

Relatório de Participação de Henrique Faulhaber na Gartner CIO & IT Executive Conference 2024

De 23 a 25 de setembro de 2024, participei da Gartner CIO & IT Executive Conference 2024, realizada no Sheraton São Paulo WTC Hotel. Esse evento, considerado um dos mais relevantes do setor de TI no Brasil, proporcionou uma imersão em tendências emergentes e estratégias inovadoras no campo da tecnologia da informação.

A conferência destacou temas como Inteligência Artificial, cibersegurança, modernização de aplicações e transformação digital e foi uma excelente oportunidade para discutir inovações e estabelecer conexões com mais de 1.500 CIOs e líderes de TI. Além das sessões plenárias, houve workshops e mesas-redondas que proporcionaram uma visão aprofundada dos desafios atuais do setor.

Em minha participação, foquei em temas como adoção de Inteligência Artificial e as melhores práticas de cibersegurança e resiliência digital.

O primeiro keynote da conferência foi sobre Inteligência Artificial, teve o título "Definindo seu ritmo na corrida de IA" e foi conduzido por Álvaro Mello e Hang Lo Hang. Em síntese, os palestrantes falaram sobre a implementação da Inteligência Artificial (IA) nas organizações, particularmente no ambiente corporativo, que enfrenta uma série de desafios complexos que exigem atenção cuidadosa dos líderes de tecnologia.

O painel descreveu dois tipos de corridas relacionadas à IA. A primeira é uma competição tecnológica entre fornecedores e empresas do setor, cujo objetivo é inovar e lançar novos modelos de IA de maneira cada vez mais rápida. Essa corrida é intensificada pela pressão constante de estar na vanguarda da inovação tecnológica. A segunda corrida, mais relevante para os CIOs (executivos de Tecnologia da Informação das empresas), envolve a aplicação segura e escalável da IA dentro das organizações, um desafio intrincado que precisa ser tratado de maneira estratégica e cuidadosa.

Um dos principais desafios mencionados está relacionado ao alinhamento entre expectativas e resultados reais da IA. Muitos CIOs relatam que, embora a IA esteja em um ponto de alta expectativa, os resultados não têm atingido o retorno sobre o investimento (ROI) esperado. Esse descompasso ocorre em parte pela dificuldade de traduzir a inovação tecnológica em benefícios práticos para os negócios. A corrida pela implementação não se resume à adoção das tecnologias mais recentes, mas envolve garantir que elas estejam alinhadas aos objetivos da organização e sejam integradas ao fluxo de trabalho de uma maneira que realmente ofereça valor agregado. Essa pressão por resultados pode criar a sensação de que a organização está simultaneamente no ápice do hype da tecnologia, mas também presa em desafios operacionais.

Além disso, um ponto importante trazido no painel é a dificuldade em obter produtividade real a partir da IA. Embora muitas empresas vejam a IA como uma ferramenta para aumentar a eficiência dos funcionários, a integração das tecnologias de IA ao trabalho cotidiano nem sempre é intuitiva. Muitos líderes acreditam que os funcionários estarão

dispostos a adotar novas ferramentas de IA, mas a verdade é que a maioria enfrenta dificuldades para incorporar essas tecnologias de forma eficaz ao seu dia a dia. O impacto varia conforme o tipo de trabalho e a experiência dos funcionários. Para trabalhadores menos experientes, a IA pode oferecer um suporte significativo, enquanto para aqueles mais experientes, os benefícios podem ser mais limitados, especialmente em tarefas que já dominam.

Outro grande desafio está relacionado aos custos de implementação da IA e à sua escalabilidade. Investir em IA não é o mesmo que investir em tecnologias convencionais de TI, pois há uma imprevisibilidade significativa nos custos envolvidos. O painel alertou que os custos de escalabilidade da IA, especialmente em processos como inferência, preparação de dados e gestão de infraestrutura, podem ser subestimados em até 500% a 1.000%. Por isso, as organizações, segundo os palestrantes, precisam de um plano robusto para gerenciar esses custos, e uma das recomendações é que as provas de conceito (PoC) não sejam usadas apenas para demonstrar a viabilidade técnica, mas também para entender os custos e as necessidades de escalabilidade, transformando as provas de conceito em "provas de valor", ajudando a garantir que o investimento seja feito de forma consciente e sustentável.

A questão da segurança e governança da IA também é destacada como um elemento crítico para uma implementação bem-sucedida. À medida que as empresas aumentam o número de iniciativas de IA, a complexidade da gestão dessas tecnologias também cresce. Para organizações que têm um número menor de iniciativas, práticas de governança, como a criação de uma comissão central de IA e a implementação de comunidades de prática, podem ser suficientes para gerenciar riscos e manter o controle sobre as operações. No entanto, para aquelas que têm um número significativo de iniciativas, essas práticas precisam ser mais robustas para evitar a proliferação descontrolada de tecnologias de IA e os riscos associados à sua má utilização..

