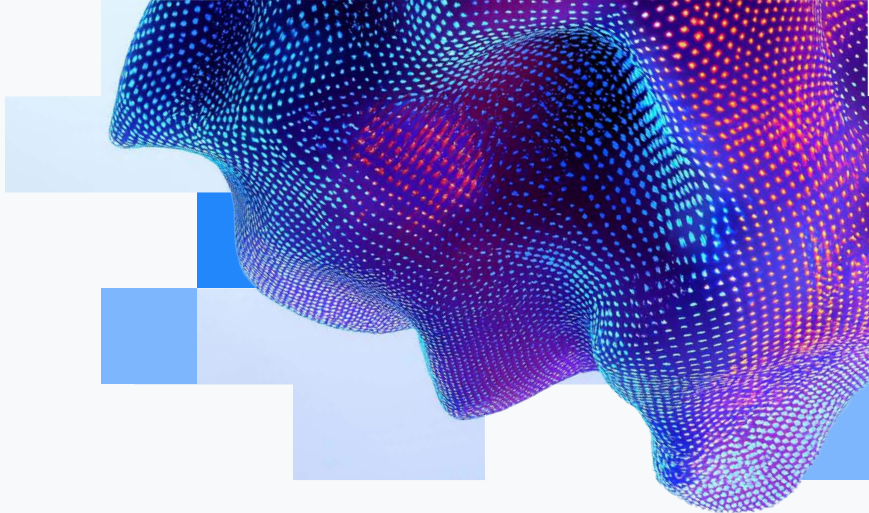


Google Cloud



Guia sobre agentes de dados

Índice



01	O que são os agentes de dados?	03
	Definições, principais competências e componentes essenciais que diferenciam os agentes de dados dos modelos de IA padrão.	
02	Por que os agentes estão em alta?	05
	Os avanços tecnológicos que estimulam o surgimento de agentes e a evolução de fluxos de trabalho multiagentes.	
03	Agentes de dados do Google Cloud	07
	Agentes pré-criados do Google Cloud para engenheiros de dados, cientistas de dados, usuários comerciais, desenvolvedores e gestores de dados.	
04	Como criar agentes de dados no Google Cloud	11
	Um guia prático para a criação de agentes de dados personalizados, desde os dados brutos até a implantação e otimização.	
05	Desenvolva uma liderança de IA	25
	As vantagens estratégicas do Google Cloud para a criação e o escalonamento de agentes.	



O que são os agentes de dados?

Um agente de dados é um sistema inteligente e orientado por objetivos que foi criado para resolver problemas complexos de dados comerciais e atuar como um parceiro autônomo do início ao fim do processo.

Os agentes de dados têm cinco competências principais:

01

Entender o contexto, como esquema, permissões ou jargões.

02

Reunir e sintetizar informações de diversas fontes.

03

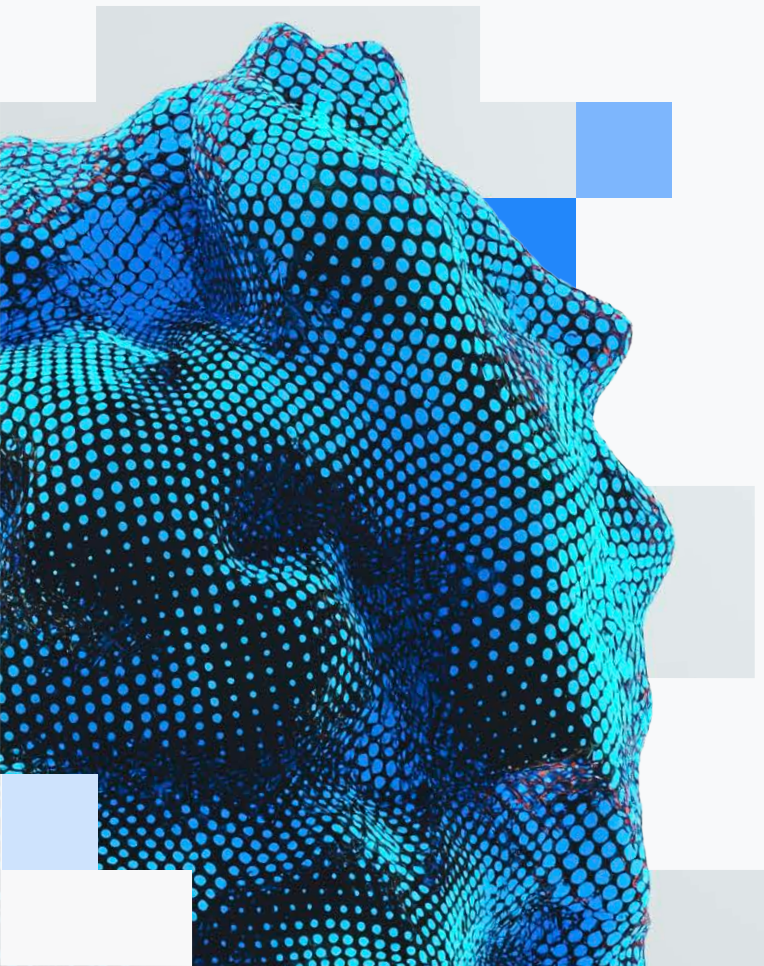
Atuar como solucionadores de problemas orientados por objetivos usando planos de várias etapas.

04

Planejar e executar usando abordagens alternativas.

05

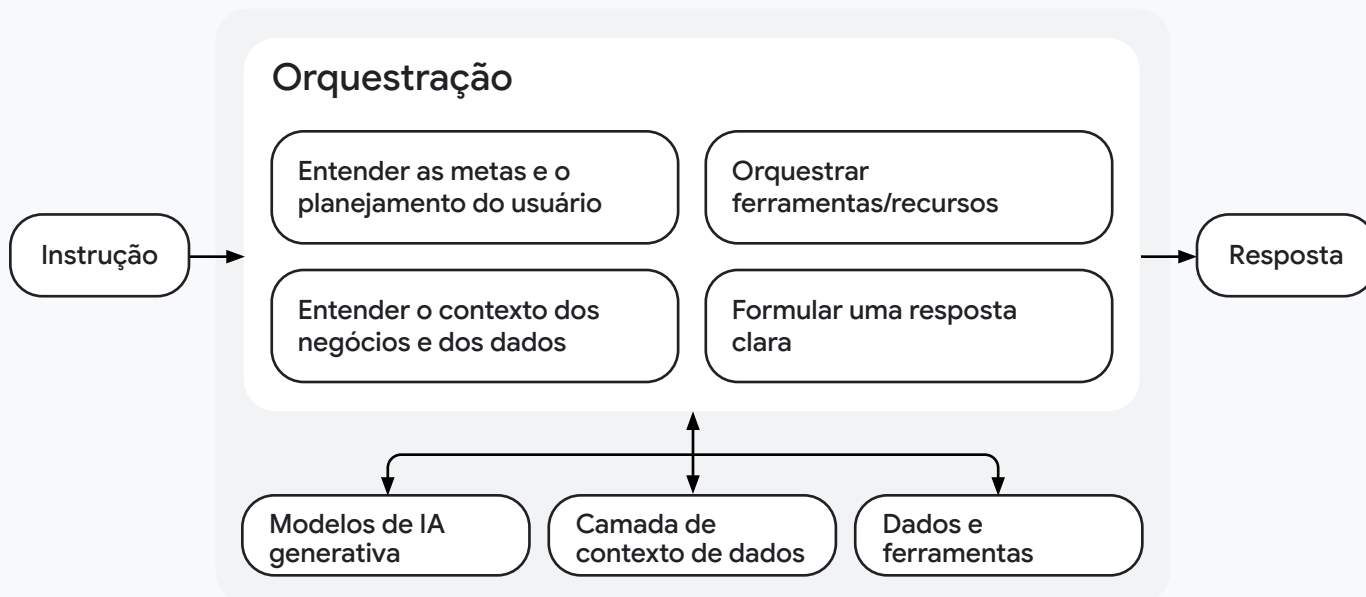
Usar várias ferramentas, incluindo modelos de machine learning (ML) e outros agentes.





