



O'REILLY®

Guia do desenvolvedor para criação de aplicações de IA

Segunda edição

Crie sua primeira aplicação
conversacional com a IA do
Microsoft Azure

Elaine Chang e Darren Jeford

RELATÓRIO

SEGUNDA EDIÇÃO

Guia do desenvolvedor para criação de aplicações de IA

*Crie sua primeira aplicação
conversacional com a IA do
Microsoft Azure*

Elaine Chang e Darren Jeford

Pequim • Boston • Farnham • Sebastopol • Tóquio

O'REILLY[®]

Guia do desenvolvedor para criação de aplicações de IA

por Elaine Chang e Darren Jeford

Copyright © 2020 O'Reilly Media. Todos os direitos reservados.

Impresso nos Estados Unidos da América.

Publicado por O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472.

Os livros da O'Reilly podem ser adquiridos para uso educacional, comercial ou de vendas promocionais. As edições online também estão disponíveis para a maioria dos títulos (<http://oreilly.com>). Para obter mais informações, contate nosso departamento corporativo/vendas institucionais: 800-998-9938 ou corporate@oreilly.com.

Editora de aquisições: Rebecca Novack

Editora de desenvolvimento: Nicole Taché

Editor de produção: Christopher Faucher

Editor: Charles Roumeliotis

Revisora: Athena Lakri

Designer de interiores: David Futato

Designer de capa: Karen Montgomery

Ilustrador: Rebecca Demarest

Julho de 2018: Primeira edição

Abril de 2020: Segunda edição

Histórico de revisões da segunda edição

17-04-2020: primeira versão

O logotipo da O'Reilly é uma marca registrada da O'Reilly Media, Inc. O *Guia do desenvolvedor para criar aplicações de IA*, a imagem de capa e a imagem comercial relacionada são marcas comerciais da O'Reilly Media, Inc.

Os pontos de vista expressos neste trabalho são os dos autores e não representam as visões do editor. Apesar dos esforços de boa fé do editor e dos autores para garantir que as informações e as instruções contidas neste trabalho sejam precisas, o editor e os autores renunciam toda a responsabilidade por erros ou omissões, incluindo, sem limitação, a responsabilidade por danos resultantes do uso ou dependência deste trabalho. O uso das informações e instruções contidas neste trabalho é por seu próprio risco. Se houver exemplos de código ou outra tecnologia contida ou descrita nesse trabalho sujeitos a licenças de Open Source ou a direitos de propriedade intelectual de terceiros, você será responsável por garantir que o uso cumpra essas licenças e/ou direitos.

Este trabalho faz parte de uma colaboração entre a O'Reilly e a Microsoft. Confira nossa [declaração de independência editorial](#).

978-1-492-08058-9

[LSI]

Sumário

Prefácio	vii
-----------------------	------------

Guia do desenvolvedor para criação de aplicações de IA	1
Introdução	1
O cruzamento entre dados, IA e a nuvem	3
IA do Microsoft Azure.....	5
IA Conversacional	6
Principais recursos de assistentes virtuais.....	17
Desenvolver seu assistente virtual	23
Conectar assistentes a clientes e canais.....	24
Opcional: adicionar inteligência ao assistente com habilidades	26
Habilitar ferramentas analíticas para o assistente virtual.....	27
Roteiro e mais recursos	29
Criar IA responsável	30

Experimente a IA

Atenda melhor aos clientes com soluções com IA. Crie um chatbot. Treine e implante modelos de aprendizado de máquina. Descubra insights em seu conteúdo. Crie aplicações inteligentes.

Explore 12 serviços de IA: gratuitos por 12 meses com a sua conta. **Comece gratuitamente >**

Obtenha ajuda com seu projeto.
Fale com um especialista em vendas >

Agradecimentos

Agradecemos aos seguintes indivíduos da Microsoft que contribuíram com o relatório como assessores e revisores técnicos:

Lili Cheng, Anand Ramen, Ben Brown, Chris Mullins, Cindy Noteboom, Deborah Harrison, Dewain Robinson, Em Ivers, Lauren Mills, Patrick Volum, Robert Standefer, Ryan Lengel Isgrig, Steve Sweetman, Ted Li, Tyler Mays-Childers, William Mendoza e Prem Prakash.

Prefácio

Os criadores do computador moderno queriam aumentar o intelecto humano e expandir nossas capacidades além das limitações do poder intelectual coletivo. Quando pesquisadores e cientistas da computação na década de 1950 conceberam o "computador pensante", eles começaram uma rápida evolução rumo à IA, lançando as bases para a tecnologia que teria um enorme impacto no mundo nas próximas décadas.

A IA agora está em todos os lugares. As pessoas nem percebem que a IA está potenciando suas experiências. Além de estar presente em praticamente tudo que usamos, desde termostatos até sistemas de aspersão, a IA também potencializa oportunidades em novas fronteiras, como veículos autônomos, pesquisa farmacêutica e agricultura de precisão. A IA é uma ótima fronteira para os desenvolvedores criarem formas poderosas e inovadoras de abordar problemas difíceis de resolver.

Embora a IA fosse domínio apenas de pesquisadores e instituições, a disponibilidade recente de infraestrutura, plataformas e serviços de IA significa que todo o potencial da IA agora está disponível para desenvolvedores criarem soluções com inteligência. Os serviços e ferramentas de IA na nuvem viabilizam soluções baseadas em IA, incluindo aprendizado de máquina, mineração de conhecimento, agentes inteligentes e modelos pré-treinados, sem exigir conhecimento especializado.

Nos últimos anos, houve um aumento na adoção da IA, catalisado por vastos volumes de dados digitais, serviços de aplicações e enorme poder de computação. Como resultado, tecnologias de IA, como compreensão de linguagem natural, análise de sentimento, reconhecimento de fala e compreensão de imagens agora podem capacitar as aplicações em uma ampla gama de indústrias.

Uma das aplicações mais interessantes da IA é facilitar e melhorar a vida cotidiana. Desde o desenvolvimento da computação, as pessoas imaginam ter diálogos significativos com computadores, expressando necessidades e ideias na forma como nos comunicamos uns com os outros por meio da linguagem natural: diga algo ao computador e ele responderá. A IA conversacional muda o modelo de interação de comandos específicos de domínio, controlados por máquina para interfaces de conversação que se concentram em pessoas e expressão. Com a IA conversacional, os desenvolvedores podem fazer com que os computadores se comuniquem como pessoas, reconhecendo palavras, entendendo a intenção e respondendo de maneiras que pareçam naturais e familiares.

