

O Papel da Governança de Dados na Implementação de IA Generativa, um estudo de caso da adoção por empresas digitais brasileiras.

Jessica Miyazaki Nogueira

Trabalho de Conclusão de Curso
MBA em Inteligência Artificial e Big Data

UNIVERSIDADE DE SÃO PAULO

Instituto de Ciências Matemáticas e de Computação

O Papel da Governança de Dados na
Implementação de IA Generativa, um estudo de
caso da adoção por empresas digitais brasileiras.

Jessica Miyazaki Nogueira

Jessica Miyazaki Nogueira

O Papel da Governança de Dados na Implementação de IA Generativa, um estudo de caso da adoção por empresas digitais brasileiras.

Trabalho de conclusão de curso apresentado ao Departamento de Ciências de Computação do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo - ICMC/USP, como parte dos requisitos para obtenção do título de Especialista em Inteligência Artificial e Big Data.

Área de concentração: Inteligência Artificial

Orientador: Prof. Dr. Alneu de Andrade Lopes

USP - São Carlos

2024

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

N778p Nogueira, Jessica Miyazaki
O Papel da Governança de Dados na Implementação de
IA Generativa, um estudo de caso da adoção por
empresas digitais brasileiras. / Jessica Miyazaki
Nogueira; orientador Prof. Dr. Alneu de Andrade
Lopes Lopes. -- São Carlos, 2024.
47 p.

Trabalho de conclusão de curso (MBA em
Inteligência Artificial e Big Data) -- Instituto de
Ciências Matemáticas e de Computação, Universidade
de São Paulo, 2024.

1. . I. Lopes, Prof. Dr. Alneu de Andrade Lopes,
orient. II. Título.

DEDICATÓRIA

*Ao meu esposo pela compreensão,
carinho e apoio incansável.*

AGRADECIMENTOS

Ao Professor Doutor Alneu, por sua empatia e gentileza nas recomendações e compreensão quanto as dificuldades particulares apresentadas durante o desenvolvimento desse trabalho. A Professora Doutora Solange por seu apoio, liderança, acolhimento e aconselhamento fundamentais para a conclusão desse trabalho. Aos colegas de curso e todos os colaboradores que foram extremamente generosos.

RESUMO

NOGUEIRA, Jessica Miyazaki. **O Papel da Governança de Dados na Implementação de IA Generativa, um estudo de caso da adoção por empresas digitais brasileiras.** 2024. Trabalho de conclusão de curso (MBA em Inteligência Artificial e Big Data) do Departamento de Ciências de Computação do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo - ICMC/USP, 2024. Este trabalho explora a relevância da inovação e da inteligência artificial generativa (IA generativa) no ambiente corporativo, abordando os impactos e desafios que essa tecnologia traz para a competitividade empresarial. Desde o lançamento do ChatGPT em 2022, a IA generativa tem ganhado ampla adoção e revelado seu potencial para transformar operações e processos de negócios. A rápida integração dessa tecnologia em diversas indústrias demonstra que, para manter sua competitividade, as empresas precisam adotar novas abordagens tecnológicas com agilidade e responsabilidade. A pesquisa se concentra especialmente na governança de dados como um fator crucial para o sucesso na implementação da IA gen. A qualidade dos dados utilizados e a gestão eficiente desses recursos impactam diretamente o desempenho das organizações que adotam essa tecnologia. O trabalho explora estudos de caso de três empresas, analisando como a governança de dados influencia os diferentes estágios de adoção da IA generativa, categorizando essas empresas em níveis de maturidade — desde amador até proficiente. O estudo conclui que, embora a IA generativa ofereça oportunidades significativas de inovação e eficiência, sua implementação enfrenta desafios críticos, sobretudo relacionados à gestão de dados. A pesquisa destaca a governança de dados como um determinante essencial para o sucesso organizacional na era digital, promovendo uma compreensão das práticas que melhor suportam a eficácia e o impacto da IA generativa nas operações empresariais.

ABSTRACT

NOGUEIRA, Jessica Miyazaki. **The Differential Role of Data Governance in the Implementation of Generative AI, Case Study in Brazilian Digital Companies.** 2024. Completion of course work (MBA in Artificial Intelligence and Big Data) – Institute of Mathematics and Computer Science, University of São Paulo, São Carlos, 2024. This paper explores the relevance of innovation and generative artificial intelligence (generative AI) in the corporate environment, addressing the impacts and challenges that this technology brings to business competitiveness. Since the launch of ChatGPT in 2022, generative AI has gained widespread adoption and revealed its potential to transform operations and business processes. The rapid integration of this technology in various industries demonstrates that, to remain competitive, companies need to adopt new technological approaches with agility and responsibility. The research focuses specifically on data governance as a crucial factor for the success of the implementation of generative AI. The quality of the data used, and the efficient management of these resources directly impact the performance of organizations that adopt this technology. The paper explores case studies of three companies, analyzing how data governance influences the different stages of adoption of generative AI, categorizing these companies into maturity levels — from amateur to proficient. The study concludes that, although generative AI offers significant opportunities for innovation and efficiency, its implementation faces critical challenges, especially related to data management. The research highlights data governance as a key determinant of organizational success in the digital age, fostering an understanding of the practices that best support the effectiveness and impact of generative AI on business operations.

SUMÁRIO

INTRODUÇÃO	12
ADOÇÃO ACELERADA DE IA GENERATIVA	15
IMPACTO NOS NEGÓCIOS	15
DESAFIOS DE IMPLEMENTAÇÃO	16
GOVERNANÇA E QUALIDADE DE DADOS	16
ESTRATÉGIAS DE IMPLEMENTAÇÃO E HABILIDADES EMERGENTES	17
ORGANIZAÇÃO DO TRABALHO	17
DESAFIOS DE DADOS E GOVERNANÇA NA IMPLEMENTAÇÃO DE IA GENERATIVA EM TRABALHOS RELACIONADOS.	18
EM SERVIÇOS FINANCEIROS.	18
EM ATENDIMENTO AO CLIENTE.	18
EM DESENVOLVIMENTO DE PRODUTOS.	19
ADOÇÃO DE APLICAÇÕES DE IA GENERATIVA	19
HABILIDADES EMERGENTES	21
LACUNA IDENTIFICADA	21
METODOLOGIA DE ESTUDO DE CASO	22
ESTRATÉGIAS DO ESTUDO DE CASO	23
<i>Sistema de Classificação e Análise de Respostas</i>	23
ANALISE DE DOCUMENTOS INTERNOS	23
<i>Análise de Conteúdo e Convergência de Respostas</i>	23
<i>Critério de Seleção das Empresas</i>	24
COLETA DE DADOS E RESULTADOS	24
ESTUDO DE CASO DA EMPRESA A	24
<i>Questionário da Empresa A</i>	24
ESTUDO DE CASO DA EMPRESA B	25
<i>Questionário da Empresa B</i>	26
ESTUDO DE CASO DA EMPRESA C	27
<i>Questionário da Empresa C</i>	27
ANÁLISE	29
GOVERNANÇA DE DADOS EM LLMs EMPRESARIAIS	29
IMPORTÂNCIA DA GOVERNANÇA DE DADOS	30
ELEMENTOS-CHAVE DA GOVERNANÇA DE DADOS PARA LLMs	30
DESAFIOS NA GOVERNANÇA DE DADOS PARA LLMs	31
PROMPTING RAG	31
MECANISMO DE SOLICITAÇÃO RAG	31
BENEFÍCIOS DA SOLICITAÇÃO RAG	32
EFICIÊNCIA DE RECURSOS	32
TENDÊNCIAS FUTURAS NO PROMPT RAG	32
FINE-TUNING (AJUSTE-FINO)	33
EXISTEM VÁRIOS MÉTODOS E ESTRATÉGIAS EMPREGADOS NO PROCESSO DE FINE-TUNING, QUE PODEM INCLUIR:	33
<i>Completo</i>	33
<i>Com eficiência de parâmetros (PEFT)</i>	33
<i>Sequencial</i>	33
<i>Supervisionado</i>	34
<i>Ajuste de instruções</i>	34

APRENDIZADO POR REFORÇO A PARTIR DO FEEDBACK HUMANO (RLHF)	34
AVALIAÇÃO E IMPLANTAÇÃO	34
FERRAMENTAS E TECNOLOGIAS	35
META PROMPTING	35
ESTRUTURAS DE AJUSTE FINO	35
LLM OPS E GERENCIAMENTO DE DADOS	35
PROMPT ENGINEERING	36
GERENCIAMENTO DE DADOS:	36
AJUSTE FINO DO MODELO	36
FERRAMENTAS DE AVALIAÇÃO E MONITORAMENTO	36
MONITORAMENTO FUNCIONAL	36
MONITORAMENTO RÁPIDO	36
MONITORAMENTO DE RESPOSTA	36
TENDÊNCIAS EMERGENTES	36
DEMOCRATIZAÇÃO E ACESSIBILIDADE DA IA	37
INOVAÇÕES EM ESTRUTURAS DE AVALIAÇÃO	37
CONSIDERAÇÕES ÉTICAS E GOVERNANÇA DE DADOS	37
PERSPECTIVAS FUTURAS	37
ANÁLISE COMPARATIVA DAS EMPRESAS A, B E C	38
EMPRESA A: IMPLEMENTAÇÃO AVANÇADA	38
EMPRESA B: IMPLEMENTAÇÃO EM DESENVOLVIMENTO	38
EMPRESA C: IMPLEMENTAÇÃO EMBRIONÁRIA	38
COMPARAÇÃO E CLASSIFICAÇÃO	39
CLASSIFICAÇÃO FINAL:	39
CONCLUSÃO	40
NÍVEIS DE MATURIDADE NA IMPLEMENTAÇÃO DE IA	40
GOVERNANÇA DE DADOS COMO FATOR CRÍTICO	40
DESAFIOS E OPORTUNIDADES	41
TENDÊNCIAS EMERGENTES	41
IMPLICAÇÕES PARA A PRÁTICA	41
CONSIDERAÇÕES FINAIS	41
REFERÊNCIAS	42