Um agente de dados é um sistema coordenado com um mecanismo de raciocínio de IA, uma camada de contexto de dados e um mecanismo de orquestração que seleciona as melhores ferramentas para chegar a resultados mais abrangentes.

Por exemplo, ao responder a uma pergunta como *"Por que as vendas caíram no Nordeste?"*, um agente de dados geraria e executaria consultas sobre dados de vendas, marketing e gestão de relacionamento com o cliente (CRM), preparando um resumo parecido com este: *"A queda nas vendas está relacionada a uma redução de 50% nos gastos com publicidade digital na região Nordeste"*, apresentado recursos visuais para corroborar as afirmações..





Por que os agentes estão em alta?

As organizações têm exigido níveis cada vez mais altos de automação, eficiência e tomada de decisões baseada em dados. Essas demandas agora podem ser atendidas com o advento dos **agentes de IA**, que são ativados por avanços tecnológicos rápidos, como modelos de fundação mais poderosos, recursos de dados especializados e mecanismos de raciocínio sofisticados.

Os agentes de IA evoluíram, saindo de modelos de tarefa única para sistemas sofisticados capazes de raciocínio complexo e orquestração de fluxos de trabalho com vários agentes para automatização de processos inteiros.





Em um relatório de junho de 2025, a Gartner previu que a análise aumentada se tornará autônoma até 2027, gerenciando 20% dos processos de negócios com benefícios proativos, colaborativos, contextuais e contínuos.

Fonte: [Gartner](#)

Com o aumento da demanda das equipes de análise de dados, a criação de processos mais automatizados é fundamental. Os agentes de dados são uma solução poderosa porque ampliam os recursos das equipes.

Eles oferecem análises de autoatendimento por conversa e orientadas por eventos para **usuários comerciais**, atuam como assistentes incansáveis de análise detalhada de dados para **analistas de dados**, automatizam ciclos de vida de ML para **cientistas de dados** e traduzem linguagem natural em consultas SQL e pipelines de dados para **desenvolvedores** e **engenheiros de dados**.

Porque conseguem lidar com tarefas repetitivas, os agentes mudam o foco da equipe da execução para a supervisão estratégica e a inovação.





Agentes de dados do Google Cloud

O Google Cloud entende que cada organização tem uma jornada de IA única. Por isso, nossa abordagem aos agentes de dados se baseia em dois princípios fundamentais: poder e escolha. No lugar de uma solução única, o Google Cloud oferece um conjunto de agentes pré-criados e prontos para uso, projetados para funções específicas, para gerar valor imediato às suas equipes. Além disso, o Google Cloud dá acesso a várias ferramentas para quem precisa criar experiências de agentes totalmente personalizadas.

Neste capítulo, vamos conhecer os agentes de dados prontos para uso do Google, criados para ajudar e automatizar fluxos de trabalho para todos os usuários de dados.

Agentes de dados

Analista de dados
Consultas e análises



Usuário comercial
Recebimento de insights e tomada de decisões



Desenvolvedor
Criação de aplicativos inteligentes



Engenheiro de ML
Implantação de modelos escalonáveis



Cientista de dados
Criação de modelos eficientes



Engenheiro de dados
Recebimento de dados limpos e úteis



Agente de engenharia de dados

Os engenheiros de dados enfrentam diversos desafios causados pelo esforço manual necessário para criar, corrigir e modernizar infraestruturas de dados. A criação de pipelines complexos, a resolução de falhas, a migração de sistemas legados e a administração de ingestão de dados complexos são tarefas demoradas, que consomem tempo que poderia ser gasto em inovação de alto valor.

O [agente de engenharia de dados](#) do Google é um parceiro inteligente que automatiza esse ciclo de vida do começo ao fim. Na criação de pipelines, ele se baseia em vários nós e gera documentação de código. Na solução de problemas, ele encontra a correção dos jobs que você escolher.

Na migração de sistemas, ele traduz o código legado do Spark ou local em pipelines modernos e padronizados para acelerar o processo. Por fim, para simplificar a ingestão de dados complexos, ele analisa formatos específicos do setor e busca dados de APIs.

Com a automatização dessas tarefas enfadonhas, o agente permite aos engenheiros se concentrarem em tarefas mais estratégicas, aumentando a produtividade e gerando dados de maior qualidade mais rápido e em maior escala.

Exemplos de comandos

"Extraia os dados da tabela sales.customers na região us_west_1 e envie-os para a tabela reporting.dim_customers no BigQuery. Faça a correspondência com o esquema da tabela de destino."

"Crie um pipeline seguindo estas etapas:

Extraia os dados da tabela ecomm.orders

Mescle os dados extraídos com a tabela marts.customers em customer_id

Carregue o resultado final na tabela reporting.customer_orders."

Demonstração



Como o agente de engenharia de dados usa a IA para gerar SQL e criar pipelines de dados.

[Assistir](#)

Agente de ciência de dados

Os cientistas de dados precisam lidar com tarefas manuais e repetitivas que prejudicam a produtividade e atrasam os insights. A natureza extensa da preparação de dados, a análise detalhada e o desenvolvimento de modelos complexos e iterativos consomem um tempo valioso.

O [agente de ciência de dados](#) do Google está integrado diretamente aos notebooks do Colab para simplificar essas tarefas demoradas. Em uma única interface, o agente de ciência de dados oferece assistência de IA com tecnologia do Gemini para automatizar a análise de dados, gerar visualizações e escrever código Python ou SQL. Ele também acelera a produção ao fornecer acesso unificado ao BigQuery e à Vertex AI e dá acesso a ferramentas poderosas, como o TimesFM pré-treinado, para previsão e avaliação de modelos integrada diretamente ao seu notebook do Colab.

Com a automatização dessas tarefas complexas de configuração, o agente permite que os cientistas de dados trabalhem com modelos prontos para produção em vez de dados brutos. Assim, eles podem se concentrar no que importa de verdade: refinar a precisão do modelo e fornecer insights comerciais de alto impacto ainda mais rápido.

Exemplos de comandos

"Crie e avalie um modelo de classificação em `bigquery-public-data.ml_datasets.census_adult_income` usando o BigQuerySQL. Usando SQL, preveja o tráfego futuro do meu site para o próximo mês com base em `bigquery-public-data.google_analytics_sample.ga_sessions_*`. Em seguida, crie uma representação gráfica dos valores históricos e estimados."

"Crie um DataFrame do pandas para os dados em `project_id:dataset.table`. Analise os dados para verificar se há valores nulos. Em seguida, crie uma representação gráfica da distribuição de cada coluna usando o tipo de gráfico. Use gráficos de violino para valores medidos e gráficos de barras para categorias."