Foi demonstrado no painel que os desafios da implementação da IA no ambiente corporativo vão muito além da simples aquisição da tecnologia, pois envolvem desde a escolha de uma estratégia de adoção e a busca por resultados práticos até a necessidade de lidar com custos imprevisíveis e assegurar uma governança que garanta o uso responsável e seguro dessas ferramentas.

Outra palestra interessante foi sobre a importância da infraestrutura tecnológica na IA generativa, conduzida por Jonathan Forest, analista do Gartner. A palestra abordou os desafios e estratégias relacionadas à implementação da infraestrutura para IA generativa em ambientes corporativos, destacando a importância de entender os requisitos de computação, armazenamento e rede, além dos custos associados, para garantir o sucesso na adoção da IA.

Durante a palestra, foram destacados os diferentes tipos de modelos de IA, como grandes modelos de linguagem (LLMs) e modelos menores, cada um com necessidades distintas de infraestrutura. Foi enfatizado que as organizações precisam decidir se irão utilizar um modelo pronto, como o ChatGPT, que exige infraestrutura mínima, ou se irão investir em treinar e ajustar modelos personalizados, o que demanda capacidades mais robustas de computação e habilidades especializadas em ciência de dados.

Decisões sobre a escolha da infraestrutura têm impactos significativos sobre a execução dos projetos, como falhas de desempenho, problemas de escalabilidade e custos inesperados. Além disso, a falta de habilidades adequadas em IA dentro da equipe pode prejudicar o desenvolvimento dos projetos. A implementação de IA também requer uma distinção clara entre o treinamento de modelos e a inferência.

O treinamento de modelos envolve processos complexos e centralizados, muitas vezes realizados em data centers ou na nuvem, enquanto a inferência, que diz respeito à interação constante do modelo com os usuários, é geralmente distribuída e pode ocorrer em ambientes cloud e edge. Assim, a infraestrutura para inferência precisa ser adaptada para oferecer respostas rápidas e estáveis, atendendo às necessidades de latência e volume de dados.

A escolha entre usar a nuvem, um data center local ou uma infraestrutura edge depende de diversos fatores, como custos, habilidades internas, privacidade de dados e soberania. Organizações que não possuem habilidades internas suficientes geralmente optam pela nuvem, que oferece maior flexibilidade e acesso a ferramentas prontas. No entanto, aquelas que possuem mais experiência em IA e têm requisitos rigorosos de privacidade de dados preferem soluções on-premise, garantindo maior controle sobre suas operações.

Sobre as tendências futuras, a palestra apontou que haverá um aumento na utilização de data centers e infraestruturas edge, principalmente devido à necessidade crescente de privacidade e soberania dos dados. Com o tempo, as organizações precisarão ser flexíveis para se mover entre ambientes de nuvem, data center e edge, criando uma abordagem híbrida que permita atender às diferentes necessidades de cada aplicação e garantindo a eficiência na gestão de suas infraestruturas de IA.

Assisti também a uma palestra sobre oportunidades para o uso de IA relacionadas à cibersegurança das organizações, conduzida por Oscar Isaka do Gartner, que trouxe uma visão sobre o futuro da cibersegurança em um mundo cada vez mais marcado pela volatilidade, incerteza, complexidade e ambiguidade. A palestra enfatizou o uso da IA para aumentar a capacidade dos profissionais de segurança, e não substituí-los. Ele também defendeu o uso de tecnologias como "Zero Trust", Gestão de Identidades e a luta contra a desinformação como fundamentais para mitigar ameaças emergentes e fortalecer as operações de segurança digital.

Na palestra "10 Boas Práticas para Escalar a IA Generativa", de Anthony Mullenn, foram abordadas as melhores estratégias para implementar e ampliar o uso da IA generativa em uma organização. Ele destacou tanto o potencial transformador dessa tecnologia quanto os desafios que as empresas enfrentam ao integrá-la em seus processos. A palestra enfatizou a importância de priorizar casos de uso estratégicos.

Foi citado que muitas iniciativas de IA generativa falham devido a problemas de qualidade dos dados, controles de risco insuficientes, custos elevados ou uma falta de clareza no valor que essa tecnologia pode agregar ao negócio.

Em termos de gerenciamento de dados e riscos, a IA generativa exige dados de alta qualidade, mas muitas empresas ainda operam com práticas de dados inadequadas ao potencial da IA. Além disso, a tecnologia traz novos riscos, como alucinações dos modelos

e deepfakes, que representam ameaças à integridade da informação e à reputação das empresas. Portanto, é essencial estabelecer uma abordagem robusta de governança de dados para sustentar a escalabilidade da IA.