INTRODUÇÃO

Desde o início do século XXI, as organizações têm enfrentado intensas pressões para inovar, a fim de manter sua competitividade e alcançar resultados que não apenas satisfaçam suas expectativas, mas também garantam sua sobrevivência em um contexto global acelerado pela era digital contemporânea. Essa era é caracterizada por rápidas e disruptivas mudanças tecnológicas do que em qualquer outro período anterior. A inovação tecnológica destaca-se como um dos fatores mais relevantes e fundamentais para a produção e prestação de serviços, promovendo processos internos cada vez mais eficientes. Atualmente, vivemos em uma era dominada por negócios digitais, onde a adaptação e a inovação contínua são essenciais para o sucesso organizacional (Inovação e qualidade de vida no trabalho: uma visão integrada da gestão a partir de estudo de caso na indústria petroquímica do Grande ABC, 2010). O comportamento de investir em tecnologia para trazer resultados maiores e mais rápidos pelas diversas industriais já é bem de perto acompanhado e conhecido, porém, a pandemia do Covid-19 mudou totalmente o comportamento da nossa sociedade, seu meio e suas relações. Esse acontecimento trouxe um fator importante para a priorização e desenvolvimento de determinadas tecnologias e pesquisas, não só pela motivação de inovar e se manter competitivo, mas, sim pela obrigatoriedade em se adaptar ao novo cenário a partir da aceleração na adoção de soluções digitais diversas. De acordo com a Gartner (2022) As empresas que negligenciam optando por manter os riscos relacionados com evolução/expansão de experiências imersivas, automação acelerada com inteligência artificial e entrega otimizada de tecnólogos como por exemplo para áreas de identidade, acesso, dados, segurança da informação, comunicação de pessoas e sistemas sem priorizar aquelas com potencial de máximo benefício para a sua organização mesmo que arriscado, perderiam os benefícios que se tem para os primeiros que a adotam. Esse tipo de recomendação provou-se pagar de volta quando se faz uma inovação por inovação quando comparado a acontecimentos de empresas que não fazem esse tipo de investimento e acabam adotando de qualquer maneira só que mais tarde de forma reativa, improvisadas e não programado. O Portal da FGV (2022) apontou que o uso e os gastos e investimentos em TI representa 8,7% da receita e continuava a crescer. *Essa expansão foi ainda mais notável em 2021 e 2022, em aspectos como valor, maturidade e importância para os negócios existentes e para viabilizar novos modelos de negócios. (FUNDAÇÃO GETULIO VARGAS, [2022]) Essas empresas que não adotam e não desenvolvem estratégias relacionadas a esse aspecto formam um grupo econômico relevante. Algumas julgadas dispensáveis porque, anteriormente, não eram consideradas expressivamente necessárias, como é o caso de plataformas de videoconferência como o Zoom ou o Google Meet. Essas empresas acabam se beneficiando de maiores oportunidades ao explorar intensamente a demanda de outras empresas por produtos inovadores e de fácil adoção. Isso se torna uma vantagem competitiva, já que concentram seus esforços exclusivamente para serem a opção mais óbvia nesses momentos. Esse tipo de gestão de TI acaba ficando sem alternativas quando pressionados a agir e por consequência de suas não escolhas antecipadas, se juntando ao grupo de empresas que adotaram pelo mesmo motivo e ao mesmo tempo determinada tecnologia e assim, causando o seguinte resultado, por exemplo: *“Zoom alcançou bons resultados financeiros no 1º trimestre de 2020. Segundo a*

companhia nesta terça-feira (2), a receita do serviço de chamadas de vídeo apresentou aumento de quase 200% conforme o uso da plataforma crescia devido aos impactos da COVID-19, doença causada pelo novo coronavírus”. (De Blasi Bruno, 2022)

<https://tecnoblog.net/noticias/zoom-receita-cresce-quase-200-durante-pandemia/> . O impacto profundo do período de tempo da COVID-19 eu penso que foi o elemento fundamental para difundir que as empresas devem evoluir e ganhar maturidade sobre os gastos de TI pela a importância que elas têm e são cada vez mais atribuídas como fundamentais para os negócios existentes e para viabilizar novos modelos de negócios. No começo de 2023, o cenário socioeconômico global estava cuidadosamente recuperando o fôlego após enfrentar altos e baixos em um curto período. Enquanto algumas empresas fecharam suas portas, outras registraram um crescimento de 200% em seus resultados. Contudo, nesse ambiente ainda tímido, anúncios significativos na área de tecnologia destacaram-se, como o lançamento do ChatGPT pela OpenAI em

Nenhuma entrada de sumário foi encontrada.

novembro de 2022 (OpenAI, 2022). Eu penso que para algumas empresas e inovações, o ritmo estabelecido em 2023 foi ideal para lançar algo transformador, oferecendo à humanidade uma notícia promissora e cheia de esperança. As pessoas precisavam de novidades revolucionárias e otimistas para fortalecer a narrativa de superação pós-pandemia.

Desde este lançamento em particular, a adoção dessa nova tecnologia foi a mais rápida já registrada, superando o crescimento de aplicativos como TikTok e Instagram. Apesar da transparência quanto às limitações do modelo recém-lançado, a UBS afirmou em relatório, citando dados da empresa de análise SimilarWeb, que cerca de 13 milhões de usuários únicos utilizaram o ChatGPT diariamente em janeiro, mais que o dobro dos níveis de dezembro de 2022 (Forbes Brasil, 2023). O ano de 2023 foi marcado por um grande entusiasmo e elevadas expectativas em relação às tecnologias emergentes, com os modelos de LLMs (Large Language Models) ocupando uma posição de destaque como uma das principais tendências tecnológicas do período. Este reconhecimento gerou impactos significativos e levou a mudanças estratégicas substanciais nas abordagens empresariais, particularmente no que tange à integração da inteligência artificial nos modelos operacionais e de negócios. Estamos rodeados por expectativas crescentes de que as empresas operem de forma mais enxuta, com menos pessoal, mas entregando maior eficiência, qualidade e desempenho. Isso por requer a utilização de especialidades específicas que garantam o funcionamento adequado dentro do domínio de cada negócio como a IA Generativa. Entre as tecnologias emergentes, a Inteligência Artificial Generativa destacou-se como uma das mais proeminentes. Espera-se que a Inteligência Artificial Generativa atinja o platô de produtividade nos próximos 2 a 5 anos e, junto com outras tecnologias, tem a responsabilidade de causar um impacto significativo nos negócios e na sociedade ao longo dos próximos 2 a 10 anos. É o que aponta a ferramenta de visualização da

Gartner Hype Cycle 2023 que ilustra a maturidade, adoção e aplicação social de tecnologias específicas. Esta ferramenta oferece um relatório para ser usado como guia para compreender as tendências tecnológicas emergentes com objetivo de auxiliar líderes de inovação e arquitetos empresariais na avaliação do impacto dessas tecnologias, bem como na elaboração de estratégias para capitalizar sobre elas (Gartner, 2023). Desde estes acontecimentos após a COVID-19, fez com que as organizações refletissem melhor sobre suas ações e prioridades. Certamente, a disponibilidade dessa tecnologia promoveu sua democratização, impactando significativamente uma ampla gama de indústrias e mercados globais, que reagiram prontamente (Forbes Brasil. ChatGPT coloca IA na agenda de CEOs brasileiros. Disponível em: [forbes.com.br](https://www.forbes.com.br/tecnologia/chatgpt-coloca-ia-na-agenda-de-ceos-brasileiros/). Acesso em: 29 out. 2024). A inteligência artificial generativa é, indubitavelmente, uma tecnologia recente com potencial revolucionário para a sociedade contemporânea. Desde o lançamento da OpenAI em 2022 pode se perceber, tanto usuários comuns quanto empresas começaram a transformar suas abordagens em diversas atividades, especialmente no que tange ao processamento de texto. Desconsiderar que vivemos em uma era marcada pela produção intensiva de grandes volumes de dados está fora da realidade das maiores empresas do mundo, estes dados de comportamento nunca antes registrados na história, jamais foram usados com a importância como hoje para as principais dinâmicas sociais da humanidade como, trabalhar, estudar, se informar, consumir e se relacionar.

Adoção Acelerada de IA Generativa

A adoção de tecnologias de inteligência artificial, especialmente a IA generativa, tem crescido rapidamente, servindo como um catalisador para transformação organizacional. Em uma pesquisa realizada em 2023, foi evidenciado que quase metade dos CEOs considera a IA generativa uma prioridade estratégica central para os próximos anos. Embora seu uso esteja se disseminando, a taxa de adoção geral de IA permaneceu estável, com foco em poucas funções empresariais como desenvolvimento de produtos e operações de serviço, áreas com maior adoção. Essa estabilização sugere que, apesar da notoriedade da IA generativa, seu escopo de aplicação ainda é limitado. A adoção da IAG continuou a crescer em ritmo acelerado em 2024, com 70% das empresas que utilizam essa tecnologia empregando ferramentas e bancos de dados vetoriais para aprimorar modelos base. O uso de bancos de dados vetoriais cresceu impressionantes 377% ano a ano, destacando o aumento da adoção de técnicas de Recuperação Aumentada por Geração (RAG). Além disso, 76% das empresas que utilizam Modelos de

Linguagem Grande (LLMs) preferem modelos de código aberto, frequentemente combinados com modelos proprietários. Este padrão sublinha uma tendência de flexibilidade e adaptação às necessidades específicas dos negócios, proporcionando uma base robusta para inovação contínua.

Impacto nos Negócios

O impacto da IA generativa nos negócios é profundo, oferecendo disruptões que podem remodelar indústrias inteiras. A tecnologia está sendo percebida como uma força motriz para a redução da força de trabalho tradicional e a requalificação para atender às novas demandas. Entretanto, apesar da expectativa de grandes transformações, apenas uma pequena fração das empresas relata uma contribuição significativa da IA para o lucro operacional. Isso indica um potencial não explorado, enquanto a atenção ao impacto da IA continua a crescer. A pesquisa de 2023 destacou a IA generativa como uma força transformadora nos negócios, especialmente com a metade dos CEOs identificando-a como prioridade estratégica. A IA generativa promete disruptão significativa, influenciando a força de trabalho com reduções e requalificações necessárias. No entanto, apesar do rápido avanço das ferramentas de IA generativa, a adoção geral de IA nas organizações permaneceu estável. As operações de serviços e o desenvolvimento de produtos são as áreas onde a IA mais impactou, mas apenas 23% dos entrevistados afirmaram que a IA contribuiu com 5% do EBIT, indicando potencial não realizado. O impacto da IA Generativa nos negócios é evidente, com empresas alcançando reduções de custos e aumentos de receita. Setores como recursos humanos e gestão da cadeia de suprimentos relatam os benefícios mais significativos. A IAG está redefinindo processos e operações, otimizando eficiência e eficácia em diversos segmentos. As organizações estão colhendo benefícios materiais ao adotar essas tecnologias inovadoras, que não apenas melhoraram os resultados financeiros, mas também a resiliência organizacional.