Demonstração



Como o agente de ciência de dados ajuda a detectar anomalias em conjuntos de dados

[Assistir](#)

Agente de análises de conversação

Usuários comerciais e analistas de dados nem sempre conseguem responder a todas as perguntas feitas usando os painéis tradicionais, gerando atrasos e dependência em equipes técnicas para relatórios ad hoc. Esse problema prejudica o acesso a insights valiosos e reduz o ritmo dos negócios.

O [agente de análises de conversação](#) do Google transforma essa experiência porque permite ao usuário "conversar com os dados". Com tecnologia do Gemini e embasamento na camada semântica confiável do Looker, esse agente permite aos usuários fazer perguntas e acompanhamento em linguagem natural para receber respostas, gráficos e análises de dados instantâneos e confiáveis. Isso vai além dos relatórios estáticos, permitindo um diálogo dinâmico e interativo com os dados.

Com a democratização do acesso aos dados, o agente permite que as equipes de negócios usem insights confiáveis por conta própria e que os analistas de dados eliminem tarefas repetitivas e se concentrem em trabalhos de alto impacto.

Exemplos de comandos

"Qual é a tendência de devoluções de produtos ao longo do tempo, segmentada por departamento de produtos?"

"Como o preço médio dos produtos em cada departamento mudou ano a ano?"

Demonstração



Faça perguntas em linguagem simples e receba respostas diretamente dos seus dados do BigQuery

[Assistir](#)



Como criar agentes de dados no Google Cloud

Cada empresa é única. Por isso, nem sempre elas encontram agentes pré-criados que atendam às suas necessidades. A criação de um agente de dados personalizado é uma solução poderosa e adaptada, mas sair de uma demonstração promissora para um sistema de produção pode ser uma tarefa desafiadora. O desenvolvimento é complexo, mas a produção é ainda mais, com obstáculos que incluem garantir a segurança, gerenciar custos e gerar resultados consistentes e confiáveis.

O Google Cloud simplifica essa jornada, pois proporciona um caminho estruturado de cinco etapas. Esse plano ajuda você a criar, implantar e gerenciar agentes robustos e escalonáveis com confiança, transformando uma visão de solução personalizada em uma realidade de produção.

01

Contexto

Forneça informações relevantes ao agente

02

Configuração

Frameworks, ferramentas, modelos e proteções

03

Implantação

Executar, escalar e monitorar o agente

04

Publicação

Tornar os agentes acessíveis

05

Análise

Uso, custo e desempenho do agente



Quer saber se você consegue criar um agente de dados?

Antes de começar o desenvolvimento, é fundamental garantir que o projeto possa gerar bons resultados. Um ótimo agente precisa de mais do que uma tecnologia boa: precisa de uma base organizacional e de dados sólida.

Refleta sobre algumas questões importantes:

01

Você identificou um caso de uso comercial claro, com resultados mensuráveis?

02

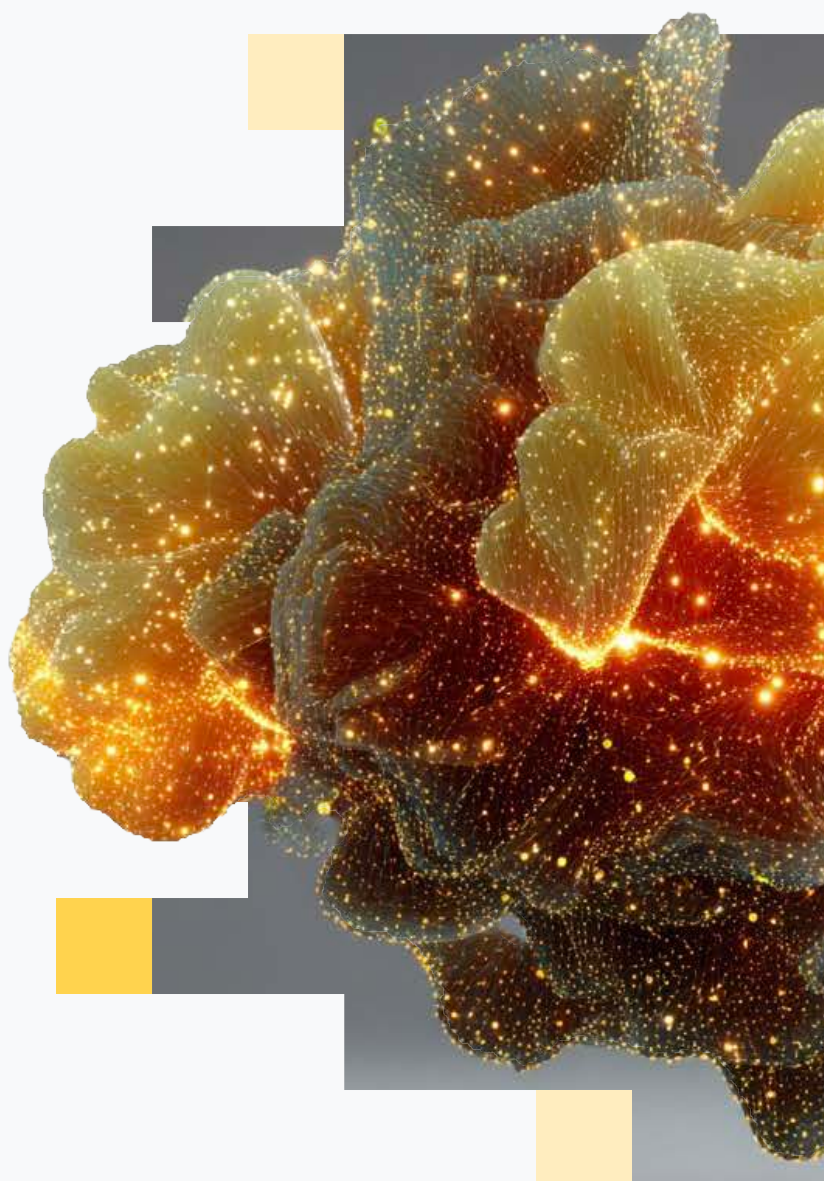
Você tem líderes de negócios e de dados que podem patrocinar o agente?

03

Você tem a experiência interna para criar um agente ou vai precisar de suporte?

04

Os dados estão bem documentados e o contexto da empresa é compreendido? A disposição de investir na preparação dos dados costuma ser o maior indicador de sucesso.





Forneça um contexto comercial detalhado

O sucesso de um agente de dados depende de um elemento fundamental: o contexto. Sem contexto, mesmo os modelos de IA mais avançados ficam perdidos. Para conseguir resultados precisos e confiáveis, é preciso criar uma "camada de contexto comercial", uma ponte que conecta os dados brutos ao mecanismo de raciocínio do agente.

O Google Cloud oferece um pacote de ferramentas projetadas para criar essa camada com o máximo de automação.

[Aumento e catalogação dos metadados:](#) recupere metadados de recursos do Google Cloud, como BigQuery, Cloud SQL, Spanner, Vertex AI, Pub/Sub, Dataform ou Dataproc Metastore, e recursos de terceiros incluídos no Dataplex Universal Catalog para gerar um catálogo de dados instantâneo. Aumente a capacidade dos agentes com dados operacionais e receba fontes de verdade da empresa em tempo real.