— *Lili Cheng, vice-presidente corporativa
de IA conversacional da Microsoft*

Guia do desenvolvedor para criação de aplicações de IA

Introdução

Neste livro, analisamos os requisitos para aplicar soluções de IA bem testadas a problemas cotidianos. Para ajudá-lo a explorar as possibilidades da IA, mostraremos como criar um assistente virtual, uma aplicação de IA conversacional que pode entender o idioma, perceber a vasta quantidade de informações e responder de forma inteligente. No processo, compartilharemos as diversas funcionalidades e recursos de IA disponíveis para desenvolvedores.

Vejamos um roteiro para o conteúdo deste livro:

“O cruzamento entre dados, IA e a nuvem”

Esta seção explica a base tecnológica para este livro e por que essas tecnologias são cada vez mais oferecidas na nuvem.

“IA do Microsoft Azure”

Esta seção apresenta a plataforma de IA do Microsoft Azure com uma variedade de serviços, infraestrutura e ferramentas para capacitar desenvolvedores e agentes para criar aplicações de IA e adicionar recursos de mineração de conhecimento e aprendizado de máquina. Este livro se concentra em aplicações de IA conversacionais e fornece indicações para recursos adicionais para outras áreas de IA do Azure.

“IA Conversacional”

Esta seção discute a evolução do processamento de linguagem natural, o serviço de Compreensão de Linguagem da Microsoft (antes chamado de LUIS) e o ecossistema do Bot Framework, casos de uso comum de IA conversacional e o ciclo de vida de desenvolvimento de aplicações de IA conversacional.

"Principais recursos de assistentes virtuais"

Esta seção destaca os principais recursos de assistentes virtuais, incluindo controle completo para desenvolvedores, habilidades de conversação predefinidas e reutilizáveis, integração flexível e conscientização contextual, insights de negócios capturados e distribuídos, entrada multimodal, cartões adaptáveis e recursos corporativos. O acelerador de solução de Assistente Virtual do Bot Framework reúne as práticas recomendadas para permitir que as organizações ofereçam experiências avançadas de assistente conversacional adaptado à sua marca, personalizadas para os usuários e disponibilizadas na gama de aplicações e dispositivos.

"Desenvolver seu assistente virtual"

Esta seção fornece orientações para a criação da aplicação de IA conversacional com o acelerador de solução de Assistente Virtual, com ponteiros para tutoriais online.

"Conectar assistentes a clientes e canais"

Esta seção fornece orientações para conectar a aplicação de IA conversacional com clientes e canais. A Microsoft oferece uma grande variedade de canais e adaptadores, permitindo que suas experiências de conversação alcancem usuários finais onde quer que estejam.

"Opcional: adicionar inteligência ao assistente com habilidades"

Esta seção apresenta habilidades, uma maneira de conectar os recursos da plataforma às experiências conversacionais, com ponteiros para tutoriais online.

"Habilitar ferramentas analíticas para o assistente virtual"

Esta seção fornece orientações para a configuração de ferramentas analíticas para a aplicação de IA conversacional. Os insights do painel de análise da aplicação de IA conversacional podem sugerir melhorias importantes na eficácia e na performance.

"Roteiro e mais recursos"

Saiba mais sobre o futuro investimento da Microsoft em IA conversacional.

"Criar IA responsável"

Concluiremos com uma discussão sobre como a plataforma de IA do Azure incentiva os desenvolvedores a criar soluções responsáveis e confiáveis que tratam pessoas de forma justa.

O cruzamento entre dados, IA e a nuvem

Hoje, possibilitamos que os computadores aprendam com grandes volumes de dados e interajam e respondam de forma mais natural com o mundo, em vez de seguir rotinas pré-programadas.¹ Considere os seguintes recursos de software moderno:

Visão computacional

A capacidade de "ver" ao reconhecer objetos e seus relacionamentos em uma pintura ou vídeo, criando dados a partir do ambiente físico.

Síntese e reconhecimento de fala

A capacidade de "ouvir" e entender as palavras que as pessoas dizem, de transcrevê-las em texto e vice-versa, e de ler o texto em voz alta com voz natural.

Compreensão de linguagem

A capacidade de "compreender" o significado das palavras e responder considerando as várias nuances e complexidades da linguagem (como gírias e expressões idiomáticas). Quando computadores podem participar de um diálogo com humanos de forma eficiente, isso é denominado *IA conversacional*.

Conhecimento

A capacidade de "deduzir" ao representar e compreender o relacionamento entre pessoas, coisas, lugares e eventos.

Como esses recursos são apresentados em aplicações corporativas? Por meio do aprendizado de máquina, as aplicações baseadas em IA *deduzem* ao revelar grandes volumes de dados variados — dados que foram coletados ao longo do tempo em repositórios e mantidos em conjuntos de dados massivos. Esses sistemas de IA *entendem* e criam significado em dados não estruturados, tais como email, chats e anotações manuscritas, itens que antes não poderiam ser processados por computadores. O mais importante é que esses sistemas agora *interagem* com clientes e os envolvem em canais diferentes, de forma que possam ser hiperpersonalizados.

¹ Lili Cheng, "Por que você não deve temer a Inteligência Artificial", *Time*, 4 de janeiro de 2018, <https://ti.me/2GEknZ>.

Ao aplicar esses recursos, as empresas usam aplicações baseadas em IA para transformar digitalmente todos os aspectos da organização. Elas estão transformando produtos por meio de insights de dados de clientes. Eles estão otimizando as operações de negócios ao prever anomalias e melhorar eficiências. Elas capacitam funcionários por meio de ferramentas inteligentes e envolvem clientes por meio de agentes conversacionais que proporcionam experiências mais personalizadas.

Para projetar a tecnologia para humanos, é importante entender o contexto de como as pessoas trabalham, atuam e vivem. As soluções de IA atuais complementam e revelam atividades criativas e o potencial humano. Essas soluções personalizadas também devem ser capazes de aprender e se adaptar a novas condições externas, assim como os seres humanos.

Uma das áreas mais fascinantes da pesquisa é unir as inteligências emocional e cognitiva para criar sistemas de IA conversacional que espelhem a linguagem humana e ter insights sobre as formas às vezes ilógicas e imprevisíveis de interação humana. Segundo Lili Cheng, vice-presidente corporativa de IA conversacional da Microsoft, "isso provavelmente significa que a IA precisa reconhecer quando as pessoas são mais eficazes por conta própria, quando não intervir, quando não ajudar, quando não gravar, quando não interromper ou distrair."²

Como os conjuntos de dados aumentam e, devido à extrema variação em tamanho, está se tornando mais importante oferecer aos desenvolvedores acesso rápido a opções flexíveis de armazenamento e processamento. Assim, as organizações adotam cada vez mais a nuvem, oferecendo essa variedade e flexibilidade. Além disso, os fornecedores de nuvem têm uma caixa de ferramentas rica e poderosa para habilitar os recursos de IA que discutimos. A enorme conectividade permite que qualquer tipo de dispositivo conectado traga grandes volumes de dados para a nuvem em tempo real para análise e processamento inteligente em grande escala. Para desenvolvedores, a nuvem fornece a infraestrutura e as ferramentas necessárias para oferecer segurança de nível empresarial, disponibilidade, conformidade e capacidade de gerenciamento para as aplicações e os serviços de negócios.