Desafios de Implementação

A transição para a IA generativa não ocorreu sem desafios. Embora democratizada, esta tecnologia requer uma infraestrutura robusta de dados para aproveitar ao máximo seu potencial. A pesquisa da Accenture sugere que 40% do tempo de trabalho poderia ser automatizado, destacando a necessidade urgente de adaptação. A IA generativa, com suas capacidades, obriga os líderes a reavaliar modelos de propriedade, infraestrutura de dados e governança. Além disso, mesmo com seu potencial, a IA continua sendo adotada em escopo limitado dentro das organizações, ressaltando desafios na implementação ampla e eficaz. Apesar do crescimento

promissor, as empresas enfrentam desafios significativos. A qualidade e a governança dos dados são preocupações primordiais, especialmente em relação à proteção de dados sensíveis, exigindo a implementação de medidas de segurança robustas. A customização de modelos, particularmente LLMs, com dados privados através de RAG, destaca a necessidade de bancos de dados vetoriais. Ferramentas como LangChain tornaram-se amplamente utilizadas, evidenciando o foco crescente na personalização de modelos para atender às demandas específicas.

Governança e Qualidade de Dados

Com a expansão da IA generativa, a governança tornou-se central. Questões de propriedade intelectual, direitos autorais e segurança de dados ganharam destaque. A complexidade aumenta à medida que modelos generativos consomem dados indiscriminadamente, exigindo práticas de governança unificadas. Regulamentações, como a Lei de IA da União Europeia, refletem preocupações sobre privacidade e uso ético. A pesquisa mostrou que executivos percebem a necessidade de equilíbrio entre acessibilidade e segurança, com a governança de dados emergindo como um tema crítico. O relatório da Databricks enfatiza a importância de modelos construídos sobre dados de alta qualidade. Soluções de governança unificada, como o Unity Catalog, estão ganhando popularidade, particularmente em indústrias regulamentadas como o setor financeiro. Isso reflete a necessidade crescente de conformidade regulatória e segurança rigorosa. A manutenção da linhagem, qualidade, controle e privacidade de dados ao longo do fluxo de trabalho de IA é crucial para o sucesso contínuo e sustentável das iniciativas de IA.

Estratégias de Implementação e Habilidades Emergentes

Os empresários visionários estão adotando estratégias que equilibram inovação e mitigação de riscos. A IA generativa, comparada a revoluções tecnológicas como o computador pessoal e a Internet, requer que as organizações sejam proativas na capacitação de sua força de trabalho. A democratização da IA está permitindo que mais pessoas contribuam para a inovação, mas também exige uma infraestrutura de dados flexível e um entendimento claro dos riscos envolvidos. A busca por profissionais capacitados na implementação e governança de IA é essencial para aproveitar as oportunidades e mitigar os riscos. A implementação de IAG varia entre indústrias, com setores como energia, tecnologia e mídia personalizando ou desenvolvendo seus próprios modelos. O tempo médio de implementação é de 1 a 4 meses, indicando a complexidade e a natureza adaptativa dessas tecnologias. Além disso, há uma crescente demanda por habilidades que combinem conhecimento técnico com compreensão dos

riscos e práticas de mitigação associadas à IAG. Apenas um terço das organizações considera essas habilidades cruciais, sublinhando a necessidade de desenvolvimento contínuo de capacidades profissionais.

Organização do trabalho

A adoção acelerada da IA generativa representa um ponto de inflexão para as empresas modernas. Ela oferece potencial sem precedentes para transformação, mas também traz à tona desafios significativos em termos de governança, infraestrutura e habilidades necessárias. Empresas que conseguirem equilibrar esses aspectos, adotando uma abordagem cuidadosa e estratégica, estarão bem-posicionadas para colher os frutos dessa tecnologia revolucionária. Como evidenciado, a IA generativa não é um simples modismo, mas um elemento central na redefinição do panorama corporativo futuro. Em resumo, enquanto a adoção da IA Generativa está crescendo rapidamente, as empresas enfrentam desafios significativos relacionados à qualidade dos dados, governança e customização de modelos. A ênfase em dados de alta qualidade e governança robusta está se tornando cada vez mais crucial para o sucesso das iniciativas de IA nas organizações. As estratégias eficazes para navegar nesse ambiente incluem a personalização de soluções, a implementação de práticas robustas de governança e o desenvolvimento de habilidades necessárias para mitigar riscos e maximizar os benefícios dessa tecnologia emergente.

Desafios de Dados e Governança na Implementação de IA Generativa em trabalhos relacionados.

A implementação de Inteligência Artificial (IA) Generativa em aplicações variadas, como serviços financeiros, atendimento ao cliente e desenvolvimento de produtos, traz à tona uma série de desafios relacionados a dados e governança de dados. Baseando-se nos problemas identificados em três artigos, torna-se evidente a necessidade de uma abordagem robusta e bem estruturada para lidar com essas questões.

Em Serviços Financeiros.

No contexto dos serviços financeiros, a qualidade e quantidade dos dados disponíveis para treinamento dos modelos de IA são insuficientes. Modelos treinados com dados deficientes correm o risco de produzir resultados imprecisos, comprometendo decisões financeiras importantes. Além disso, há riscos significativos de viés e discriminação nos dados utilizados, o que pode levar a decisões que afetam injustamente certos grupos de pessoas. A proteção da privacidade e segurança dos dados financeiros sensíveis é outro desafio crítico, dado o potencial de uso indevido de informações confidenciais. Em termos de governança, há uma necessidade

urgente de estabelecer políticas claras para o uso ético e responsável da IA. O setor financeiro é altamente regulado, e garantir a conformidade com essas regulações ao implementar IA é um desafio contínuo. A transparência nos modelos de "caixa preta" de IA também é uma preocupação, já que a falta de compreensão sobre como as decisões são tomadas pode minar a confiança nos sistemas.

Em Atendimento ao Cliente.

No atendimento ao cliente, a consistência dos dados é um problema frequente, com erros nos dados históricos de interações sendo comuns. Além disso, a integração de dados de múltiplos canais de atendimento apresenta desafios consideráveis, já que cada canal pode ter suas próprias peculiaridades e formatos de dados. A segurança dos dados pessoais dos clientes é uma preocupação constante, com riscos de vazamento que podem comprometer a confiança dos clientes. Problemas de Governança: A governança neste contexto exige a definição de responsabilidades claras sobre os resultados gerados pela IA, especialmente quando decisões automáticas podem impactar a experiência do cliente. Monitorar e auditar as decisões tomadas pelos sistemas de IA são essenciais para garantir que os resultados sejam justos e precisos. A atualização contínua dos modelos com novos dados também é crucial para manter a eficácia das soluções de IA.

Em Desenvolvimento de Produtos.

No desenvolvimento de produtos, a escassez de dados relevantes para treinamento em novos domínios representa um grande obstáculo. Garantir a qualidade e confiabilidade dos dados gerados pela IA é fundamental, pois decisões baseadas em dados defeituosos podem resultar em produtos de baixa qualidade. O uso de dados externos levanta preocupações sobre a violação de propriedade intelectual, especialmente quando não há diretrizes claras sobre a utilização desses dados. Os processos para validação humana dos resultados da IA são necessários para garantir que as soluções propostas sejam viáveis e alinhadas aos objetivos do negócio. Definir métricas adequadas para avaliar o desempenho dos modelos é um desafio que demanda atenção, dada a variedade de aplicações e expectativas. Por último, a falta de clareza sobre a propriedade intelectual dos produtos gerados por IA pode gerar conflitos e inseguranças jurídicas.

Adoção de Aplicações de IA Generativa

Os problemas destacados nos três domínios de aplicação da IA Generativa sublinham a importância de uma abordagem cuidadosa e estruturada para governança de dados e IA nas organizações. Cada aplicação traz suas particularidades, e endereçar esses desafios requer um compromisso contínuo com práticas de governança eficazes e éticas. Assim, garantir a qualidade, segurança, e responsabilidade nos usos de IA não apenas protege as organizações, mas também assegura benefícios duradouros para a sociedade. A inteligência artificial generativa tem emergido como uma das implementações tecnológicas mais amplamente adotadas no cenário corporativo atual, com 67% das empresas entrevistadas relatando seu uso. Esta tecnologia está sendo aplicada em várias áreas fundamentais que estão transformando a maneira como as organizações operam e competem no mercado. Uma das aplicações mais significativas da IA generativa é na melhoria da experiência do cliente, com 60% das empresas utilizando-a para aprimorar as interações e elevar a satisfação. A tecnologia permite personalizar experiências e antecipar as necessidades dos clientes, criando uma relação mais próxima e eficiente. Isso não só melhora a satisfação do cliente, mas também aumenta a fidelidade à marca. A IA generativa está sendo utilizada por 57% das empresas para otimizar processos internos. Ao automatizar tarefas repetitivas, as organizações conseguem liberar recursos humanos para atividades mais estratégicas, aumentando a eficiência operacional e reduzindo custos. Este tipo de automação não apenas acelera as operações, mas também melhora a precisão e a consistência das tarefas realizadas. Com 51% das empresas utilizando IA generativa para análises, a capacidade de extrair insights em tempo real é ampliada significativamente. Isso permite que as empresas tomem decisões informadas mais rapidamente, melhorando a agilidade e a capacidade de resposta às mudanças do mercado. A aceleração das análises é crucial para manter a competitividade em um ambiente empresarial dinâmico. Porém as organizações estão adotando uma variedade de abordagens para implementar a IA generativa:

- Modelos de Código Aberto: 60% das empresas utilizam modelos de código aberto sem treinamento adicional, beneficiando-se da inovação contínua da comunidade, o que traz flexibilidade e inovação sem custos excessivos.
- Treinamento com Dados Próprios: 57% das empresas estão personalizando modelos de código aberto com dados internos para atender a necessidades específicas, garantindo que as soluções de IA sejam diretamente relevantes para seus contextos de negócio.

- Ambientes Locais ou Nuvens Privadas: Metade das organizações prefere implementar IA em infraestruturas internas ou nuvens privadas, buscando um maior controle sobre segurança e privacidade dos dados.

E apesar dos benefícios, as empresas enfrentam desafios significativos na adoção da IA generativa:

- Custos de Integração: Para 50% das empresas, integrar IA às operações existentes representa um desafio financeiro e logístico, exigindo planejamento cuidadoso e uma clara justificativa de retorno sobre investimento (ROI).
- Custos de Armazenamento de Dados: 49% das empresas relatam dificuldades financeiras associadas ao armazenamento, devido ao volume crescente de dados que precisam ser geridos e protegidos.
- Riscos de Violações de Dados: Com a segurança de dados sendo crucial, 46% das empresas expressam preocupações significativas com a proteção de informações sensíveis.
- Lacuna de Habilidades: A falta de profissionais qualificados ou treinados adequadamente em IA é um desafio para 38% das empresas, ressaltando a importância de investir em capacitação.

Quanto a confiança nos dados é um pré-requisito essencial para o sucesso da IA. A esmagadora maioria dos entrevistados (94%) expressa confiança nos dados empresariais, enquanto 71% acreditam que sua infraestrutura está preparada para suportar IA empresarial. No entanto, o desafio de dados em silos (73%) e a localização (40%) permanecem obstáculos significativos. Para superar esses desafios, as empresas estão adotando arquiteturas de dados modernas, que oferecem:

- Segurança Aprimorada: 71% relatam melhorias na segurança dos dados.
- Melhoria na Análise de Dados: 59% observam melhorias significativas na capacidade de análise.
- Gerenciamento Eficiente de Dados: 58% reportam avanços no gerenciamento, facilitando governança e integração de dados.