[Descoberta de dados:](#) verifique bancos de dados, como o AlloyDB e o Bigtable, em busca de dados estruturados, e buckets do Cloud Storage, em busca de dados não estruturados, para extrair e catalogar os metadados.

[Camada semântica:](#) gerencie a terminologia e as definições relacionadas aos negócios em toda a organização. Crie cálculos, relações e lógicas sob medida para algoritmos e modelos de IA exclusivos. Traduza dados complexos em termos comerciais, permitindo a exploração e visualização de dados com linguagem natural.

[Insights de dados:](#) use a IA para gerar perguntas em linguagem natural sobre seus dados e descobrir padrões, avaliar a qualidade dos dados e realizar análises estatísticas.

[Criação de perfil de dados:](#) identifique características comuns dos dados das colunas nas suas tabelas do BigQuery, como valores de dados típicos, distribuição de dados e contagens de nulos, para classificar os dados e garantir a qualidade.

[Qualidade dos dados:](#) defina e meça a qualidade dos dados nas suas tabelas do BigQuery validando os dados em relação às políticas da organização e registrando alertas quando os dados não atenderem aos critérios de qualidade.

[Linhagem de dados:](#) acompanhe como os dados se movimentam nos seus sistemas: de onde vêm, para onde vão e quais transformações são aplicadas.



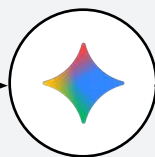


Entrada

Pergunta

Contexto

Mecanismo de raciocínio



Recursos

Resultado

Mensagens

Detecção de esquemas

Raciocínio e reflexão

Previsão

Dados de amostra

Consulta de dados

Análise dos principais fatores

Pesquisa imprecisa

Geração de gráficos

Detecção de anomalias

Consultas de ouro

Fornecimento de insights

Comparação de períodos

Conectores de dados



Looker



BigQuery



AlloyDB



Dados e contexto comerciais



A definição de relações entre tabelas, o fornecimento de exemplos de consultas de poucos disparos (few-shot) e a elaboração de instruções claras dá ao agente as diretrizes necessárias para entender seu cenário comercial. Essa abordagem automatizada transforma dados confusos e enigmáticos em uma base confiável, sendo a etapa mais importante na criação de um agente que gera valor com consistência.

Instruções e informações para orientar o agente de dados

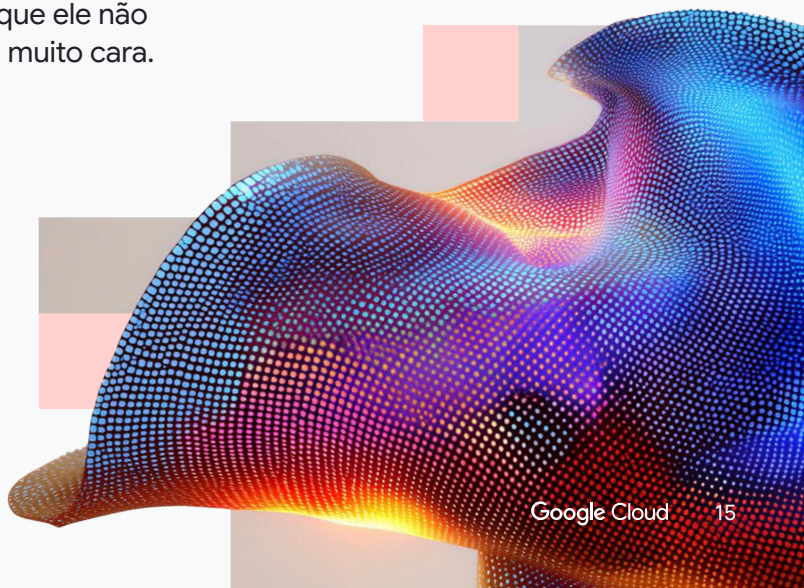
Instruções e fluxo de trabalho geral  O que é: processo geral e práticas recomendadas que o agente deve seguir Como fazer: escreva instruções ou comandos detalhados em um arquivo de configuração baseado em texto	Exemplos de poucos disparos (few-shot)  O que é: 5 a 10 exemplos de SQL de consultas comuns para o agente Como fazer: usar insights do conjunto de dados do BigQuery/mecanismo de conhecimento e repositório de exemplos	Contexto da tabela do BigQuery  O que é: receber informações sobre as tabelas (esquema, descrições de colunas, metadados, tipos de dados etc.) Como fazer: comandos do BigQuery gerados pelo modelo
Informações específicas de caso de uso  O que é: perguntas comuns, terminologia de casos de uso, campos importantes Como fazer: coleta manual de PMEs e do glossário do Dataplex	Relações entre tabelas  O que é: como as tabelas estão relacionadas (relações de entidade, chaves de junção, tipos etc.) Como fazer: extrair insights de dados do BigQuery	Perfis de dados  O que é: resultados do perfil dos dados (estatísticas de colunas, distribuição de valores, valores distintos, contagem de linhas, tamanho da tabela, padrões de dados etc.) Como fazer: perfis dos dados do BigQuery

Depois de fornecer o contexto, configure proteções para monitorar o que o agente está fazendo com os dados, garantir que ele não exclua os dados e evitar a execução de uma consulta muito cara.

Blog de investigação



[10 dicas para proteger os dados e as carteira dos agentes](#)





Como criar e configurar um agente de dados

Depois de estabelecer o contexto, a próxima etapa é configurar os principais componentes do agente. A abordagem do Google Cloud é baseada na flexibilidade e centrada em APIs e tecnologias de código aberto, como:

[API Conversational Analytics](#): permite incorporar a funcionalidade de consultas em linguagem natural nos agentes ou fluxos de trabalho, sendo compatível com acesso confiável aos dados e modelos de dados escalonáveis e robustos. Trata-se da mesma API que viabiliza as experiências de conversação prontas para uso no Looker e no BigQuery.

Ela permite criar experiências de dados personalizadas que fornecem dados, gráficos e respostas em texto, com base no modelo semântico confiável do Looker para garantir acurácia ou para fornecer contexto crítico de negócios e dados a agentes no BigQuery. Você pode incorporar essa funcionalidade para criar experiências de dados intuitivas, viabilizar análises complexas com linguagem natural e até mesmo orquestrar outros agentes de análise de conversação.

Codelab



Crie um app de chat com a Conversational Analytics API e o Looker

[Conheça o codelab](#)

[Kit de Desenvolvimento de Agente \(ADK\)](#): atua como a estrutura básica para a criação de agentes, fornecendo a estrutura de código principal, o gerenciamento do ciclo de vida e as ferramentas de avaliação necessárias para criar agentes confiáveis em Python, Java ou Go.

[Protocolo de Contexto de Modelo \(MCP\)](#): funciona como o "conector" padrão para que o agente se conecte com segurança a fontes de dados e ferramentas externas, como bancos de dados ou APIs corporativas, sem precisar de integrações personalizadas para cada novo serviço.