2 Lili Cheng, "Por que você não deve temer a Inteligência Artificial", *Time*, 4 de janeiro de 2018, <https://ti.me/2GEknZ>.

IA do Microsoft Azure

A plataforma de IA do Azure da Microsoft visa disponibilizar a inteligência artificial para desenvolvedores e capacitá-los para inovar e acelerar projetos com uma variedade de serviços, infraestrutura e ferramentas. A IA do Azure dá suporte a diversos casos de uso e aponta diferentes níveis de especialização e formas desejadas de trabalho. Por exemplo, o Azure oferece o SDK do Azure Bot Service e do Bot Framework, permitindo que desenvolvedores criem experiências de conversação sofisticadas. Além disso, os Serviços Cognitivos do Azure oferecem aos desenvolvedores serviços de IA específicos de domínio disponíveis como APIs para criar aplicações que podem ver, ouvir e entender. O Azure Machine Learning permite que desenvolvedores e cientistas de dados criem modelos de IA personalizados, com investimentos no hardware e na infraestrutura necessários para dar suporte à estrutura e ferramentas de aprendizado profundo e de aprendizado de máquina.

Agentes e aplicações de IA

Os modelos de IA líderes da indústria utilizados hoje em produtos da Microsoft como Office 365, Teams, Dynamics 365, Cortana, Xbox, Hololens e Bing estão disponíveis para suas próprias aplicações por meio de uma plataforma de serviços cognitivos democratizados. Alguns desses modelos podem ser personalizados com seus próprios dados e executados offline e online.

O livro *Criar aplicações inteligentes com APIs cognitivas* fornece uma análise mais detalhada que está por trás de aplicações que veem, ouvem, falam, entendem e interpretam as necessidades das pessoas.

O serviço de Bot do Azure, junto com o Bot Framework, permite que desenvolvedores criem experiências conversacionais avançadas. A nova oferta do Power Virtual Agents, disponível como parte do Power Platform, é criada com base no Bot Framework e capacita não desenvolvedores a criar experiências conversacionais e compor com outros componentes do Bot Framework.

Mineração de conhecimento

A Pesquisa Cognitiva do Azure opera em muitos tipos de dados para transformar informações não estruturadas em conteúdo pesquisável. Extraia insights e informações estruturadas, descubra padrões e relacionamentos, revele sentimento e muito mais.

O white paper *"Extrair insights acionáveis de todo o seu conteúdo"* abrange como a mineração de conhecimento funciona, casos de uso, soluções líderes da indústria e recursos adicionais para começar a usar a mineração de conhecimento.

Aprendizado de máquina

Os desenvolvedores podem obter acesso a recursos avançados de aprendizado de máquina de IA do Azure por meio dos serviços de Aprendizado de Máquina (AML) do Azure. O AML é um serviço de nuvem gerenciado em que você pode treinar, gerenciar e implantar modelos na nuvem ou em dispositivos de borda usando Python e ferramentas como Jupyter Notebook. Você pode até mesmo implantar modelos de classificação e reconhecimento de imagem TensorFlow, usando uma variedade de redes neurais profundas, para o hardware FPGA do Project Brainwave da Microsoft no Azure para inferência e treinamento, que oferece taxa de transferência extremamente alta e baixa latência.

O livro *Aprendizado de máquina reflexivo com Python: uma abordagem orientada a testes* fornece um ponto de partida para a programação de IA que pode ser útil para os leitores interessados em usar o AML.

Para ajudá-lo a começar a usar a IA do Azure, você pode aproveitar os recursos disponíveis no [site de IA do Azure](#).

Neste livro, vamos nos concentrar em mostrar como você pode criar uma aplicação de IA conversacional usando o Bot Framework.

IA Conversacional

O processamento de linguagem natural (NLP) fornece aos computadores a capacidade de ler, entender e derivar o significado da linguagem humana. Desde a década de 1950, os cientistas da computação têm trabalhado nos desafios do NLP, mas as limitações no poder de computação e em tamanhos de dados impediram avanços no processamento e análise de componentes textuais, sentimentos, partes da fala e as várias entidades que compõem a comunicação de linguagem natural.

Isso mudou na década de 2010. Os avanços em computação na nuvem, o aprendizado de máquina e a disponibilidade de grandes quantidades de texto digital e dados conversacionais de sistemas de mensagens e chats na Web ajudaram a fazer enormes progressos em NLP. Os avanços no NLP tornaram possível aos computadores não só identificar palavras no texto, mas também compreender o significado subjacente dessas palavras e as relações entre elas.

O NLP funciona analisando um grande corpo de texto gerado por humanos e transformando-o em dados legíveis por máquina. O NLP identifica e extrai metadados chave do texto, incluindo:

Entidades

O NLP identifica entidades em texto como pessoas, lugares e coisas. As entidades também podem ser informações que exigem extração especial, como datas e horários.

Relações

O NLP identifica como as entidades estão relacionadas usando informações semânticas.

Conceitos

O NLP extrai conceitos gerais do corpo do texto que não aparecem explicitamente. Por exemplo, a palavra "excel" pode retornar conceitos como "ferramentas de produtividade" e "números", mesmo que esses termos não apareçam no texto. Essa é uma poderosa ferramenta para criar conexões que podem não parecer óbvias à primeira vista.

Sentimento

O NLP classifica o nível de positividade ou negatividade no texto. Isso é útil, por exemplo, para medir o sentimento relacionado a um produto ou serviço. Ou, em um contexto de atendimento ao cliente, esta funcionalidade é útil ao determinar se um chat deve ser roteado para um humano (ao detectar negatividade).

Emoções

Esta é uma análise de sentimento em uma granularidade mais fina. Neste caso, o NLP classifica não apenas como "positiva" e "negativa", mas também "raiva", "tristeza" e "alegria".

Palavras-chave

O NLP extrai palavras-chave e frases para usar como base para indexação, pesquisa e classificação.

Categorias

O NLP cria uma taxonomia hierárquica para o teor dos dados e coloca essa taxonomia em uma categoria de alto nível (classificação de texto). Isso é útil para aplicações como recomendar conteúdo relevante, gerar anúncios, organizar emails e determinar o objetivo de um usuário.