Habilidades Emergentes

Para preencher as lacunas de habilidades em IA, as organizações estão investindo em:

- Programas de Treinamento Interno: 57% estão desenvolvendo iniciativas de treinamento interno para fomentar o talento.

- Colaboração Multifuncional: Também 57% promovem projetos colaborativos, integrando diversas áreas do conhecimento para soluções inovadoras.
- Aquisição de Talentos: 50% estão revisando estratégias de contratação para atrair e reter profissionais qualificados.

A IA generativa está na vanguarda da transformação digital nas empresas. Porém, para maximizar seu impacto, é fundamental que as organizações enfrentem desafios relacionados à infraestrutura de dados, governança e capacitação. Investir em arquiteturas de dados modernas e no desenvolvimento de talentos será crucial para o sucesso contínuo das iniciativas de IA e para manter a competitividade em um mercado em constante evolução. E o cenário empresarial de 2024 foi marcado por um enfoque estratégico na implementação da inteligência artificial (IA) em diversas cadeias de negócios. Empresários e líderes têm priorizado a integração da IA generativa em seus sistemas e produtos, motivados pelo potencial de transformação que esta tecnologia oferece.

No entanto, este estudo identificou uma lacuna crítica que impacta o desempenho satisfatório dessa integração: a qualidade e gestão dos dados provenientes de múltiplas e diversas fontes.

Lacuna Identificada

A principal lacuna identificada reside na governança de dados, que desempenha um papel crucial no sucesso da implementação de IA generativa. Embora a tecnologia tenha o potencial de revolucionar processos de negócios, a ausência de uma gestão eficaz dos dados pode comprometer suas vantagens. Este estudo visa explorar essa questão, destacando a necessidade de sistemas que representem o estágio de implementação da IA generativa nas empresas, além de uma categorização do nível de usabilidade da IA, que atualmente é inexistente.

O objetivo principal deste estudo é investigar a relação entre a governança de dados e a implementação da IA generativa em empresas digitais. Especificamente, o estudo busca entender como práticas específicas de governança de dados impactam o nível de usabilidade e sucesso das iniciativas de IA generativa. Identificar quais práticas e técnicas específicas de governança de dados são mais eficazes para cada nível de usabilidade de IA Generativa, classificando as empresas em níveis de amador, básico, intermediário, avançado e proficiente. Avaliar como a qualidade da gestão de dados influencia o sucesso ou fracasso das iniciativas de IA Generativa em diferentes fases de implementação, desde a exploração até o ambiente de produção. Analisar a relação entre o nível de maturidade da governança de dados de uma empresa e sua capacidade de integrar IA Generativa de forma direta ou indireta. Estabelecer em

que medida a governança de dados pode ser considerada um fator crítico na gestão eficaz da IA Generativa em empresas digitais.

Metodologia de Estudo de Caso

Este trabalho propõe uma abordagem determinística para explorar e descrever cenários relacionados à implementação e uso da IA Generativa. Para isso, realizou-se um estudo de caso com três empresas distintas, analisando padrões significativos na distribuição de características relevantes dos dados. O estudo busca validar o cenário e contexto da implementação de IA, além de identificar casos de uso específicos em cada empresa avaliada.

A pesquisa sobre a adoção de IA no ambiente empresarial é justificada pela necessidade de compreender aspectos menos evidentes e mais subjetivos desse processo. Este estudo visa desmistificar as capacidades da IA generativa, enfatizando a importância crítica da qualidade e da origem dos dados. Além disso, pretende-se estabelecer uma relação clara entre a implementação de IA e seu impacto mensurável no desempenho dos negócios, validando a governança de dados como um elemento central na gestão de IA. Pois com o rápido avanço da IA e sua crescente adoção no ambiente empresarial, este estudo promete fornecer insights valiosos para organizações que buscam implementar IA de maneira eficaz e responsável. As contribuições desta pesquisa são relevantes não apenas para empresas, mas também para gestores e profissionais de tecnologia, oferecendo uma visão abrangente e realista sobre a adoção de IA no ambiente corporativo. A proposta de criar um sistema de categorização para o nível de usabilidade da IA nas empresas representa uma contribuição original, ajudando organizações a avaliar seu progresso e identificar áreas de melhoria.

Estratégias do Estudo de Caso

O estudo de caso adotou uma abordagem meticulosa para analisar as respostas dos entrevistados e compreender o impacto da inteligência artificial generativa nas empresas selecionadas. As estratégias implementadas são descritas a seguir:

Sistema de Classificação e Análise de Respostas

As respostas dos entrevistados foram analisadas e agrupadas conforme um sistema classificatório desenvolvido especificamente para este trabalho. Importante destacar que os

participantes não tinham conhecimento prévio sobre essa categorização, assegurando que suas respostas fossem naturais e não influenciadas por expectativas externas. Este sistema foi essencial para identificar o nível de usabilidade da IA generativa em cada empresa, destacando a importância da governança de dados como um critério central para avaliação.

Análise de Documentos Internos

Uma análise detalhada de documentos internos foi realizada para avaliar o impacto real da IA generativa nas métricas de negócios. Esta análise revelou que nenhuma das empresas obteve um impacto mensurável significativo, o que é considerado um fator imprescindível para alcançar a classificação de proficiente. Essa descoberta sublinha a lacuna existente entre a implementação tecnológica e os resultados práticos de negócios.

Análise de Conteúdo e Convergência de Respostas

Foi conduzida uma análise de conteúdo focada na convergência das respostas dos entrevistados. Utilizando prospecção de palavras e ideias-chave, as respostas foram cruzadas com observações extraídas de relatórios de mercado. Estes relatórios apontam consistentemente a gestão e qualidade dos dados como fatores determinantes para o desempenho e controle eficaz da IA generativa nas cadeias produtivas. Essa análise permitiu identificar padrões e insights valiosos sobre a implementação de IA.

Critério de Seleção das Empresas

O critério para a seleção das empresas focou naquelas que operam 100% no ambiente digital, partindo da premissa de que estão mais propensas a adotar IA para alavancar a produtividade. As empresas selecionadas variam em tamanho, apresentando cenários contrastantes que ilustram a diversidade de desafios e oportunidades enfrentados. Essa variação permite uma análise comparativa das vantagens e desvantagens relativas da implementação de IA nas cadeias produtivas. Essas estratégias foram fundamentais para a realização de um estudo de caso sólido e abrangente, fornecendo uma compreensão detalhada dos fatores que influenciam a implementação e eficácia da inteligência artificial generativa em diferentes contextos empresariais.

Coleta de Dados e Resultados

Estudo de Caso da Empresa A

A entrevista discorre sobre a implementação de inteligência artificial (IA) generativa e as práticas de governança e segurança de dados na Empresa A.

Os principais pontos abordados incluem:

- Educação e Interesse em IA: Um dos interlocutores está cursando um MBA em Inteligência Artificial e Big Data, buscando migrar sua carreira para essa área.
- Desafios de Implementação de IA: Foi discutido como a Empresa A integra modelos de IA generativa, especialmente no caso de uso de modelos de linguagem de larga escala (LLMs) e sua evolução dentro da empresa.
- Integração com Provedores de IA: A Empresa A possui contratos com a OpenAI e a Microsoft para o uso de modelos como o GPT-4. Eles criam interfaces de proxy para facilitar o uso e a distribuição interna das soluções de IA. Governança e Segurança de Dados: A Empresa A assegura que dados pessoais e sensíveis não sejam utilizados indevidamente pelos modelos de IA utilizando técnicas de segurança e compliance, como as práticas SOC 2.
- Uso Interno e Casos de Uso: A empresa desenvolveu plataformas internas que utilizam IA para melhorar a experiência do cliente, responder a perguntas, e otimizar processos internos, como a atuação do agente.
- Estratégias de Produto e Desenvolvimento: A Empresa A busca facilitar a integração de IA para diversos times, mantendo a segurança e a eficiência, enquanto também avalia e testa diferentes modelos de IA para garantir o melhor desempenho em suas necessidades específicas.

Questionário da Empresa A

- ❖ Como a empresa integra a Inteligência Artificial Generativa em sua infraestrutura?
 - A Empresa A integra IA generativa através de contratos com provedores como OpenAI e Microsoft, utilizando um sistema de proxy que facilita o uso de APIs desses provedores.
- ❖ Qual é a importância da governança de dados no uso de IA na Empresa A?
 - A governança de dados é crucial para garantir que a IA não utilize indevidamente dados sensíveis. A Empresa A implementa práticas de compliance e segurança, como SOC 2, para proteger as informações dos usuários.
- ❖ Que medidas de segurança a Empresa A adota para proteger os dados utilizados por modelos de IA?
 - A Empresa A utiliza proxies para gerenciar a autenticação e garantir que as requisições estejam seguras. Além disso, tem equipes dedicadas à segurança que testam a robustez das integrações de IA.
- ❖ Como a Empresa A avalia diferentes modelos de IA antes de integrá-los em seus sistemas?
 - A avaliação inclui testes com diferentes casos de uso e análise de desempenho para garantir que o modelo escolhido atenda às necessidades específicas da Empresa A sem comprometer a segurança ou a eficiência.
- ❖ De que forma a IA está sendo utilizada para melhorar a experiência do cliente na Empresa A?
 - A IA é usada para responder a perguntas frequentes, auxiliar em interações de atendimento ao cliente, e otimizar processos internos, como o uso para diversas tarefas automatizadas.

- ❖ Quais são os desafios enfrentados na adoção de IA generativa e como a Empresa A os supera?
 - Os desafios incluem garantir a segurança dos dados, ajustar modelos às necessidades específicas, e integrar novas tecnologias de forma eficiente. A Empresa A supera esses desafios com uma estratégia robusta de governança de dados e parcerias cuidadosas com provedores de IA.
- ❖ Como a Empresa A garante que seus modelos de IA não sejam treinados com dados internos sensíveis?
 - A Empresa A assegura que os dados enviados para os provedores de IA sejam utilizados apenas nos contextos adequados, e utiliza verificações de compliance para evitar o uso indevido de informações sensíveis.

Esse questionário e resumo fornecem uma visão abrangente das práticas de implementação de IA generativa, governança e segurança de dados na Empresa A, conforme discutido na entrevista.