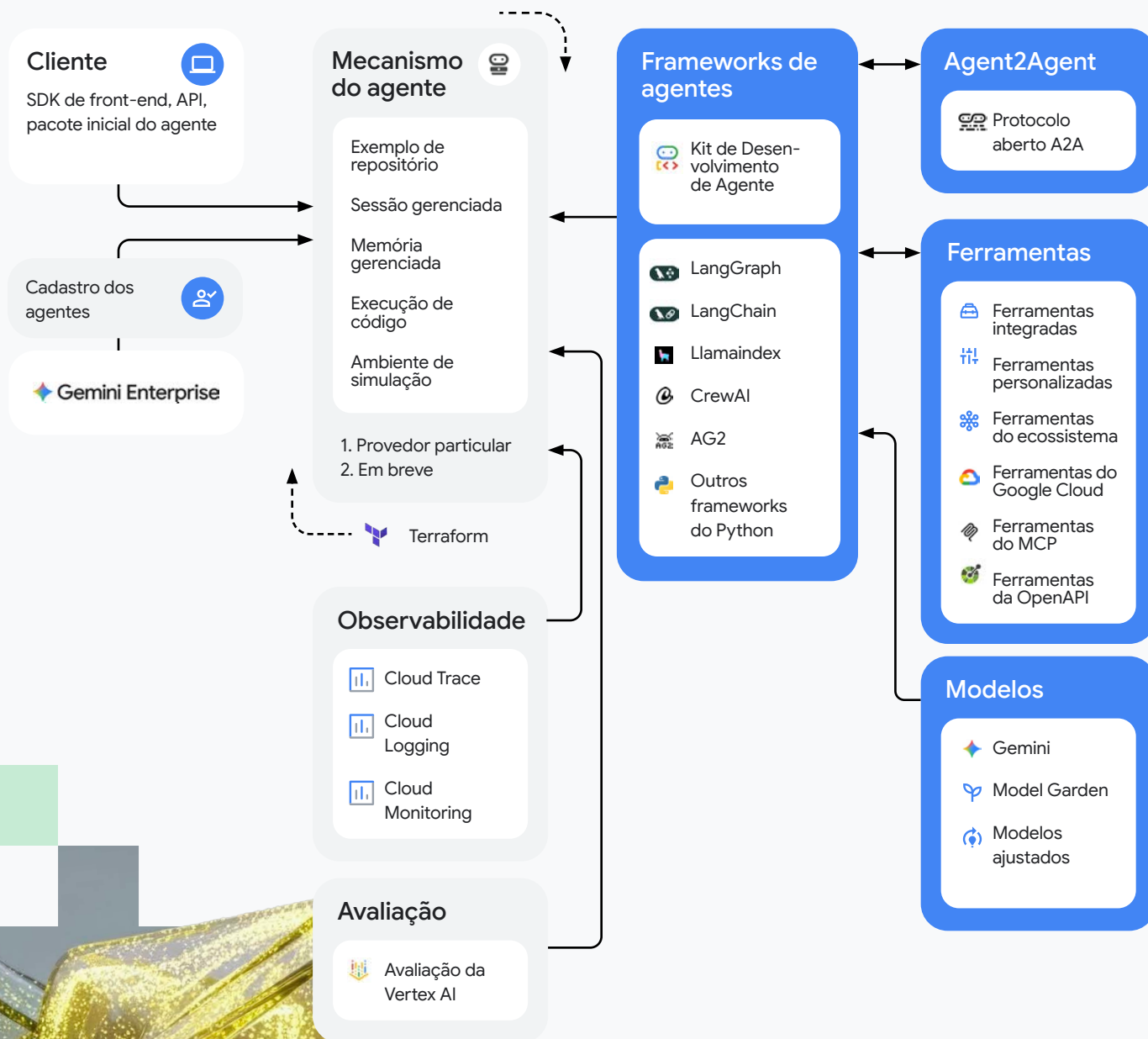
[Protocolo Agent-to-Agent \(A2A\)](#): atua como a camada de rede, permitindo que agentes individuais, independentemente de como foram criados, conversem, troquem informações e colaborem em fluxos de trabalho complexos e em várias etapas como uma equipe unificada.





Esse framework full stack simplifica o desenvolvimento de agentes e é multiagente por padrão, permitindo que você escolha os melhores modelos e ferramentas do Gemini para atender às necessidades específicas da sua organização.

Frameworks, ferramentas e modelos para a configuração do seu agente





O Google Cloud oferece um conjunto de ferramentas do BigQuery para simplificar a interação dos agentes de IA com os dados da empresa. Com essas ferramentas, os desenvolvedores não precisam criar conectores de banco de dados personalizados do zero e podem trabalhar com métodos de integração do ADK e do MCP.

Integração do ADK

Para os desenvolvedores que criam agentes usando o ADK do Google, as ferramentas de integração do [BigQuery](#) e do [Spanner](#) estão disponíveis como "habilidades" pré-criadas que podem ser atribuídas a um agente em segundos. Esse conjunto de ferramentas integrado inclui funções prontas para uso, permitindo que os agentes explorem dados, entenda esquemas e definições de colunas e executem consultas com autonomia.

Com esse conjunto de ferramentas, um agente ganha a capacidade de "explorar" um banco de dados e responder a perguntas em linguagem natural, como "qual foi o total de vendas no último trimestre?", sem que o desenvolvedor precise escrever códigos de geração de SQL.

Codelab



Saiba como criar agentes que conseguem responder a perguntas sobre dados armazenados no BigQuery usando o Kit de Desenvolvimento de Agente (ADK).

[Conheça o codelab](#)

Integração com o MCP

O [MCP no Google Cloud](#) permite que os desenvolvedores conectem um agente de IA ou um cliente MCP padrão, como a CLI do Gemini ou o Antigravity, a um endpoint consistente no mundo todo e pronto para empresas dos serviços do Google Cloud, como BigQuery e AlloyDB. Assim, os agentes interpretam esquemas e executam consultas em dados comerciais sem os riscos de segurança ou a latência associados à movimentação de dados para janelas de contexto. Assim, o MCP no Google Cloud fornece acesso direto a recursos do BigQuery, como a previsão, enquanto garante que os dados permaneçam no local com segurança e sejam governados de forma adequada.

Para mais flexibilidade e controle, use o [MCP Toolbox](#), um servidor de código aberto que centraliza a hospedagem e o gerenciamento de conjuntos de ferramentas, separando aplicativos agênticos da interação direta com o banco de dados. Em vez de gerenciar a lógica e a autenticação das ferramentas, os agentes atuam como clientes do MCP e solicitam ferramentas do MCP Toolbox for Databases. Ele também está disponível com diversos ambientes de desenvolvimento integrado (IDEs, na sigla em inglês) e ferramentas de desenvolvedor, incluindo a [CLI do Gemini](#) e o [Antigravity](#), permitindo que você conecte seus agentes de IA com segurança a serviços como [AlloyDB](#), [BigQuery](#), [Spanner](#), [Looker](#) e muito mais.

Codelab



Saiba como disponibilizar conjuntos de dados do BigQuery usando o MCP Toolbox for Databases.

[Conheça o codelab](#)



Como implantar e executar agentes de dados em escala

Para implantar um agente, você precisa de mais do que apenas um servidor: precisa de um ambiente robusto e pronto para produção. O [Vertex AI Agent Engine](#) é o ambiente de execução de preferência do Google Cloud para agentes porque foi projetado especificamente para essa finalidade. Como um serviço totalmente gerenciado, o Vertex AI Agent Engine elimina a sobrecarga operacional porque cuida da segurança, da autenticação e do escalonamento automaticamente, permitindo que o agente opere com confiança sob o tráfego real.