Antes, você poderia tentar simular recursos no estilo NLP por meio de abordagens baseadas em regras, como expressões comuns ou

árvores de decisão, em um grande esforço para entender a intenção das perguntas formuladas por um ser humano. Você poderia ter usado modelos de aprendizagem de máquina personalizados, que exigiam acesso a conhecimento especializado, grandes conjuntos de dados e ferramentas complexas, o que limitou a implementação a apenas grandes organizações com os recursos para investir.

Considere onde estamos hoje. As APIs fáceis de usar na nuvem fornecem recursos de NLP que capacitam o uso generalizado de IA conversacional. Desde o surgimento de ferramentas Open Source até a chegada das APIs de nuvem, os recursos de NLP que antes se restringiam ao meio acadêmico e à comunidade de pesquisa agora estão disponíveis a um público-alvo mais amplo em várias indústrias.

Compreensão de Linguagem (antes denominada LUIS)

Compreensão de Linguagem, um serviço desenvolvido pela Microsoft, permite que desenvolvedores criem aplicações que podem usar a entrada do usuário em linguagem natural e extrair informações estruturadas, incluindo significado e intenção. A compreensão de Linguagem é um serviço baseado em aprendizado de máquina para criar experiências de linguagem natural, permitindo criar rapidamente modelos personalizados prontos para a empresa que melhorem continuamente.

Com a Compreensão de Linguagem, você pode usar um modelo predefinido (por exemplo, clima, calendário), personalizar um existente ou criar seu próprio do zero. Um modelo começa com uma lista de intenções gerais do usuário que representa as tarefas ou ações que o usuário deseja executar, como "reservar um voo", "agendar reuniões" ou "falar com o suporte técnico". Depois de identificar a intenção, você fornece exemplos de frases, denominadas expressões, para a intenção. Depois, você rotula as expressões com os detalhes específicos que deseja que a Compreensão de Linguagem extraia da expressão. Os dados extraídos da expressão são uma entidade.

Uma entidade representa informações detalhadas que são relevantes na conversa. Ao reconhecer e rotular as entidades que são mencionadas na entrada do usuário, a Compreensão de Linguagem ajuda você a escolher a ação específica a ser adotada para responder à requisição do usuário. Você pode definir suas próprias entidades, como terminologia específica de domínio, ou extrair entidades comuns predefinidas, como datas e horários, nomes próprios, medições e números. Com os **domínios predefinidos**, você tem um conjunto de entidades e expressões por categorias comuns, como calendário, entretenimento, comunicação e automação residencial.

A Compreensão de Linguagem também permite que os desenvolvedores melhorem continuamente a aplicação por meio de aprendizado ativo. A Compreensão de Linguagem armazena consultas de usuários e seleciona expressões incertas. Você pode revisar as expressões, selecionar a intenção e marcar entidades para expressões reais. Isso retreina o modelo de linguagem com mais dados.

O serviço se integra a outras ferramentas de IA na nuvem para potencializar o processamento e a compreensão de linguagem natural em aplicações, bots e dispositivos de Internet das Coisas (IoT). Por meio do Bot Framework, a Microsoft incorpora a Compreensão de Linguagem e outros serviços cognitivos para o desenvolvimento de bots.

Ecosistema do Bot Framework

O Microsoft Bot Framework (Figura 1) tem um ecossistema de ferramentas e serviços que fornece uma experiência abrangente para a criação de aplicações de IA conversacional.

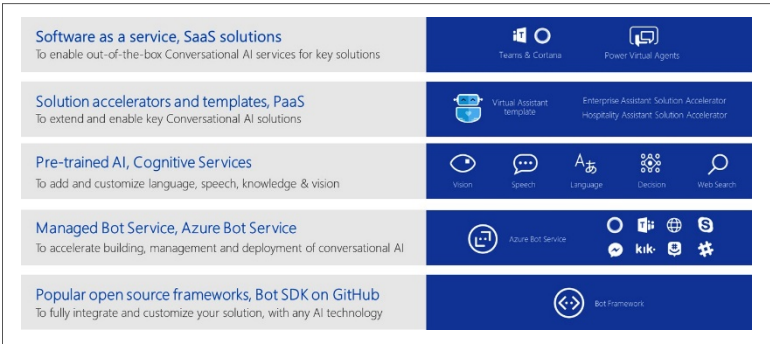


Figura 1. Ecosistema do Bot Framework

Com o SDK do Bot Framework, os desenvolvedores podem facilmente modelar e criar conversas sofisticadas usando as linguagens de programação favoritas. Os desenvolvedores podem criar aplicações de IA conversacionais que conversam de forma livre ou também podem ter interações mais guiadas em que a aplicação fornece opções de usuário ou possíveis ações. A conversa pode usar texto simples ou cartões avançados mais complexos contendo texto, imagens e botões de ação. Os desenvolvedores podem adicionar interações e perguntas e respostas de linguagem natural que permitem aos usuários interagir com bots de forma natural.

O Serviço de Bot do Azure permite hospedar aplicações de IA de conversação inteligentes e de nível empresarial com total propriedade e controle

dos dados. Os desenvolvedores podem registrar e conectar bots a usuários no Microsoft Teams e Web Chat, no Facebook Messenger e muito mais.

Para adicionar mais inteligência a um aplicação de IA conversacional, você pode adicionar e personalizar modelos de API pré-treinados e Serviços Cognitivos, incluindo recursos de linguagem, fala, conhecimento e visão.

O Bot Framework também fornece um conjunto de aceleradores de solução e modelos para ajudar a criar experiências de conversação sofisticadas. O acelerador de soluções de assistente virtual reúne todos os componentes de suporte e simplifica muito a criação de um novo projeto, incluindo intenções conversacionais básicas, integração de envio, QnA Maker, Application Insights e uma implantação automatizada.

A oferta **Power Virtual Agents** se baseia na plataforma do Bot Framework, fornecendo uma interface gráfica sem código para criar experiências conversacionais.

Casos de uso de IA conversacional

Os clientes familiarizados com assistentes virtuais e aplicações de mensagens estão se envolvendo cada vez mais com interfaces conversacionais, o que pode apresentar uma experiência mais natural em que seres humanos expressam suas necessidades por meio da linguagem natural e completam tarefas rapidamente. Aplicações de IA conversacional estão se tornando um diferencial competitivo para muitas empresas. Muitas organizações disponibilizam estrategicamente bots nas mesmas plataformas de mensagens em que clientes dedicam tempo. Organizações do mundo inteiro estão transformando seus negócios com a IA conversacional, que pode promover interações mais eficientes e naturais com clientes e funcionários. Aqui estão alguns casos de uso comuns:

Suporte ao cliente

As organizações estão usando a IA conversacional para transformar relacionamentos com clientes, fornecendo interações fáceis e naturais em vários canais e plataformas, como dispositivos domésticos, aplicações móveis, canais sociais como o Facebook Messenger e sites. As experiências conversacionais não só permitem que as organizações alcancem os clientes onde quer que estejam, mas também personalizem e melhorem continuamente as interações.