Estudo de Caso da Empresa B

A entrevista detalha a implementação de inteligência artificial (IA) generativa na Empresa B, uma empresa focada em CRM e melhoria de vendas consultivas. Os principais tópicos discutidos são:

- Utilização da IA para Melhorar Vendas: A Empresa B usa IA para apoiar vendedores em suas atividades diárias, como sumarização de conversas de vendas e cadastramento de informações, buscando tornar o CRM mais atraente e produtivo para seus usuários.
- Estratégia e Implementação de IA: A adoção de IA pela Empresa B é impulsionada por um dos fundadores que é entusiasta da tecnologia. A empresa começou a utilizar o ChatGPT para tarefas internas e está estudando a melhor forma de integrar IA em seus produtos para oferecer mais valor aos clientes.
- Infraestrutura de TI e Governança de Dados: A Empresa B é uma empresa enxuta, sem um departamento específico de TI, mas com áreas terceirizadas para suporte e RH. Utiliza serviços da AWS para armazenamento e segurança dos dados e está sempre buscando maneiras de otimizar suas operações internas.
- Integração e Personalização: A Empresa B oferece soluções integradas no CRM e outras plataformas como WhatsApp, utilizando APIs para incorporar funcionalidades de IA. No entanto, a integração ainda está em fase de desenvolvimento e testes.
- Desafios e Perspectivas da IA: A empresa reconhece os desafios de confiança e eficácia da IA, mas percebe um grande potencial no futuro. Está desenvolvendo soluções internas e para clientes que possam beneficiar os hábitos de trabalho dos vendedores e gerentes, tornando as operações mais ágeis e precisas.

Questionário da Empresa B

- ❖ Como a Empresa B utiliza a inteligência artificial em suas operações diárias?
 - A Empresa B utiliza IA para sumarizar conversas de vendas, cadastrar automaticamente informações no CRM, e facilitar o trabalho dos vendedores, tornando o CRM mais acessível e eficiente.

- ❖ Qual é a estratégia da Empresa B para integrar IA em seus produtos?
 - A estratégia é liderada por um fundador entusiasta da tecnologia, que começou explorando o ChatGPT. A integração está sendo testada internamente antes de ser oferecida aos clientes, focando em agregar valor real as suas operações.
- ❖ Como a Empresa B garante a segurança e a confidencialidade dos dados dos clientes?
 - A segurança dos dados é garantida através de armazenamento em servidores AWS, com instâncias separadas para cada cliente, evitando qualquer cruzamento de dados entre as contas.
- ❖ Quais são os principais desafios enfrentados pela Empresa B ao implementar IA?
 - Os principais desafios incluem garantir a exatidão e confiança dos dados sumarizados pela IA integrar adequadamente a IA nas operações existentes, e medir o real valor agregado pela IA em termos de produtividade e eficiência.
- ❖ De que maneira a IA está sendo utilizada para melhorar a performance de vendas na Empresa B?
 - A IA melhora a performance de vendas ao automatizar tarefas repetitivas, fornecer insights sobre o comportamento dos clientes, e ajudar os vendedores a priorizar leads mais quentes e oportunidades mais promissoras.
- ❖ Como a IA influencia a cultura organizacional da Empresa B?
 - A IA está sendo incorporada na cultura organizacional como uma ferramenta de apoio, não substituindo a interação humana, mas aumentando a produtividade e a eficiência das operações de vendas.
- ❖ Qual é a visão de futuro da Empresa B quanto ao uso de IA para aumentar a competitividade?
 - A Empresa B vê a IA como uma oportunidade de se diferenciar através de inovações que melhorem a experiência do usuário e a produtividade, estando sempre à frente do mercado sem correr os riscos associados as grandes disruptões.
- ❖ Quais são os impactos esperados da IA nos processos internos da Empresa B?
 - Espera-se que a IA aumente a agilidade dos processos, reduza erros humanos, e permita que a equipe se concentre em tarefas de maior valor agregado, como estratégias de vendas e relacionamento com clientes.

Este questionário e resumo oferecem uma visão abrangente sobre a implementação de IA generativa, governança e segurança de dados na Empresa B, conforme discutido na entrevista.

Estudo de Caso da Empresa C

A entrevista detalha o uso ainda amador de IA na Empresa C, um e-commerce de moda feminina focado em vendas pelo site. Os principais tópicos discutidos são:

- Utilização da IA para Melhorar Vendas: A Empresa C ainda não faz nenhum uso de IA diretamente para melhorar as vendas.
- Estratégia e Implementação de IA: A adoção de IA pela Empresa C ainda é muito embrionário, os funcionários receberam um treinamento introdutório de como usar o prompt do Chat GPT para ajudar em funções diárias e gerar insights, principalmente nas postagens em redes sociais, se faz o uso da ferramenta para ajuda com os textos das postagens.

- Infraestrutura de TI e Governança de Dados: A Empresa C não conta com servidor próprio, os dados principais de sua operação ficam delegados a seus fornecedores de plataforma de E-commerce, ERP e CRM. Como não possui uma área de TI e desenvolvimento interno, depende da evolução dessas empresas para um avanço tecnológico. Hoje é feito um esforço para controlar esse fluxo de dados de vendas e marketing a parte utilizando serviços como o Google Sheets e o Dashboard Looker Studio, também do Google.
- Desafios e Perspectivas da IA: A Empresa C reconhece que tem um longo caminho a frente no que se refere a IA, por enquanto se preocupa mais com a organização de dados e entende a importância de ter um controle interno desses dados, que hoje estão delegados 100% a empresas terceiras. Por outro lado, entende que parte desses dados delegados em empresas que têm maior maturidade em TI, permite a ela acesso a ferramentas impossíveis de se desenvolver internamente.

Questionário da Empresa C

- ❖ Como a Empresa C utiliza a inteligência artificial em suas operações diárias?
 - A Empresa C utiliza IA de forma amadora e pessoal, com um treinamento inicial que receberam, os funcionários e gestores da empresa utilizam Chat GPT e Google Gemini para gerar insights, sem ainda uma estruturação exata de tarefas que essas ferramentas poderiam auxiliar de forma massiva e escalável. Se destaca o uso em geração de textos para postagens em redes sociais e geração de textos para e-mail marketing.
- ❖ Qual é a estratégia da Empresa C para integrar IA em seus produtos futuramente?
 - A empresa entende o potencial que a utilização de IA pode oferecer para o setor que atua (e-commerce). Enxerga ótimas oportunidades para explorar seu uso na automação de atendimento e automação de fluxo de ponto de contato com seus leads. Apesar do entusiasmo, prefere acompanhar o caminho das grandes empresas e tomar decisões mais acertadas de onde investir seus esforços no campo da IA.
- ❖ Como a Empresa A garante a segurança e a confidencialidade dos dados dos clientes?
 - A empresa, como mencionado, armazena 100% dos dados de clientes em três ferramentas principais: plataforma de E-commerce, CRM e ERP, e com todos esses parceiros pactuou um contrato de confidencialidade de dados.
- ❖ Quais são os principais desafios enfrentados pela Empresa C ao implementar IA?
 - O desafio central hoje é ter a governança e gestão de dados, sem isso a empresa entende que não conseguirá utilizar a IA em seu máximo potencial. Hoje a empresa conta com o uso de IA de terceiros, por exemplo: Nas campanhas pagas que faz na plataforma da Meta, conta com a

IA da mesma para encontrar o melhor público-alvo para seus anúncios e entende que isso pode escalar exponencialmente seu alcance.

❖ Qual é a visão de futuro da Empresa C quanto ao uso de IA para aumentar a competitividade?

□ A Empresa C se enxerga cercada por incontáveis concorrentes lutando para se destacar em um mercado extremamente concorrido. Sobretudo no posicionamento de marketing e metrificação de campanhas entende que a implementação da IA pode revolucionar o mercado. Hoje estuda como implementar um sistema de Business Intelligence para se destacar de seus concorrentes na tomada de decisão.

Este questionário e resumo oferecem uma visão abrangente sobre a implementação de IA generativa, governança e segurança de dados na empresa C, conforme discutido na entrevista.

Análise

A ascensão dos LLMs é acompanhada pela integração de métodos Retrieval-Augmented Generation (RAG), que aprimoram as capacidades das técnicas tradicionais de solicitação e ajuste fino. O RAG emprega uma combinação de recuperação e geração de dados para fornecer respostas mais precisas e contextualmente relevantes. Essa abordagem dupla não apenas melhora a relevância dos resultados gerados, mas também atenua alguns riscos associados a alucinações e respostas fora do tópico, que são considerações críticas na implantação de LLMs para aplicativos empresariais [5][9]. À medida que as organizações adotam cada vez mais IA Generativa para várias aplicações— incluindo pesquisa empresarial, geração de conteúdo e análise automatizada de dados — a importância da governança de dados se torna primordial. Estruturas de governança eficazes são necessárias para garantir a conformidade com os padrões éticos e para abordar riscos potenciais, como vazamento rápido, toxicidade em resultados e desvio de desempenho ao longo do tempo [6][7]. Isso exige o estabelecimento de sistemas de monitoramento que avaliem continuamente os resultados do modelo em relação a benchmarks predefinidos, ajudando as organizações a manter a confiabilidade e a responsabilidade de seus sistemas de IA.[9][8]. Além disso, as organizações devem navegar pelos desafios associados à integração em sistemas existentes. Isso inclui a necessidade de pessoal qualificado para gerenciar essas tecnologias sofisticadas, bem como a criação de plataformas flexíveis que podem se adaptar ao cenário de IA em rápida evolução [19]. Com investimentos contínuos nessa tecnologia, a convergência da governança de dados e estratégias de incentivo inovadoras é fundamental para maximizar o potencial das IA Generativas, garantindo ao mesmo tempo o uso responsável e ético em ambientes empresariais [4][8].

Governança de dados em LLMs empresariais

A governança de dados é um aspecto crucial da implementação de modelos de linguagem de grande porte (LLMs) em ambientes empresariais. À medida que as organizações adotam cada vez mais AI Generativas fundamentais, uma estrutura abrangente de governança de dados é essencial para garantir a qualidade, a conformidade e a segurança dos dados, mantendo assim uma vantagem competitiva no mercado [20]. Essa estrutura deve facilitar o gerenciamento eficaz de dados durante todo o seu ciclo de vida, abordando as oportunidades e os desafios apresentados.

Importância da governança de dados

A governança de dados eficaz é vital para as organizações, pois alinha as práticas de gerenciamento de dados com os requisitos regulatórios e os objetivos organizacionais. Ela abrange os processos de coleta, armazenamento, categorização e utilização de dados para dar suporte à tomada de decisões e à eficiência operacional [17][21]. Um programa de governança de dados bem definido não apenas aprimora a qualidade dos dados, mas também dá suporte à conformidade com as leis de proteção de dados, protegendo assim as informações pessoais [10]. Isso é particularmente significativo na era da IA, onde as organizações devem navegar por cenários regulatórios complexos enquanto maximizam o valor derivado de seus ativos de dados.

Elementos-chave da governança de dados para LLMs

Uma estrutura de governança de dados robusta para LLMs empresariais deve incluir vários componentes críticos.