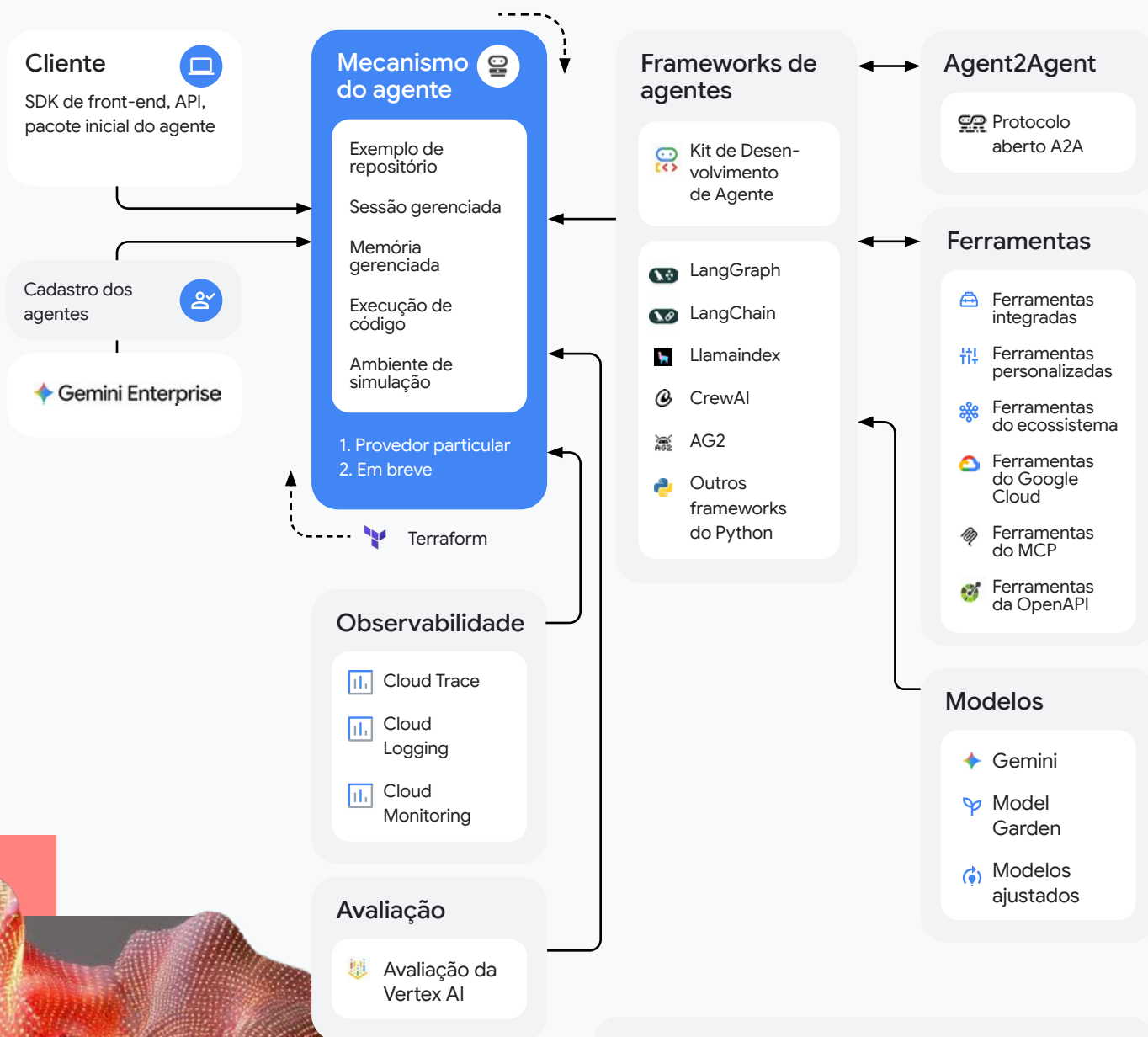
O design independente de framework do Vertex AI Agent Engine também permite implantar agentes criados com ADK, LangChain ou frameworks personalizados, oferecendo máxima flexibilidade sem sobrecarregar a infraestrutura.

Embora o Vertex AI Agent Engine seja recomendado para a maioria dos casos de uso, o [Cloud Run](#) oferece flexibilidade máxima para quem busca controle preciso do contêiner ou níveis máximos de portabilidade. É simples: se você quer reduzir o tempo de lançamento e aproveitar um ecossistema empresarial gerenciado, o Vertex AI Agent Engine é a melhor opção. Se você se preocupa mais com políticas de segurança rigorosas ou binários personalizados, o Cloud Run é a opção que vai proporcionar o nível de controle que você procura.





Ferramentas para implantar, executar e avaliar o agente



Codelab



Siga nosso codelab para aprender a criar um sistema multiagente com o ADK, o Agent Engine e o AlloyDB.

[Conheça o codelab](#)



Publicação dos agentes para torná-los acessíveis

O valor de um agente só é percebido quando é possível acessá-lo com facilidade. O Google Cloud transforma agentes personalizados de projetos isolados em recursos empresariais detectáveis com o [Gemini Enterprise](#) ou suas próprias plataformas de distribuição, que atuam como uma "app store de agentes" interna.

O processo é bastante simples. Depois de implantar o agente em um ambiente de execução gerenciado, como o Vertex AI Agent Engine, use uma ferramenta de registro para publicar o agente formalmente.

Com essa etapa simples, os recursos e o endpoint do seu agente são disponibilizados para toda a organização, transformando uma ferramenta personalizada em um serviço reutilizável.

Depois que o agente é publicado no Gemini Enterprise, ele aparece na [Galeria de Agentes](#). Nesse hub centralizado, os usuários podem navegar, pesquisar e conhecer os agentes disponíveis e têm acesso a descrições detalhadas das funções, aos proprietários e às instruções de uso. A galeria vai além da descoberta: é um ambiente governado em que os administradores gerenciam permissões para garantir que os usuários certos tenham acesso aos agentes apropriados e às respectivas fontes de dados. Essa abordagem estruturada fomenta um ecossistema seguro e colaborativo, evitando esforços de desenvolvimento redundantes e maximizando o impacto de cada agente personalizado criado.



Publicar os agentes no Gemini Enterprise



AgentOps: analisar e otimizar o uso, o custo e o desempenho

Com o agente de referência configurado, a pergunta é: "esse agente funciona mesmo em um ambiente de produção real e consegue mesmo lidar com consultas complexas e cheias de nuances feitas por usuários reais?" Especificamente, ele tem mesmo a capacidade de identificar as tabelas corretas em um esquema, gerar consultas SQL acuradas e, o mais importante, resistir à tendência crítica de alucinar? É preciso realizar uma avaliação abrangente para validar a prontidão dos agentes.

Embora o Kit de Desenvolvimento de Agente (ADK) inclua um serviço de avaliação, ele está limitado a métricas que avaliam apenas a resposta final do agente. O [serviço de avaliação de IA generativa da Vertex AI](#) permite avaliar e entender os agentes de IA de um jeito mais rigoroso. Ele conta com um conjunto avançado de métricas de avaliação feitas especialmente para agentes criados com diferentes frameworks e inclui recursos integrados de inferência para simplificar o processo de avaliação.

A criação de um agente é um processo iterativo. Para melhorar um agente criado, é preciso entender como ele é usado. Antes, a coleta de dados sobre o uso, o custo e o desempenho dos agentes exigia um grande esforço da equipe de engenharia na criação de pipelines de registro e painéis personalizados.

