tratamento de dados de usuários e ética na IA, especialmente no que tange à privacidade e ao uso de dados para treinamento de seus modelos.

IBM

Fortalezas: A IBM, com seu Watson, foi uma das pioneiras em IA comercial, aplicando a tecnologia em áreas como saúde e finanças. A empresa tem forte presença em IA empresarial, com capacidades robustas de análise de dados e aprendizado de máquina.

Debilidades: O Watson tem enfrentado críticas por não atender às expectativas em alguns setores, e a IBM tem desafiado a manter sua liderança diante de concorrentes mais ágeis e inovadores.

xAI

Fortalezas: A recém-lançada xAI propõe uma nova abordagem para entender fenômenos complexos do universo através da IA. Com forte financiamento e uma visão ambiciosa, espera-se que a xAI introduza inovações disruptivas.

Debilidades: Sendo uma novidade, a xAI enfrenta o desafio de estabelecer sua credibilidade e aplicabilidade prática, além de potenciais questões éticas associadas às ambições de seus projetos.

Principais Tendências de Mercado

A adoção da GenAI está crescendo exponencialmente, com várias tendências emergindo:

- **Personalização em Massa:** Empresas usam GenAI para criar experiências personalizadas para os usuários, desde recomendações de produtos até conteúdo personalizado.
- **Automação de Design e Conteúdo:** Setores de marketing e design gráfico utilizam GenAI para gerar imagens, vídeos e textos, reduzindo custos e aumentando a eficiência.
- **Desenvolvimento de Software Assistido por AI:** GenAI está ajudando programadores a escrever e revisar códigos, acelerando o desenvolvimento de software.
- **Ética e Regulação:** Conforme a GenAI se torna mais prevalente, cresce o foco em criar normas éticas e regulatórias para seu uso adequado.

Expectativas para o Futuro da GenAI

As expectativas em torno da GenAI são altamente positivas e ambiciosas:

- **Expansão da Capacidade Criativa:** Acredita-se que a GenAI ampliará as capacidades criativas humanas, permitindo a criação de obras de arte, literatura e inovações técnicas a um ritmo antes inimaginável.
- **Colaboração Homem-Máquina:** Prevê-se uma colaboração cada vez maior entre humanos e máquinas, onde a GenAI servirá como uma ferramenta de ampliação das capacidades humanas, não apenas substituindo tarefas.
- **Democratização da Criação de Conteúdo:** Com ferramentas de GenAI, indivíduos e pequenas empresas terão poder para gerar conteúdos de qualidade comparável às

grandes corporações.

Principais Desafios

Apesar das grandes promessas, a GenAI enfrenta vários desafios significativos:

- **Questões Éticas e de Direitos Autorais:** A geração de conteúdo que parece autêntico levanta questões sobre originalidade e propriedade intelectual.
- **Viés:** Os dados usados para treinar modelos de GenAI podem conter vieses, resultando em saídas também enviesadas.
- **Segurança e Privacidade:** As implicações de segurança da GenAI são profundas, especialmente se usada para gerar desinformação ou conteúdo prejudicial.
- **Impacto no Emprego:** Existe a preocupação de que a GenAI possa deslocar trabalhos, especialmente na criação de conteúdo e design.

Concluindo

Considero que o ChatGPT é uma ferramenta poderosa e inovadora e que já está transformando a forma como interagimos e atuamos no trabalho e vida pessoal.

Ao vermos esse tipo de tecnologia sendo cada vez mais embutida nas plataformas de produtividade que usamos cotidianamente, é de se esperar uma transformação ainda mais acelerada.

Aqui comentei o exemplo da Microsoft, mas o conceito em si se enquadra em outras plataformas, como o Google, assim como muito provavelmente teremos novidades nesse sentido com Apple e outros vendedores.

Mas ao mesmo tempo que existe o lado bonito da história, existem também os seus

desafios e limitações, que ainda devem ser superados.

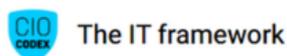
Acredito que é essencial investir na pesquisa e no desenvolvimento de soluções que possam tornar o ChatGPT mais eficiente, confiável e seguro, assim como monitorar e avaliar o seu uso e o seu impacto na sociedade.

O ChatGPT pode ser uma grande oportunidade para melhorar a comunicação e a interação entre as pessoas, mas também exige cautela e consciência dos seus riscos e desafios.



Arthur De Santis

Arthur De Santis é um executivo com mais de 20 anos de atuação na indústria de serviços financeiros, com destaque para bancos, processadoras de cartões, adquirentes e seguradoras, formando e liderando equipes e iniciativas ao longo de toda a cadeia de valor de Tecnologia da Informação.



O conteúdo apresentado neste website, incluindo o framework, é protegido por direitos autorais e é de propriedade exclusiva do CIO Codex. Isso inclui, mas não se limita a, textos, gráficos, marcas, logotipos, imagens, vídeos e demais materiais disponíveis no site. Qualquer reprodução, distribuição, ou utilização não autorizada desse conteúdo é estritamente proibida e sujeita às penalidades previstas na legislação aplicável