As seguradoras, por exemplo, facilitam tarefas para os clientes, como obter respostas rápidas para as perguntas frequentes, enviar reivindicações ou até mesmo gerar uma cotação para um plano de seguro.

As empresas de varejo também estão permitindo que usuários rastreiem rapidamente pacotes e obtenham atualizações de status do pedido, permitindo que um cliente seja transferido para o chat com um agente humano. As empresas de telecomunicações estão usando assistentes virtuais com recursos de IA para saber mais sobre os clientes para fornecer interações sofisticadas personalizadas, aumentar a receita e aumentar a produtividade das equipes de suporte ao cliente.

Assistente corporativo

As organizações estão usando a IA conversacional para melhorar o envolvimento dos funcionários, conectando pessoas, tarefas, informações e serviços de forma mais eficaz com interfaces mais naturais e intuitivas. Integrando assistentes de funcionários com interfaces de voz e texto em dispositivos corporativos e telas de conversa existentes (por exemplo, Microsoft Teams, Slack e Web Chat), as organizações aceleram o processo de gerenciamento de calendários, encontram salas de reuniões disponíveis, identificam pessoas com habilidades específicas ou contatam o RH. A integração com Dynamics, PowerApps, ServiceNow e outros provedores de TI simplifica os acessos dos funcionários e permite que encontrem facilmente os dados e executem as tarefas que estão procurando. A integração em pesquisas adiciona o poder de fornecer dados corporativos de forma natural para os usuários também.

Otimização de call center

A integração de uma experiência conversacional em um sistema de comunicação via telefone de call center pode reduzir os tempos de chamada com agentes humanos, esclarecendo informações com antecedência ou resolvendo solicitações simples sem precisar de um agente humano. Somado a isso, a solução substitui as soluções clássicas de resposta de voz interativa (IVR) com uma experiência conversacional moderna e permite uma experiência de usuário consistente durante a chamada, ou até a transferência para um agente humano.

A análise pós-chamada avalia a qualidade da chamada e os comentários dos clientes, com insights disponíveis para melhorar o fluxo de chamadas e otimizar a experiência do usuário, aumentar a resolução de primeiro contato e atender a outros indicadores-chave de performance (KPIs).

O mesmo assistente pode ser exposto por meio de canais adicionais somente de texto, permitindo que os usuários finais interajam por meio do canal de sua preferência e aumentando a recompensa do investimento ao garantir a participação de todos os usuários, estejam eles usando SMS ou canais mais avançados.

Assistente de voz no carro

Os assistentes habilitados para voz integrados a carros fornecem aos motoristas e passageiros a capacidade de realizar operações de carro tradicionais (por exemplo, navegação, rádio), junto com cenários focados em produtividade, como adiar reuniões quando você está atrasado, adicionar itens à lista de tarefas e experiências proativas em que o carro pode sugerir tarefas a serem concluídas com base em eventos, como dar partida no motor, ir para casa ou ativar o piloto automático. Outros casos de uso incluem agendamento de serviço para um veículo com base nas preferências de um usuário para o provedor de serviços, a localização do veículo, a disponibilidade de agendamento do provedor, a gravidade do problema, a preferência de empréstimo, os horários pessoais e de trabalho e muitas outras variáveis. Esse é o poder de trazer os dados de um fornecedor automotivo para a imagem e ilustra a experiência totalmente integrada possível por meio da solução de assistente virtual.

Assistente de hospitalidade

Um assistente virtual integrado a um dispositivo de quarto de hotel pode fornecer uma ampla variedade de cenários focados em hospitalidade: estender uma estadia, solicitar checkout tardio, serviço de quarto, serviços de concierge e encontrar restaurantes e atrações locais. A aplicação pode ser vinculada a uma conta de produtividade, abrindo experiências mais sofisticadas, como chamadas de alarme, avisos meteorológicos e padrões de aprendizagem em estadias.

Estes são alguns exemplos dos tipos de aplicações de IA conversacional a serem criadas que abordaremos neste livro. Vejamos o fluxo de trabalho típico para desenvolver uma aplicação de IA conversacional.

Fluxo de trabalho de desenvolvimento de aplicações de IA conversacional

O fluxo de trabalho típico para o desenvolvimento de uma aplicação de IA conversacional assemelha-se a outros tipos de projetos: as principais fases são *projetar*, *criar*, *testar*, *implantar*, *conectar* e *avaliar* (Figura 2).³

³ Essas fases são descritas mais adiante em [Documentos de linha do Azure](#).



Figura 2. O fluxo de trabalho típico para desenvolver uma aplicação de IA conversacional

Vejam os detalhes de cada fase deste fluxo de trabalho.

Projeto

Desenvolver um bot, como desenvolver sites e aplicações, deve começar com um projeto de uma ótima experiência. Quando seres humanos interagem com bots, esperamos que compreendam o que dizemos, que recebamos uma resposta apropriada e que o serviço obtido seja proveitoso. Esperamos que, se sairmos no meio da conversa, o bot se lembre onde paramos.

Seu bot representa sua marca, produtos e serviços para clientes e funcionários; portanto, é imperativo começar com uma abordagem orientada a design para garantir que o objetivo do bot atenda à necessidade explícita ou latente do serviço de TI humano. Para projetar uma experiência proveitosa, utilize as práticas recomendadas de pesquisa de usuários direcionados, definindo perfis de bots, cenários de bot de roteiro, projetando fluxo de conversas e definindo um plano de avaliação, *sem* especificar detalhes de desenvolvimento técnico.

Para cada uma dessas atividades de design, aqui estão as perguntas-chave para responder:

Pesquisar usuários direcionados

Quem são seus usuários? Quais são os objetivos, as necessidades e as expectativas deles? Qual é o contexto da interação com o bot? Como é o ambiente deles? Como seu bot ajudará eles? Quais serviços o bot deve fornecer a eles?

Definir perfis de bots

Qual é a aparência do seu bot (por exemplo, um avatar)? Como deve ser nomeado? O bot cumpre os valores da sua organização? Qual é a personalidade do seu bot? Seu bot tem um gênero? Ele pode responder perguntas fora do tópico? Que tom de voz seu bot deve usar? Como seu bot lidaria com diferentes situações? Como o bot deve responder (com gerenciamento proativo, reativo ou de exceção)?