O Google Cloud elimina esse desafio com o [plug-in de análise de agentes do BigQuery](#), uma solução eficiente e pronta para uso. Com esse plug-in, você ativa um pipeline de dados escalonável em segundos e pode capturar, analisar e visualizar o uso, o custo e o desempenho do agente.

Essa ferramenta permite passar da implantação à análise em menos de cinco minutos. É possível monitorar os custos, entender o comportamento do usuário e identificar quais ferramentas são usadas com mais frequência. Ter acesso a esses insights é fundamental para a criação de um ciclo de feedback eficiente que permita usar dados reais para melhorar continuamente a qualidade do agente e demonstrar o valor que ele tem para a empresa.

Codelab



Aprenda a avaliar os agentes usando o serviço de avaliação de IA generativa da Vertex AI.

[Conheça o codelab](#)

Codelab



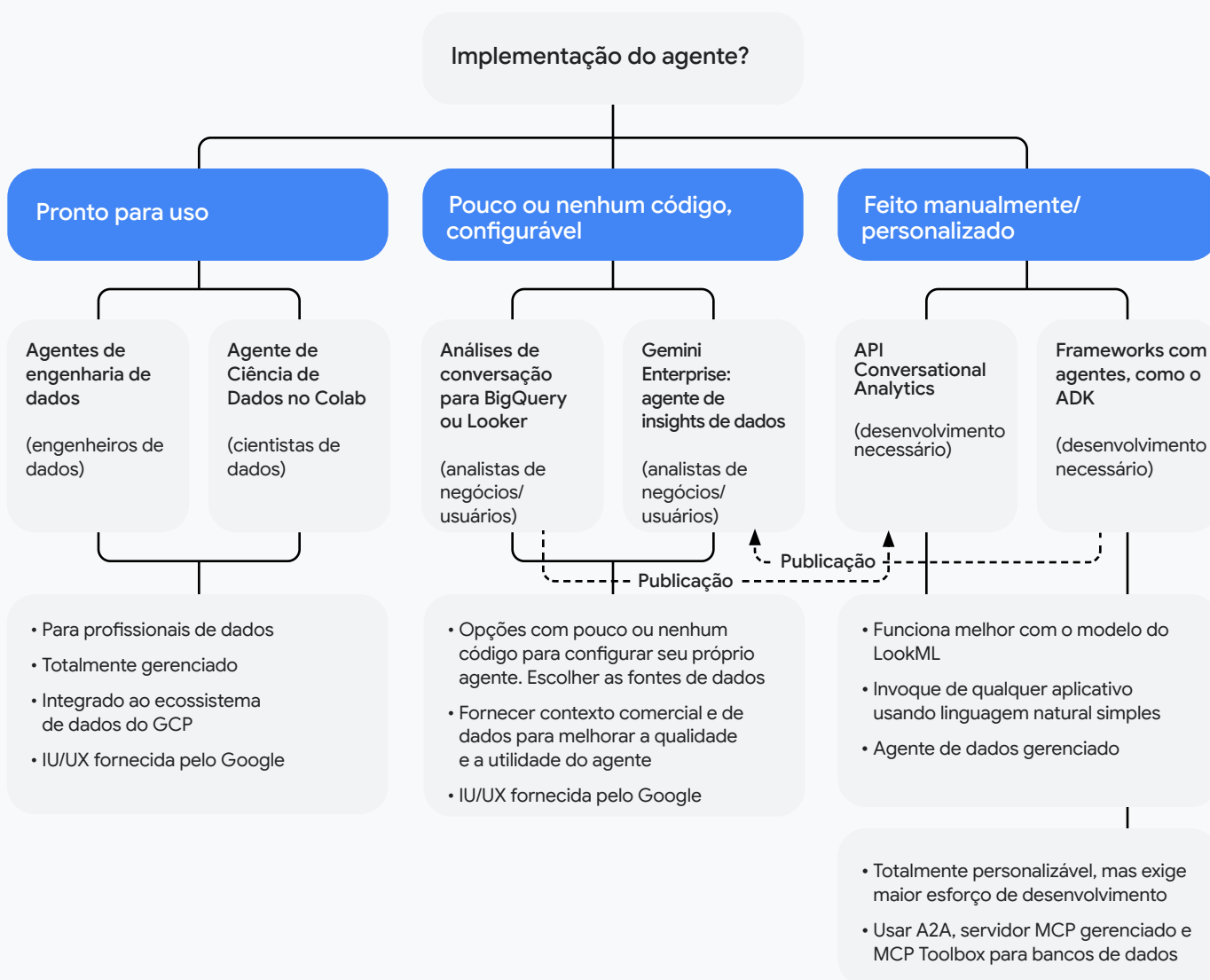
Saiba como ativar a observabilidade do agente para analisar traces de conversas e o uso de ferramentas pelo agente.

[Conheça o codelab](#)



Um resumo da abordagem do Google Cloud aos agentes de dados

Abordagem do Google Cloud aos agentes de dados



A abordagem do Google Cloud aos agentes de dados se baseia em uma filosofia de escolha e oferece várias soluções desenvolvidas para capacitar todos os usuários, desde analistas de negócios até desenvolvedores especializados. Assim, as equipes podem adotar a IA generativa da maneira mais adequada ao caso de uso e às capacidades técnicas.



A abordagem é dividida em dois caminhos principais:

01

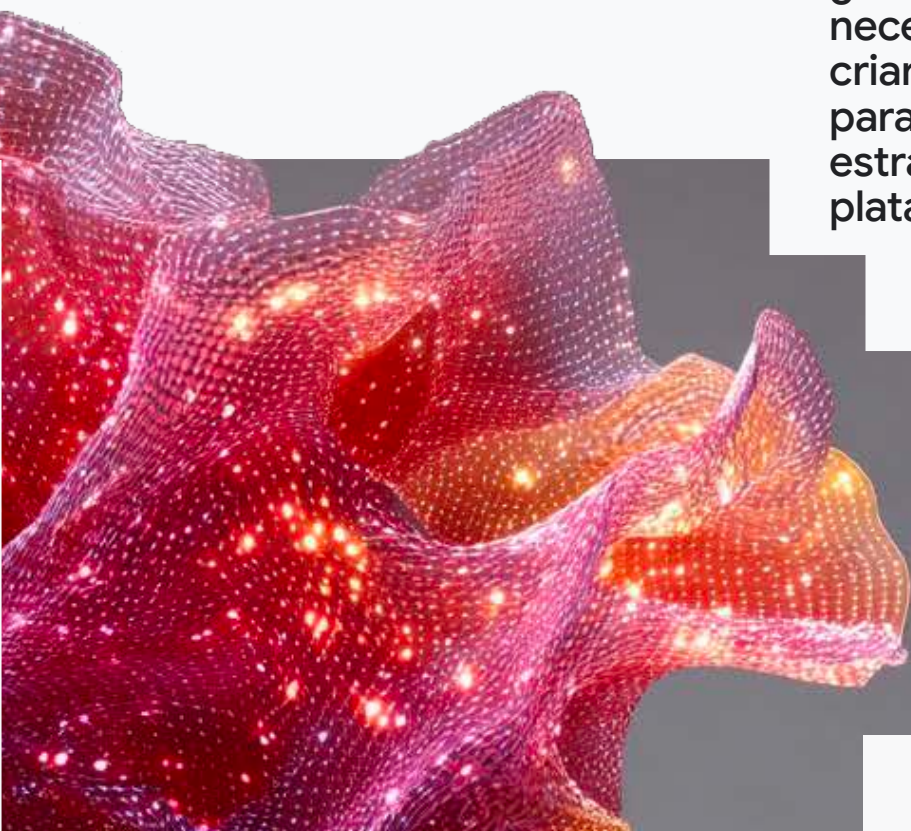
Agentes pré-criados, totalmente gerenciados e prontos para uso: para profissionais de dados que buscam acelerar os fluxos de trabalho, o Google Cloud oferece um conjunto de agentes pré-criados e específicos para cada função. Atualmente, eles incluem o [Agente de Engenharia de Dados](#) e o [Agente de Ciência de Dados](#).