Qualidade dos dados: garantir a precisão, integridade e usabilidade dos dados é fundamental para uma governança eficaz. Os LLMs podem ajudar a identificar e retificar inconsistências em conjuntos de dados, melhorando a qualidade geral dos dados [10][21].

Conformidade regulatória: automatizar o monitoramento do uso e acesso aos dados pode ajudar as organizações a aderir aos requisitos regulatórios, reduzindo o risco de violações[- 10]. As estruturas de governança de dados devem incluir mecanismos para rastrear a conformidade com leis e padrões relevantes.

Privacidade de dados: implementar técnicas de aprimoramento de privacidade (PDTs) é essencial para proteger informações confidenciais. Isso envolve tornar os dados anônimos antes de seu uso em LLMs para garantir que as organizações possam aproveitar os insights sem comprometer a privacidade [18-][10].

Gerenciamento de mudanças: a adoção bem-sucedida de estratégias de governança de dados necessita de práticas eficazes de gerenciamento de mudanças. Garantir que as partes interessadas sejam treinadas e estejam cientes de novas políticas e procedimentos é essencial para o sucesso das iniciativas de governança de dados [17].

Propriedade e responsabilidade: estabelecer papéis claros de propriedade e administração de dados é crucial para gerenciar dados de forma responsável. Um conselho de governança de

dados pode supervisionar iniciativas e promover uma cultura de responsabilidade de dados em toda a organização [22].

Desafios na governança de dados para LLMs

Os desafios exclusivos apresentados pelos LLMs exigem estratégias de governança personalizadas. As organizações devem abordar questões como viés em saídas de IA e as implicações éticas do uso de dados. O monitoramento contínuo e o retreinamento de LLMs podem ajudar a mitigar o viés e garantir a justiça no manuseio de dados [20][10]. Além disso, as organizações precisam gerenciar os riscos associados a silos de dados, que podem dificultar o compartilhamento e a integração eficazes de dados [22].

Prompting RAG

O prompting Retrieval-Augmented Generation (RAG) é uma abordagem sofisticada que alavanca bases de conhecimento externas juntamente com modelos de linguagem tradicionais para aumentar a precisão e a relevância das respostas geradas. Esta técnica combina dois componentes principais: um recuperador e um gerador. O recuperador atua como um mecanismo de busca inteligente que identifica os documentos ou passagens mais relevantes com base nas consultas do usuário, enquanto o gerador, normalmente um modelo de linguagem avançado, processa essas informações selecionadas para produzir saídas coerentes e contextualmente apropriadas [11][5].

Mecanismo de solicitação RAG

O processo RAG começa com a fase de ingestão, onde os dados são obtidos de diversas bases de conhecimento, incluindo documentos internos e repositórios externos, como a web. Esta fase garante que o sistema tenha acesso a informações atualizadas e contextualmente relevantes. Ao organizar e dividir documentos apropriadamente, o RAG pode recuperar dados relevantes de forma eficiente em resposta a solicitações do usuário [23][24]. Na fase de recuperação, o RAG usa a entrada do usuário para pesquisar sua base de conhecimento em busca de documentos pertinentes. Isso envolve examinar os principais resultados para garantir que as informações mais precisas e contextualmente apropriadas sejam priorizadas. Os documentos recuperados

são então combinados com a entrada original do usuário, fornecendo um contexto rico que aprimora a qualidade de saída do gerador [12][25].

Benefícios da solicitação RAG

Precisão e relevância aprimoradas é uma das principais vantagens da solicitação RAG é sua capacidade de produzir respostas que são atuais e relevantes. Isso é particularmente benéfico em ambientes onde as informações estão evoluindo rapidamente, como os setores de saúde ou jurídico, onde as decisões dependem dos dados mais recentes[12][5]. O RAG permite que os modelos de linguagem contornem as limitações do conhecimento estático integrando informações em tempo real, o que reduz instâncias de alucinação — onde os modelos geram saídas imprecisas ou sem sentido.

Eficiência de recursos

O prompt RAG também oferece uma abordagem equilibrada em termos de personalização e gerenciamento de recursos. Ao contrário dos métodos tradicionais que exigem amplo poder computacional para ajuste fino, o RAG permite operações flexíveis, tornando-o mais acessível a uma gama maior de usuários e desenvolvedores [12][23]. Essa eficiência se estende a atualizações periódicas de dados, permitindo que as organizações mantenham seus modelos treinados nas informações mais recentes sem incorrer em tempo e custos extras associados ao ajuste fino completo [26][24].

Tendências futuras no prompt RAG

A evolução dos sistemas RAG abriu vários caminhos para pesquisa e desenvolvimento. Há um interesse crescente em melhorar a qualidade da recuperação, gerenciar dados não estruturados e aprimorar a segurança dos dados [25]. Além disso, o desenvolvimento de metodologias híbridas e técnicas de autorrecuperarão são áreas que estão sendo aktivamente exploradas para otimizar o desempenho de aplicativos RAG em vários domínios. À medida que as organizações reconhecem cada vez mais o potencial do RAG, há uma demanda significativa por melhores ferramentas de avaliação e métricas para avaliar a eficácia desses sistemas em aplicativos do mundo real [25][23].

Fine-Tuning (Ajuste-Fino)

O Fine-Tuning é um processo crucial na adaptação de modelos de grandes linguagens (LLMs) a tarefas ou domínios específicos, refinando seu conhecimento pré-existente com dados específicos da tarefa. Esse processo geralmente envolve a atualização dos pesos de um modelo pré-treinado para aprimorar seu desempenho em um determinado conjunto de dados, tornando-o mais especializado para aplicativos específicos em vários setores, como saúde, finanças e comércio eletrônico [13][27]. O ajuste fino contrasta com a fase de treinamento inicial, onde os modelos são construídos do zero usando pesos inicializados aleatoriamente dedicados a uma tarefa específica. Em vez disso, o ajuste fino ajusta um modelo já treinado modificando seus parâmetros com base em novos dados direcionados, o que ajuda o modelo a aprender as nuances do domínio específico [13].

Existem vários métodos e estratégias empregados no processo de Fine-Tuning, que podem incluir:

Completo

O ajuste fino completo envolve a atualização de todos os pesos e parâmetros de um modelo pré-treinado. Essa adaptação abrangente permite que o modelo alinhe efetivamente suas representações internas com as especificidades do conjunto de dados de destino, tornando-o adequado para uma ampla gama de aplicações [14].

Com eficiência de parâmetros (PEFT)

O PEFT se concentra na otimização do processo de ajuste fino atualizando apenas um pequeno subconjunto de parâmetros do modelo. Técnicas como LoRA (Low-Rank Adaptation) e adaptadores permitem reduções significativas em custos computacionais, mantendo o desempenho. Essa abordagem permite um uso mais eficiente de recursos, tornando-a prática para implantação em ambientes empresariais [14][9].

Sequencial

Esse método envolve treinar o modelo em uma tarefa primária e, posteriormente, refiná-lo por meio de ajustes incrementais. Por exemplo, um modelo inicialmente treinado em dados de texto

diversos pode ser ainda mais especializado em tarefas como resposta a perguntas ou geração de texto específico de domínio, melhorando seu desempenho nessas áreas [13][9].

Supervisionado

O ajuste fino supervisionado emprega conjuntos de dados rotulados adaptados a tarefas específicas, onde cada entrada é pareada com uma saída correta. Isso permite que o modelo aprenda previsões precisas aproveitando seu conhecimento pré-treinado, melhorando assim seu desempenho em aplicativos direcionados [28].

Ajuste de instruções

O ajuste de instruções melhora a capacidade de um modelo de seguir instruções de linguagem natural em várias tarefas, aumentando sua capacidade de resposta e generalização. Ao fornecer instruções e exemplos explícitos durante o treinamento, essa abordagem permite que o modelo entenda e execute melhor tarefas que não encontrou antes [29].

Aprendizado por Reforço a partir do Feedback Humano (RLHF)

Este método inovador treina modelos usando feedback de avaliadores humanos, permitindo o refinamento contínuo das capacidades do modelo com base em interações do mundo real. O RLHF ajuda o modelo a produzir respostas mais precisas e contextualmente relevantes, aumentando assim a eficácia geral na execução de tarefas [28].

Avaliação e Implantação

Após a conclusão do processo de ajuste fino, é essencial avaliar o desempenho do modelo usando várias métricas, como exatidão, precisão, recall e pontuação F1. Esta avaliação garante que o modelo generalize bem para novos dados e atenda aos requisitos de desempenho específicos das tarefas para as quais foi ajustado [27]. Uma vez validado, o modelo ajustado pode ser implantado em ambientes de produção, onde pode ser integrado em sistemas maiores e monitorado quanto à sua eficácia em aplicações do mundo real [27]. O ajuste fino permite que as empresas aproveitem os pontos fortes de modelos pré-treinados enquanto os personalizam para atender a necessidades específicas, criando uma abordagem mais ágil e eficiente para implantar recursos avançados de processamento de linguagem.

Ferramentas e tecnologias

A implantação eficaz de Large Language Models (LLMs) em ambientes empresariais depende do design cuidadoso de procedimentos que facilitam o acesso ao conhecimento e às ferramentas. Esse design capacita os modelos a aproveitar os recursos disponíveis de forma eficiente, garantindo que os resultados desejados sejam alcançados. Uma distinção fundamental nesse cenário é entre fluxos de trabalho baseados em agentes e aplicativos de geração aumentada de recuperação (RAG). Em fluxos de trabalho baseados em agentes, os LLMs recebem autonomia para decidir quais ferramentas utilizar e como aplicá-las, aumentando significativamente sua flexibilidade e capacidade [30].

Meta Prompting

A evolução da engenharia de prompts levou a uma mudança em direção à meta prompting, um conceito que envolve o uso de IA para gerar prompts de sistema, exemplos de poucos disparos e rotinas de procedimentos. Essa abordagem, descrita como “promptception”, permite que os modelos criem seus próprios prompts sob a orientação de um meta prompt, agilizando processos e maximizando as capacidades dos modelos [30].

Estruturas de ajuste fino

As principais empresas de tecnologia, como Hugging Face, Microsoft Azure e OpenAI, foram pioneiras em estruturas e serviços que simplificam e democratizam o processo de ajuste fino para LLMs. Essas plataformas fornecem ferramentas que atendem a várias necessidades do setor, desde fluxos de trabalho de ajuste fino automatizados até ambientes de treinamento escaláveis baseados em nuvem. A OpenAI introduziu o "ajuste fino como serviço", permitindo que as organizações adaptem modelos poderosos como o GPT-4 por meio de APIs fáceis de usar, minimizando a necessidade de experiência ou infraestrutura interna em IA[9].