Cenários de bot de roteiro

Qual é a jornada do usuário para os usuários direcionados do bot?
O que o bot deve fazer e não fazer? Quais são os objetivos e as prioridades dos casos de uso do seu bot?

Projetar fluxos de conversas

Quais fluxos de conversa você pode esperar para seus principais casos de uso? P e R simples, notificações por push, instruções passo a passo ou interações mais complexas?

Definir um plano de avaliação

Como medir o sucesso? Quais medidas você deseja usar para melhorar o serviço e onde inserir a instrumentação?

Antes de escrever código, revise as **diretrizes de design de bot** da documentação do Bot Framework da Microsoft para obter práticas recomendadas.

O Bot Framework fornece um conjunto de ferramentas para a fase de design, incluindo:

- Arquivos **.chat** para criar um modelo de conversas entre o usuário e o bot para cenários específicos
- O comando **bf chatdown** para converter arquivos **.chat** em transcrições avançadas
- Emulador do Bot Framework, que **abre um arquivo .transcript** para exibir uma renderização realista das conversas (**Figura 3**)

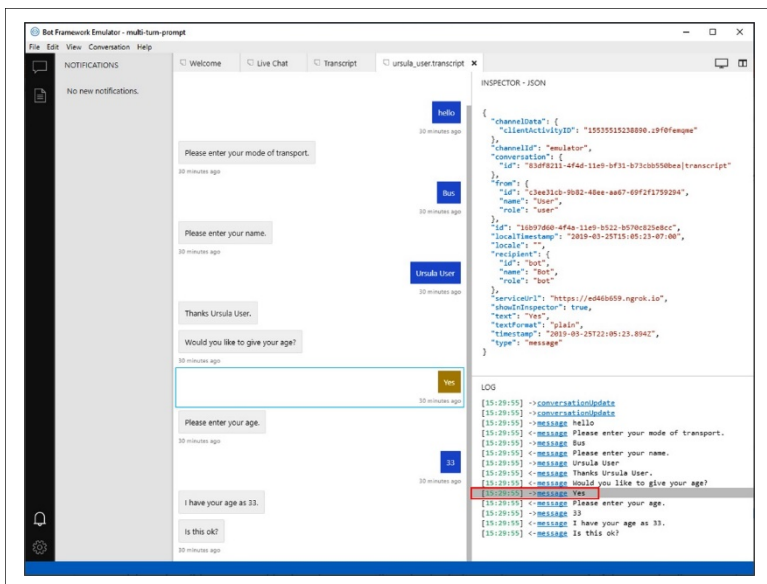


Figure 3. Exibição de um arquivo .transcript no emulador do Bot Framework

Criar

Um bot é um serviço Web REST (Representational State Transfer) que se comunica com o usuário enviando e recebendo mensagens e eventos de interfaces conversacionais, como salas de chat ou widgets de Web Chat. Com o Serviço de Bot do Azure e o Bot Framework da Microsoft, você pode criar bots em vários ambientes e linguagens de desenvolvimento. Você pode iniciar o desenvolvimento de bot no [portal do Azure](#) ou usar um dos modelos de SDK do Bot Framework para desenvolvimento local. Os modelos são compatíveis com as linguagens **C#**, **JavaScript** e **Python** com suporte a Java na visualização antecipada no momento da gravação.

Depois de criar o bot básico, estenda a funcionalidade da forma exigida pelo seu design. Você pode adicionar recursos de NLP usando a **Compreensão de Linguagem**, adicionar uma base de conhecimento para responder perguntas comuns usando o **QnA Maker**, adicionar recursos para gerenciar fluxos de conversas complexos e vários domínios de conhecimento usando a ferramenta **Dispatch** e adicionar gráficos ou menus usando **Cartões Adaptáveis**. Além disso, a Microsoft fornece **ferramentas de linha de comando** para ajudá-lo a criar, gerenciar e testar esses ativos de bot como parte de um processo de DevOps.

Você pode acessar diversos **exemplos** que demonstram os recursos conversacionais disponíveis por meio do SDK, incluindo recursos de diálogo básico,

como diálogos multiturnos por meio de recursos avançados, como mensagens proativas e autenticação.

Além disso, a Microsoft fornece um **modelo de Assistente Virtual** mais avançado, que é recomendado como um ponto de partida para criar uma experiência conversacional mais sofisticada. Ele reúne muitas práticas recomendadas para criar experiências conversacionais e automatiza a integração de componentes que foram considerados altamente benéficos pelos desenvolvedores do Bot Framework.

Por exemplo, uma experiência conversacional criada no modelo de assistência virtual permite que desenvolvedores lidem com várias linguagens, modelos de NLP para intenções conversacionais de base, personalidades personalizadas para responder perguntas mais gerais, geração de linguagem integrada para respostas mais naturais, uma experiência de introdução para novos usuários, alternância de contexto e suporte de habilidades.

Na próxima seção deste livro, utilizaremos o modelo de assistente virtual para criar uma aplicação de IA conversacional.

Testar

Para testar sua aplicação de IA conversacional, a Microsoft fornece o **Emulador do Bot Framework** permitindo que os desenvolvedores testem conversas com rapidez e facilidade. Você também pode **gravar testes de unidade** usando o SDK do Bot Framework, que pode se concentrar no teste de funcionalidade de diálogos específicos. Uma vez configurado por meio do portal do Azure, seu bot pode ser alcançado por meio de uma interface de web chat, permitindo testes mais amplos pelos usuários finais no início do processo de desenvolvimento.

Publicar

Quando você estiver pronto para disponibilizar o bot na Web, publique-o no **Azure** ou em seu próprio serviço Web ou data center — onde uma aplicação Web normal possa ser hospedada.

Conectar

O Serviço de Bot do Azure faz a maior parte do trabalho necessário para conectar bots a diversos canais e dispositivos. Configurado por meio do portal do Azure, você pode conectar bots ao Facebook Messenger, Slack, Microsoft Teams, Cortana, email, Telegram, Twilio, LINE e outros canais. Você também pode usar widgets de Web Chat para incorporar bots em sites ou aplicações móveis.

Você pode usar o canal Direct Line para conectar o bot à sua própria aplicação de cliente, ou o canal do Direct Line Speech que permite interfaces de fala de baixa latência com aplicações cliente usando o SDK do Microsoft Speech. Assim, você pode incorporar experiências de texto e fala em aplicações de área de trabalho, aplicações móveis e dispositivos, como carros, alto-falantes e despertadores.

O Bot Framework e os membros da comunidade Open Source também fornecem **adaptadores baseados em código** para conectar bots a outros canais, como Google Assistant, Amazon Alexa, Webex Teams, WebSockets e webhooks.