Essas são soluções totalmente gerenciadas, integradas ao ecossistema do Google Cloud e desenvolvidas para acelerar a geração de valor com o mínimo de esforço. Para usuários comerciais, o [Agente de Análises de Conversação](#) torna a interação com os dados mais imediata e intuitiva.

02

Agentes personalizados: para desenvolvedores que buscam máxima flexibilidade e controle, o Google Cloud oferece os elementos básicos para a criação de experiências de agentes personalizadas. Com o [ADK](#), o [MCP](#) e a [API Conversational Analytics](#), as equipes criam agentes totalmente personalizáveis com os modelos, frameworks e lógica que quiserem e podem incorporá-los diretamente nos próprios aplicativos.

Essa abordagem de dois caminhos significa que você nunca vai precisar usar uma solução única para todos os casos de uso. Você pode começar usando um agente gerenciado para resolver uma necessidade imediata e depois criar um agente personalizado para uma oportunidade estratégica única, tudo em uma plataforma unificada.





Capítulo 05:

Desenvolva uma liderança de IA

O Google Cloud oferece a única pilha totalmente integrada e otimizada para IA do setor, projetada especificamente para a era da IA generativa. Ao contrário de soluções fragmentadas que exigem a integração de bancos de dados vetoriais, mecanismos de inferência e modelos de fundação de fornecedores diferentes, o Google Cloud oferece tecnologia própria em todas as camadas da pilha. Isso vai desde uma infraestrutura de IA segura e personalizada até as análises de dados, camadas semânticas e plataformas de ML que alimentam os próprios agentes. Essa abordagem completa elimina a dívida técnica, reduz a latência e garante o desempenho máximo. Isso cria um ambiente para desenvolver e implantar agentes que podem raciocinar, planejar e agir com rapidez.

Na base, nossa arquitetura de Hipercomputador de IA integra hardware com desempenho otimizado, como TPUs e GPUs líderes do setor, com software aberto e modelos de consumo flexíveis. Essa é a mesma infraestrutura usada para treinar e disponibilizar os modelos mais avançados do Google, garantindo que os agentes operem em uma base projetada para entregar alta capacidade de processamento e raciocínio de baixa latência.

A qualidade dos agentes depende dos dados que podem acessar. O Google Cloud cria um ciclo de vida de dados para IA único no mercado atual. Ao unificar o armazenamento em data warehouse sem servidor do BigQuery e o banco de dados de alto desempenho do AlloyDB com a modelagem semântica do Looker, oferecemos aos agentes definições de negócios de confiança. Isso permite que os agentes respondam de acordo com sua realidade operacional, entendendo métricas, governança e fatos em tempo real, em vez de depender apenas do conhecimento pré-treinado. Essa integração profunda de dados reduz alucinações e garante que seus agentes de dados estejam prontos para uso na empresa.

Esses agentes são alimentados pela família de modelos do Gemini, que oferece competências multimodais líderes do setor e grandes janelas de contexto. Dessa forma, os agentes podem processar grandes quantidades de dados não estruturados, como documentos, códigos, vídeos ou áudios, além de dados da empresa estruturados para maximizar o impacto.

Também é possível personalizar esses modelos facilmente usando a Vertex AI e mecanismos de raciocínio avançado, permitindo aos agentes dividir metas complexas do usuário em fluxos de trabalho de várias etapas.

O Google Cloud tem mais de 20 anos de liderança em dados e IA. Essa é a mesma tecnologia que sustenta a Pesquisa Google e o YouTube, entregando aos clientes grande economia e confiabilidade incomparáveis diariamente. O Google foi pioneiro na arquitetura Transformer, que deu início à revolução da IA generativa, e continua liderando nesse espaço com as inovações do DeepMind. Com o Google Cloud, você ganha mais do que apenas ferramentas individuais: você tem acesso a um ecossistema completo e flexível com caminhos práticos para começar mais rápido. Essa combinação de uma pilha unificada e superior e décadas de experiência comprovada em dados e IA faz do Google Cloud a solução ideal para criar, gerenciar e escalar seus agentes inteligentes, acelerando a geração de valor comercial real.



Crie seus apps inteligentes de IA no lugar certo

Com uma plataforma de IA abrangente, flexível e integrada, da infraestrutura a soluções personalizadas, o Google Cloud oferece um ecossistema completo, não apenas ferramentas.



Aproveite a experiência mais avançada em IA

Tendo como exemplo o produto de IA mais bem-sucedido do mundo, a Pesquisa Google, o Google oferece aos clientes escala incomparável, profundidade de pesquisa e economia em grande escala.



Comece agora mesmo

O Google Cloud oferece caminhos práticos (além de ferramentas, recursos e suporte) para começar e acelerar o tempo de retorno do investimento com a IA em um plano de 30 dias.

Entre na era dos agentes



Neste guia, definimos o que é um agente de dados, analisamos os agentes avançados e orientados por persona disponíveis no Google Cloud e definimos as etapas da criação do seu próprio agente. Lembre-se: a era dos agentes não é um futuro distante: ela já chegou e oferece soluções tangíveis às enormes pressões que as equipes de dados enfrentam atualmente.

A próxima etapa é passar do aprendizado para a implementação.

Sua transformação agêntica começa hoje.

Experiência prática

O melhor jeito de entender o poder desse novo paradigma é testá-lo. Os agentes précriados do Google Cloud para engenharia, ciência de dados e análise estão integrados diretamente ao BigQuery Studio e prontos para uso. Você já pode começar a automatizar pipelines, gerar insights e fazer perguntas sobre seus dados em linguagem natural para ver como a IA pode transformar seus fluxos de trabalho.

Experimente o BigQuery sem custo financeiro



Experimente o AlloyDB sem custo financeiro



Experimente o Looker sem custo financeiro



Planeje sua estratégia com nossos especialistas

O caminho de cada organização rumo à adoção da IA generativa é único. Se quiser conversar sobre como os agentes de dados podem resolver seus desafios específicos e ajudar a sua empresa a alcançar as metas de negócios, nossa equipe está aqui para ajudar. Vamos criar sua estratégia em conjunto.

Entre em contato para agendar uma consultoria personalizada



Colaboradores

Manoj Gunti,
Sr Product Marketing Manager,
Google Cloud

Geeta Banda,
Head of Outbound Product Management,
Google Cloud

Sean Zinsmeister,
Director of Outbound Product Management,
Google Cloud

Sean Rhee,
Senior Product Manager,
Google Cloud

Ani Jain,
Senior Outbound Product Manager,
Google Cloud