LLMops e Gerenciamento de dados

LLMops representa um aspecto crucial da implantação e gerenciamento de LLMs em ambientes de produção, enfatizando o gerenciamento eficaz de dados. Um relatório da AWS

destaca que 93% dos Chief Data Officers reconhecem a estratégia de dados como essencial para maximizar o valor derivado da IA generativa [31].

Prompt Engineering: Desenvolvendo prompts eficazes que melhoram o desempenho do modelo.

Gerenciamento de dados: Garantindo o manuseio adequado de dados, controle de versão e análise exploratória.

Ajuste fino do modelo: Personalização de LLMs para tarefas específicas para melhorar sua eficácia[32].

Ferramentas de avaliação e monitoramento

Para avaliar o desempenho de LLMs, várias estruturas de avaliação e ferramentas de teste são empregadas. Essas ferramentas podem medir as capacidades do modelo e identificar vieses em tarefas relacionadas à linguagem, empregando conjuntos de dados de benchmarking e métricas de avaliação para fornecer insights abrangentes sobre o desempenho do modelo.

Monitoramento funcional: Métricas de rastreamento, como tempo de resposta e taxas de erro.

Monitoramento rápido: Avaliação de entradas do usuário para conteúdo prejudicial e injeções rápidas.

Monitoramento de resposta: Análise de saída para detectar problemas como alucinações e desvio de tópico [33].

Tendências emergentes

O cenário de grandes modelos de linguagem (LLMs) é caracterizado por rápido crescimento e inovação, classificando-se entre as 14% principais tendências em tecnologia de acordo com o TrendFeedr[2]. Com uma impressionante taxa de crescimento anual superior a 90%, os LLMs não estão apenas remodelando o mercado tecnológico, mas também gerando interesse significativo entre as organizações, como evidenciado pelas 562 entidades que se envolvemativamente com esses sistemas avançados de IA [2]. Essa trajetória ascendente levou a uma

taxa de crescimento mensal de 10,64% em novas organizações que adotaram a tecnologia LLM nos últimos cinco anos [2].

Democratização e acessibilidade da IA

Um dos aspectos mais transformadores dos LLMs é sua capacidade de democratizar a IA. Ao tornar os aplicativos de IA acessíveis por meio da linguagem natural, os LLMs permitem que organizações de todos os tamanhos aproveitem soluções orientadas por IA adaptadas às suas necessidades específicas [15]. Essa acessibilidade é crucial para desbloquear novos insights de grandes quantidades de dados organizacionais, permitindo que as empresas analisem tendências e padrões que podem informar significativamente a tomada de decisões estratégicas [15]. A crescente dependência de soluções de IA ressalta a importância de integrar essas tecnologias para permanecer competitivo em um mercado em evolução [15].

Inovações em estruturas de avaliação

À medida que os LLMs ganham força, a necessidade de estruturas de avaliação eficazes se torna cada vez mais aparente. Várias metodologias foram desenvolvidas para avaliar a qualidade e a eficácia dos LLMs em diferentes aplicativos [16]. Estruturas notáveis incluem Prompt Flow, Weights & Biases do Microsoft Azure em combinação com LangChain e DeepEval da confidence-ai[16]. Essas ferramentas são essenciais para garantir a confiabilidade dos resultados do LLM e para aprimorar sua implantação em vários setores.

Considerações éticas e governança de dados

Com a rápida adoção de LLMs, a governança ética de dados surgiu como uma preocupação urgente. O potencial de exposição de dados sensíveis por meio de interfaces de programação de aplicativos (APIs) exige conformidade cuidadosa com regulamentações e políticas da empresa ao ajustar LLMs em dados proprietários [34]. Além disso, as organizações devem abordar questões relacionadas a vieses inadvertidos e às implicações éticas do conteúdo gerado por IA, particularmente em relação a regulamentações de privacidade de dados, como o GDPR, que exige consentimento explícito para uso de dados [18][17].

Perspectivas futuras

Olhando para o futuro, os avanços em LLMs estão definidos para transformar significativamente as indústrias. Os marcos previstos incluem o código aberto de modelos como

o LLaMA-3 da Meta e a competição entre gigantes da tecnologia e startups para expandir os limites das capacidades do LLM [35]. À medida que as empresas continuam a explorar o potencial dos LLMs, equilibrar a inovação com a governança de dados responsável será crucial para moldar um futuro sustentável para as tecnologias de IA.

Análise Comparativa das Empresas A, B e C

A análise das três empresas revela diferentes níveis de maturidade na implementação e uso de Inteligência Artificial (IA) generativa, bem como distintas abordagens para governança e segurança de dados. Vamos examinar cada empresa individualmente e, em seguida, compará-las.

Empresa A: Implementação Avançada

A Empresa A demonstra uma abordagem sofisticada e estruturada para a implementação de IA generativa:

- **Integração Robusta:** Utiliza contratos com provedores como OpenAI e Microsoft, implementando um sistema de proxy para facilitar o uso de APIs.
- **Governança de Dados:** Prioriza a segurança e compliance, adotando práticas como SOC 2 para proteger informações sensíveis.
- **Casos de Uso:** Desenvolveu plataformas internas que utilizam IA para melhorar a experiência do cliente e otimizar processos internos.
- **Avaliação de Modelos:** Realiza testes rigorosos com diferentes casos de uso antes da integração.

Empresa B: Implementação em Desenvolvimento

A Empresa B está em um estágio intermediário de adoção de IA:

- **Foco em CRM:** Utiliza IA para sumarizar conversas de vendas e automatizar o cadastramento de informações no CRM.
- **Estratégia em Evolução:** Liderada por um fundador entusiasta, a empresa está testando internamente antes de oferecer soluções aos clientes.
- **Infraestrutura:** Utiliza serviços AWS para armazenamento e segurança de dados, sem um departamento de TI dedicado.
- **Desafios:** Enfrenta questões de confiança e eficácia da IA, buscando integrar a tecnologia de forma a agregar valor real.

Empresa C: Implementação Embrionária

A Empresa C está nos estágios iniciais de adoção de IA:

Critério	Empresa A	Empresa B	Empresa C
Integração de IA	Avançada	Intermediária	Básica
Governança de dados	Robusta	Moderada	Limitada
Casos de uso	Múltiplos	Focados	Mínimos
Infraestrutura de TI	Sólida	Adequada	Terceirizada
Estratégia e visão de futuro	Clara	Em desenvolvimento	Embrionária

- Uso Limitado: Utilização amadora de ferramentas como ChatGPT para tarefas diárias e geração de conteúdo para redes sociais.
- Infraestrutura: Não possui servidor próprio, delegando dados a fornecedores de plataforma de e-commerce, ERP e CRM.
- Governança de Dados: Reconhece a importância, mas ainda depende totalmente de terceiros para gestão de dados.
- Perspectivas Futuras: Entende o potencial da IA, mas prefere uma abordagem cautelosa, seguindo tendências de grandes empresas.

Comparação e Classificação

Para classificar as empresas, consideraremos os seguintes critérios:

1. Nível de integração de IA
2. Governança e segurança de dados
3. Casos de uso implementados
4. Infraestrutura de TI
5. Estratégia e visão de futuro

Classificação Final:

1. **Empresa A: Implementação Avançada**
2. **Empresa B: Implementação em Desenvolvimento**
3. **Empresa C: Implementação Embrionária**

A Empresa A se destaca por sua abordagem abrangente e madura na implementação de IA, com forte ênfase em segurança e governança de dados. A Empresa B mostra um progresso significativo, focando em aplicações específicas de IA para melhorar suas operações de CRM e vendas. A Empresa C, embora reconheça o potencial da IA, ainda está nos estágios iniciais de adoção, com uso limitado e dependência de terceiros para infraestrutura e gestão de dados. Esta análise revela uma clara diferenciação entre as empresas em termos de maturidade na adoção de IA, com implicações significativas para sua competitividade e potencial de inovação futura no mercado.

Conclusão

A análise das empresas A, B e C, juntamente com as tendências observadas no mercado, revela insights valiosos sobre a implementação da IA generativa e sua relação com a governança de dados em empresas digitais. Esta conclusão sintetiza os principais achados e implicações do estudo.

Níveis de Maturidade na Implementação de IA

O estudo identificou uma clara diferenciação nos níveis de maturidade na implementação de IA generativa entre as empresas analisadas:

1. Empresa A: Avançada
2. Empresa B: Intermediária
3. Empresa C: Amadora/Embrionária

Esta classificação reflete não apenas o uso da tecnologia, mas também a sofisticação das práticas de governança de dados e a integração da IA nos processos de negócios.

Governança de Dados como Fator Crítico

A pesquisa estabeleceu uma forte correlação entre a maturidade da governança de dados e o sucesso na implementação de IA generativa:

- Empresa A: Demonstrou práticas robustas de governança, incluindo compliance SOC 2 e sistemas de proxy para segurança.
- Empresa B: Apresentou práticas moderadas, com foco em segurança de dados na AWS.
- Empresa C: Revelou práticas limitadas, dependendo fortemente de terceiros para gestão de dados.

Esta correlação sugere que a governança de dados é, de fato, um fator crítico na gestão eficaz da IA generativa em empresas digitais.

Desafios e Oportunidades

O estudo identificou desafios comuns enfrentados pelas empresas:

- Equilibrar o custo-benefício da implementação de IA
- Estabelecer métricas claras para avaliar o impacto da IA nos negócios
- Garantir a segurança e compliance dos dados utilizados em modelos de IA

Ao mesmo tempo, oportunidades significativas foram observadas:

- Automação de processos e melhoria na experiência do cliente
- Otimização de operações internas e tomada de decisão baseada em dados
- Potencial para inovação e diferenciação competitiva

Tendências Emergentes

O estudo também apontou para tendências importantes no mercado de IA generativa:

- Crescimento da demanda por consultores especializados em IA
- Aumento gradual da participação da IA em atividades geradoras de valor
- Evolução para modelos híbridos, combinando IA com intervenção humana

Implicações para a Prática

Com base nos resultados, recomenda-se que as empresas:

1. Priorizem o desenvolvimento de práticas robustas de governança de dados
2. Invistam em infraestrutura e habilidades necessárias para implementação eficaz de IA
3. Estabeleçam métricas claras para avaliar o impacto da IA nos negócios
4. Adotem uma abordagem estratégica e gradual na implementação de IA generativa

Considerações Finais

Este estudo contribui para preencher uma lacuna importante na avaliação da implementação de IA generativa em empresas digitais. Ao estabelecer uma relação clara entre governança de dados e sucesso na implementação de IA, o estudo fornece um framework valioso para empresas que buscam avançar nessa área.

A jornada de implementação da IA generativa ainda está em seus estágios iniciais, e muitas descobertas serão feitas à medida que novos desafios surgiem. No entanto, fica claro que a qualidade e a gestão eficaz dos dados são fundamentais para o sucesso dessas iniciativas. As empresas que priorizarem a governança de dados estarão mais bem posicionadas para aproveitar o potencial transformador da IA generativa e obter vantagens competitivas significativas no mercado digital em rápida evolução.