Avaliação

Gravações de conversas entre bots e usuários fornecem insights de negócios valiosos para ajudá-lo a avaliar a performance do seu bot. Nesta fase, as práticas recomendadas incluem avaliar as métricas de sucesso que você definiu durante a fase de design, revisar logs de instrumentação, coletar comentários de usuários, refinar e iterar. O Bot Framework fornece exemplos de consultas do Application Insights e um painel do Power BI para ajudá-lo a compreender a amplitude das conversas do bot com usuários e obter insights importantes sobre a integridade e o comportamento do bot.

Principais recursos de assistentes virtuais

Com base no **SDK do Bot Framework**, a solução Open Source de **Assistente Virtual** da Microsoft (disponível em `c#` e TypeScript) é um modelo de projeto que engloba as práticas recomendadas para desenvolver um bot na plataforma Microsoft Azure.

As organizações estão percebendo uma necessidade maior de oferecer experiências avançadas de assistente de conversacional personalizadas de acordo com a marca e os usuários e disponibilizadas em uma ampla variedade de aplicações e dispositivos. Com o Assistente Virtual, você controla o nome, a voz e a personalidade para atender às suas necessidades. O Bot Framework fornece soluções que simplificam a criação de um Assistente Virtual, permitindo que você comece a usar e estenda o bot com uma gama de ferramentas de desenvolvimento completas.

O Assistente Virtual reúne Bot Framework, Serviço de Bot do Azure e Compreensão de Linguagem na plataforma de IA do Azure para simplificar a criação de seu próprio Assistente Virtual (consulte a **Figura 1**). O Bot Framework e o Serviço de Bot do Azure fornecem recursos conversacionais básicos

para o Assistente Virtual, incluindo gerenciamento de diálogos, solicitações de linguagem natural, alternância de contexto, memória e geração de linguagem. O Assistente Virtual fornece recursos adicionais e conjuntos predefinidos de interações específicas de domínio chamadas *Habilidades* para ajudar as organizações a criar suas próprias experiências de assistente em diversas linguagens, como Calendar e ToDo.

O Assistente Virtual pretende facilitar o trabalho do desenvolvedor e torná-lo mais produtivo. Nesta seção, vamos examinar alguns recursos principais do Assistente Virtual.

Controle completo para desenvolvedores

Com o Assistente Virtual, todos os aspectos da experiência do usuário são de propriedade e controlados por você. Isso inclui marca, nome, voz, personalidade, respostas e avatar. A Microsoft fornece **cinco personalidades de chat** baseadas no QnA Maker do Azure Cognitive Service, permitindo adaptar a personalidade do bot. O código-fonte para o Assistente Virtual e as habilidades de suporte são oferecidos como exemplos para você personalizar. O Assistente Virtual será implantado na assinatura do Azure. Portanto, todos os dados gerados pelo assistente (perguntas, comportamento do usuário etc.) estão totalmente contidos na assinatura do Azure. Consulte **detalhes de conformidade e privacidade dos serviços cognitivos** e a **seção do Azure da Central de Confiabilidade** para obter mais informações.

Habilidades predefinidas e reutilizáveis

Cenários comuns do Assistente Virtual são oferecidos como habilidades conversacionais reutilizáveis e incluem tarefas como encontrar pontos de interesse próximos, verificar um item em uma lista de tarefas e responder um email. Habilidades, entregues no formulário de código-fonte, são totalmente personalizáveis e consistem em modelos de linguagem para várias linguagens naturais, diálogos e código de integração. Habilidades adicionais podem ser criadas e disponibilizadas por meio de seu próprio assistente ou de um ecossistema de habilidades mais amplo. Isso permite selecionar os recursos que fazem sentido no seu cenário e que funcionam em todas as indústrias. Como o Assistente Virtual aproveita o Serviço de Bot do Azure, você pode fornecer a usuários acesso ao seu assistente por meio de um dos canais e adaptadores compatíveis, permitindo alcançar os usuários finais onde quer que estejam e usando experiências de UI/UX com as quais eles já estejam familiarizados e confortáveis.

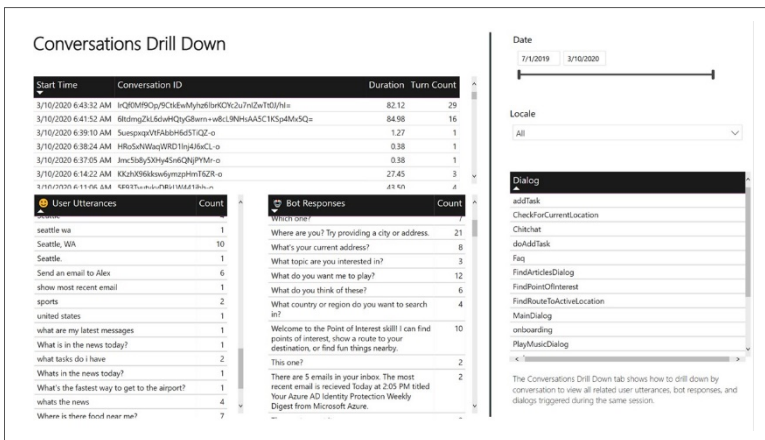
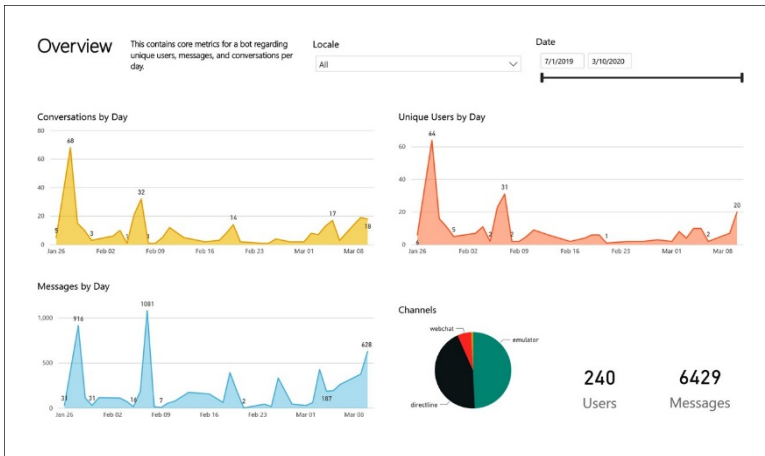
Integração flexível e conscientização contextual

A arquitetura do Assistente Virtual é flexível e pode ser integrada a recursos de fala ou NLP existentes, sistemas de back-end, APIs e dispositivos para habilitar a personalização e adaptar as respostas conforme necessário para o contexto dos usuários. O assistente também pode estar ciente do dispositivo ou canal por meio do qual o usuário está se comunicando, que pode ser usado para otimizar a experiência (por exemplo, voz x texto) e permitir que o usuário alterne entre canais diferentes, se necessário.