No contexto da adoção da IA generativa, o palestrante abordou que a maioria dos projetos ainda está em fase de experimentação e que poucos avançaram para a produção em larga escala, com investimentos concentrados em funções internas, como finanças, recursos humanos e TI. Estas escolhas visam reduzir riscos antes de expandir a IA para interações diretas com clientes, evitando incidentes como decisões incorretas de assistentes de IA, que resultaram em prejuízos significativos.

Também foram discutidas abordagens para decidir entre construir ou comprar soluções de IA. Projetos-piloto, além de testar a viabilidade técnica dos modelos, são importantes para avaliar o impacto sociotécnico — ou seja, como a tecnologia interage com as pessoas e processos existentes. Essa abordagem permite não só verificar a funcionalidade técnica da IA, mas também ajustar o modelo operacional de forma a maximizar sua adoção e eficácia.

Práticas de IA responsável são um dos pilares fundamentais discutidos. Elas incluem garantir a segurança, a privacidade e a responsabilidade nos sistemas de IA. Para isso, as organizações devem adotar práticas de governança rigorosas para lidar com os riscos da IA, como falsificações profundas e a complexidade da propriedade intelectual. Ferramentas para monitorar e garantir a segurança dos modelos, como inspeções de injeção e compliance, são recomendadas para assegurar a confiabilidade da IA.

A colaboração entre IA e seres humanos também foi identificada como um ponto crucial, pois a verdadeira maximização do valor da IA ocorre quando esta é utilizada como complemento das capacidades humanas, enriquecendo a eficiência e a produtividade.

O palestrante abordou o impacto crescente das regulamentações na adoção da IA generativa. Atualmente, as legislações ainda são limitadas, mas estão evoluindo rapidamente, especialmente em regiões como a União Europeia e a China, que já começaram a implementar regras específicas para a IA. As regulamentações impactarão diretamente os custos e a frequência de treinamento de modelos, pois as empresas precisarão garantir que seus sistemas cumpram requisitos legais, aumentando os custos de compliance e limitando a flexibilidade para usar certos tipos de dados. Isso também cria barreiras de entrada para novos desenvolvedores de IA, favorecendo grandes empresas que já possuem infraestrutura robusta para atender aos requisitos regulatórios.

Além disso, as regulamentações estão impulsionando o desenvolvimento de modelos específicos de domínio, especialmente em setores sensíveis como saúde e finanças. Esses modelos devem ser desenvolvidos de acordo com normas que garantam a privacidade e a segurança dos dados dos usuários. O documento também menciona que veremos um aumento na regulamentação de agentes de IA, capazes de interagir diretamente com sistemas de software e interfaces de usuário, elevando os riscos associados à automação e destacando a necessidade de regras claras para definir a responsabilidade em caso de falhas.

Na palestra "Impulsione a Adoção de IA Generativa pelos Funcionários de Forma Correta", apresentada por Alicia Mullery, foram abordadas estratégias para promover a adoção eficaz

da IA generativa nas organizações. Alicia enfatizou que a principal barreira para a adoção não está na resistência dos funcionários, mas na falta de simplicidade no uso das ferramentas. Muitas empresas experimentam um entusiasmo inicial com novas tecnologias, mas acabam retornando aos métodos tradicionais após algumas semanas. Ela argumentou que, para maximizar o potencial da IA generativa, é preciso torná-la acessível e relevante para os colaboradores.

Alicia compartilhou experiências de líderes empresariais que estão focados em garantir que as equipes entendam e estejam envolvidas no processo de adoção da IA. Um ponto crítico identificado é a necessidade de se concentrar tanto nos aspectos importantes da tecnologia quanto nos fundamentos humanos, como a mudança de papéis e responsabilidades e o impacto na identidade dos trabalhadores. Além disso, pesquisas mostraram que muitos empregados não compreendem a visão estratégica da empresa para a IA, sentindo-se “no escuro” sobre seu impacto.

A abordagem sugerida envolve quatro passos principais: colaborar com os empregados na criação dos novos papéis que a IA irá transformar, personalizar as medições de produtividade de acordo com os diferentes papéis, estabelecer recursos dedicados para suporte e adotar uma mentalidade que valorize a capacitação e o apoio às equipes. A criação de comunidades de prática e a presença de especialistas dedicados são elementos fundamentais para ajudar os funcionários a entender e adotar essas novas tecnologias, levando a um ambiente de trabalho mais produtivo e menos ansioso.

Conclusão

Minha participação no Fórum da Gartner foi extremamente proveitosa, pois pude aprofundar meus conhecimentos sobre os desafios atuais da adoção da inteligência artificial em organizações do ponto de vista dos gestores de negócios e de tecnologia da informação. Esses conhecimentos são de grande importância para o trabalho do GT IA do CGI.br, em que uma de suas vertentes é o estímulo à adoção de IA no país de forma responsável e ética.