Referências

(Inovação e qualidade de vida no trabalho: uma visão integrada da gestão a partir de estudo de caso na indústria petroquímica do Grande ABC, 2010).

(FUNDAÇÃO GETULIO VARGAS, [2022])

(Gartner, 2023).

GARTNER. Gartner survey finds CEOs cite AI as the top disruptive technology impacting industries. Disponível em:

[**www.gartner.com](<https://www.gartner.com/en/newsroom/press-releases/2023-05-17-gartner-survey-finds-ceos-cite-ai-as-the-top-disruptive-technology-impacting-industries>)**
Acesso em: 2 dez. 2024.

QUANTUMBLACK AI BY MCKINSEY. *The state of AI in 2023: Generative AI's breakout year: As organizations rapidly deploy generative AI tools, survey respondents expect significant effects on their industries and workforces*. [S.l.]: QuantumBlack AI by McKinsey, ago. 2023.

GREEN, Adam. *The great acceleration: CIO perspectives on generative AI*. [S.l.]: MIT Technology Review Insights, 2023. Relatório patrocinado por Databricks. Editor: Teresa Elsey. Publicador: Nico Crepaldi.

SINGLA, Alex; SUKHAREVSKY, Alexander; YEE, Lareina; CHUI, Michael, com HALL, Bryce. *The state of AI in early 2024: Gen AI adoption spikes and starts to generate value*. QuantumBlack, AI by McKinsey and McKinsey Digital, 2024. DATABRICKS. *State of Data+AI: Data intelligence*. [S.l.]: Databricks, 2024.

CLOUDERA. *The State of Enterprise AI and Modern Data Architecture*. [S.l.]: Cloudera, 2024. Relatório de pesquisa conduzido por Researchscape.

BILALLI, Besim; JOVANOVIC, Petar; NADAL, Sergi; QUERALT, Anna; ROMERO, Oscar. *There is no Data Science without Data Governance: a Proposal Based on Knowledge Graphs*. Universitat Politècnica de Catalunya, UPC-BarcelonaTech, [s.d.]

(KODEX AI, 2024, p x) no texto na refs: KODEX AI. Adopting generative AI in banking. Kodex AI, 2024

Borges (2023) discute como a inteligência artificial está sendo utilizada para promover melhorias operacionais nas organizações. BORGES, Aline de Fátima Soares. Uso estratégico da inteligência artificial nas organizações. São Paulo, 2023

TADDEO, Mariarosaria; ZIOSI, Marta; TSAMADOS, Andreas; GILLI, Luca; KURAPATI, Shalini. Artificial Intelligence for National Security: The Predictability Problem. Oxford: Centre for Emerging Technology and Security Research Report, Oxford Internet Institute, set. 2022.

[1]: [Somaieh Nikpoor] DATA GOVERNANCE IN THE AGE OF LARGE-SCALE DATA-DRIVEN LANGUAGE MODELS., (<https://independent.academia.edu/SomaiehNikpoor>) [2022].

[2]: BEHIND THE AI BOOM: Large Language Model (LLM) Trends. TrendFeedr, [2024].

[3]: Small Language Models vs. Large Language Models - Synergy TechnicalSMALL LANGUAGE MODELS VS. LARGE LANGUAGE MODELS. Synergy Technical, [2024].

[4]: SROCH ROBIT AND RAMACHANDRAN JAYACHANDRAN. EVOLVING TRENDS IN PROMPT ENGINEERING FOR LLMS WITH OPEN DATA SCIENCE., [2023]

[5]: RETRIEVAL AUGMENTED GENERATION (RAG) FOR LLMS. In: Prompt Engineering Guide. [Gao et al., 2023]

[6]: The Ultimate Guide to Fine-Tuning LLMs from Basics to Breakthroughs: An Parthasarathy, V. B., Zafar, A., Khan, A., & Shahid, A. (2024, August). The Ultimate Guide to Fine-Tuning LLMs from Basics to Breakthroughs: An Exhaustive Review of Technologies, Research, Best Practices, Applied Research Challenges and Opportunities (Version 1.0). CeADAR: Ireland's Centre for AI, University College Dublin.

[7]: AAYS Analytics. (n.d.). Enhancing techniques for effective governance of LLMs. AAYS Analytics Blog. <https://blog.aaysanalytics.com/post/enhancing-techniques-for-effective-governance-of-llms> [2024]

[8]: Bose, A. (2023, September 21). Bolstering enterprise LLMs with machine learning operations foundations: The benefits of MLOps must balance with well-managed operational practices and risk management.

[9]: Evidentlyai. (2024, November 27). 45 real-world LLM applications and use cases from top companies. Evidentlyai Blog. <https://www.evidentlyai.com/blog/llm-applications>

[10]: Mairena, A., & Liu, C. (2023, May 24). The Moveworks Enterprise LLM Benchmark: Evaluating large language models for business applications. Moveworks

[11]: Shakudo. (n.d.). Safeguarding AI Innovation - Implementing Robust Data Governance Framework for LLMs in Enterprise Environments. Shakudo Blog. <https://www.shakudo.io/blog/implementing-data-governance-framework-for-llms-in-enterprise-environments>

[12]: LinkedIn. (2024). Navigating Data Governance Challenges in the Era of Large Language Models (LLMs) and AI. LinkedIn Pulse. <https://www.linkedin.com/pulse/navigating-data-governance-challenges-era-large-fnq5e/>

[13]: Capital One. (2024, November 4). Building a data governance framework. Capital One Software Blog. <https://www.capitalone.com/software/blog/data-governance-framework/>

[14]: Sanjay T S. (2023, May 1). Mastering the Data Realm: Unraveling LLMs in Data Governance. LinkedIn Pulse. <https://www.linkedin.com/pulse/mastering-data-realm-unraveling-llms-governance-sanjay-t-s/>

[15]: Keyser, C. (2023, August 23). Why data governance is essential for enterprise AI. IBM Think. <https://www.ibm.com/think/topics/data-governance-for-ai>

[16]: Farmer, D. (2024, September 10). 5 data governance framework examples: Data governance isn't plug and play: Organizations must select which data governance framework best fits their business goals and needs. TechTarget. <https://www.techtarget.com/searchDataManagement/tip/5-data-governance-framework-examples>

[17]: Khangarot, S. (2024, April 24). The RAG Effect: How AI Is Becoming More Relevant And Accurate. Forbes. <https://www.forbes.com/councils/forbesbusinesscouncil/2024/04/24/the-rag-effect-how-ai-is-becoming-more-relevant-and-accurate/>

[18]: Ezra, O. (2024, September 25). Enterprise LLM: The Challenges and Benefits of Generative AI via RAG. K2view. <https://www.k2view.com/blog/enterprise-llm>

- [19]: McKinsey & Company. (2024, October 30). What is retrieval-augmented generation (RAG)? McKinsey & Company. <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-retrieval-augmented-generation-rag>
- [20]: MyScale. (2024, April 17). Prompt Engineering vs Fine-tuning vs RAG. Medium. <https://medium.com/@myscale/prompt-engineering-vs-finetuning-vs-rag-cfae761c6d06>
- [21]: Sarmah, A. (2024, November 29). Retrieval Augmented Generation (RAG) and Knowledge Management. KM Insider. <https://kminsider.com/topic/retrieval-augmented-generation-rag-and-knowledge-management/>
- [22]: Microsoft. (2024, August 28). Retrieval Augmented Generation using Azure Machine Learning prompt flow (preview). Microsoft Learn. <https://learn.microsoft.com/en-us/azure/machine-learning/concept-retrieval-augmented-generation?view=azureml-api-2>
- [23]: Dhaduk, H. (2023, July 3). A Complete Guide to Fine Tuning Large Language Models: Empower your models, elevate your results with this expert guide on fine-tuning large language models. Simform. <https://www.simform.com/blog/completeguide-finetuning-llm/>
- [24]: Dilmegani, C. (2024, August 31). LLM Fine-Tuning Guide for Enterprises. AIMultiple. <https://research.aimultiple.com/llm-fine-tuning/>
- [25]: Nexla. (2024). LLM Fine-Tuning—Overview with Code Example. Nexla. <https://nexla.com/enterprise-ai/llm-fine-tuning/>
- [26]: Analytics Insight. (2024, November 28). Improving Large Language Models: Advances in Fine-Tuning Strategies - Learn advancements in fine tuning strategies of LLMs for enhanced accuracy. Analytics Insight. <https://www.analyticsinsight.net/llm/improving-large-language-models-advances-in-fine-tuning-strategies>
- [27]: Weng, B. (2024). Navigating the Landscape of Large Language Models: A Comprehensive Review and Analysis of Paradigms and Fine-Tuning Strategies. arXiv. <https://arxiv.org/html/2404.09022v1>
- [28]: Varghese, L. (2024, July 24). How LLMs are Shaping Enterprise-Scale Applications: By implementing structured frameworks such as AI agents, integrating structured prompts, and incorporating human interaction patterns and guardrails, businesses can navigate LLMs'

creative potential while ensuring reliability. MIT Sloan Management Review Middle East. <https://www.mitsloanme.com/article/how-llms-are-shaping-enterprise-scale-applications/>

[29]: Atlan. (2024, April 3). Data Management for LLM Deployments: Importance, Challenges, Best Practices. Atlan. <https://atlan.com/know/ai-readiness/data-management-for-llm-deployments/>

[30]: Dilmegani, C. (2024, October 9). Comparing 10+ LLMOps Tools: A Comprehensive Vendor Benchmark. AIMultiple. <https://research.aimultiple.com/llmops-tools/>

[31]: Aisera. (2024, October). Aisera Named a Leader in October, 2024 Gartner Magic Quadrant for AI Applications in IT Service Management. Aisera. <https://aisera.com/blog/llm-evaluation/>

[32]: Hughes, D. (2023, April 28). Large Language Models (LLMs) for Enterprises: A Comprehensive Guide to Navigating the New AI Frontier. Graphable. <https://www.graphable.ai/blog/large-language-models-llms/>

[33]: Huang, J. (2024, March 5). Evaluating Large Language Model (LLM) systems: Metrics, challenges, and best practices. Data Science at Microsoft. <https://medium.com/data-science-at-microsoft/evaluating-llm-systems-metrics-challenges-and-best-practices-664ac25be7e5>

[34]: Red Hat. (2024, October 1). LLMs vs SLMs. Red Hat. <https://www.redhat.com/en/topics/ai/llm-vs-slm>

[35]: Huynh, D. (2024, January 27). State Of LLM In 2023: A Quick Recap On Latest Advancements. Medium. <https://medium.com/@vndee.huynh/state-of-llm-in-2023-a-quick-recap-on-latest-advancements-46a55dfe1fe5>