Insights de negócios capturados e distribuídos

O Bot Framework fornece um conjunto sofisticado de métricas prontos para uso, permitindo que qualquer experiência conversacional colete dados nos níveis desejados, para uma melhor compreensão do que os clientes estão fazendo e onde estão prováveis oportunidades de melhorar o serviço.

O Azure Application insights captura análises completas em toda a arquitetura. A Microsoft também forneceu painéis de Power BI de exemplo que fornecem gráficos fáceis de ler, mostrando conversas populares, duração da conversa, usuários únicos e outras estatísticas importantes (Figuras 4 e 5). Você pode estender ainda mais o pipeline de insights com seu próprio aprendizado de máquina para fornecer experiências de IA personalizadas aos usuários. Como todos os dados que fazem parte do seu bot, as métricas permanecem sob controle, ajudando-o a cumprir regulamentos de privacidade, como o GDPR (Regulamento Geral sobre a Proteção de Dados).



O Assistente Virtual também se integra ao serviço de fala, parte da família de Serviços Cognitivos, junto com o NLP e o gerenciamento de diálogo, gerando a conscientização contextual de conversas. O streaming em tempo real do áudio do usuário permite que o NLP e os diálogos comecem assim que o usuário termina de falar, oferecendo uma experiência mais natural e de baixa latência. A capacidade de voz neural personalizada no serviço de fala permite que os clientes desenvolvam vozes personalizadas altamente realistas para interfaces conversacionais naturais, começando com apenas 30 minutos de áudio.

Cartões adaptáveis

Os **Cartões Adaptáveis** fornecem recursos gráficos, como cartões, imagens e botões no assistente. Os cartões são peças independentes de plataforma de Interface do usuário, criado em JSON, que pode ser trocado entre aplicações e serviços. Quando entregue a uma aplicação específica, o JSON é transformado na interface do usuário nativa que se adapta automaticamente ao seu entorno. Ele permite que você crie e integre a interface do usuário leve para todas as principais plataformas e estruturas.

Se a tela de conversa tiver uma tela, esses cartões poderão ser renderizados em diversos dispositivos e plataformas, fornecendo um UX consistente com o serviço ou contexto em que o cartão é incorporado. Dispositivos sem telas podem utilizar as respostas amigáveis à fala, oferecidas junto com os cartões adaptáveis ou qualquer combinação de mecanismos de entrega apropriados ao contexto.

O Assistente Virtual e as habilidades relacionadas operam de forma abrangente com cartões adaptáveis, e o design e a identidade visual podem ser totalmente personalizados para atender ao seu cenário. A **Figura 6** mostra alguns exemplos.

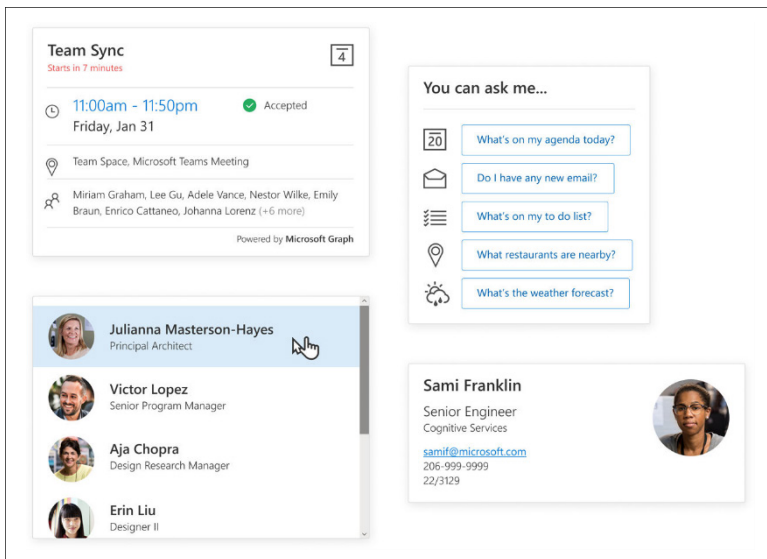


Figura 6. Exemplos de Cartão Adaptável

Recursos para a empresa

Uma típica experiência conversacional baseada no Bot Framework aproveita diversos recursos do Azure, como Serviço de Bot do Azure, Compreensão de Linguagem e Serviços Cognitivos de Fala, junto com um amplo conjunto de componentes de suporte do Azure. Isso significa que você se beneficia da **infraestrutura global do Azure** que inclui ISO 27018, HIPAA, PCI DSS e certificação SOC 1, 2 e 3. Além disso, a Compreensão de Linguagem **é compatível com muitas linguagens**. O **serviço Translator** fornece recursos de tradução automática para estender ainda mais o Assistente Virtual.

Agora que sabemos a capacidade do assistente virtual, vamos criar um. Nas próximas seções, vamos orientá-lo no processo de criação e personalização do Assistente Virtual, adicionando Inteligência ao assistente com habilidades, conectando o assistente a clientes e canais, e possibilitando análises do assistente. No percurso, forneceremos links para tutoriais online, que podem ser feitos com C# ou TypeScript.

Desenvolver seu assistente virtual

Nesta seção, vamos orientá-lo sobre como criar um assistente virtual altamente funcional em sua própria assinatura do Azure.⁴

O modelo do assistente virtual

O modelo de assistente virtual,⁵ como observamos anteriormente, é uma ferramenta Open Source que fornece um ponto de partida predefinido para criar um assistente personalizado. Usando o modelo, você gerará um projeto de assistente, que segue a estrutura recomendada para um projeto de bot. Você é livre para reestruturar isso conforme necessário, mas lembre-se de que os scripts de implantação fornecidos esperam que alguns arquivos estejam em locais consistentes.

Criar uma experiência conversacional usando o modelo de assistente virtual exige os seguintes recursos:

- Registro de Bot do Azure (para definir a configuração do ponto de extremidade do bot e do canal)
- Aplicação Web do Azure (para hospedar sua aplicação de bot)
- Conta de Armazenamento do Azure (para armazenar transcrições)
- Azure Application Insights (telemetria)
- Azure Cosmos DB (estado conversacional e estado do usuário — isso pode ser trocado para Armazenamento do Azure para fins de desenvolvimento)
- Compreensão de Linguagem
- QnA Maker (incluindo Pesquisa Cognitiva do Azure e uma aplicação Web do Azure)

Para permitir que você comece rapidamente e preveja esses recursos, a Microsoft fornece um modelo do Azure Resource Manager (ARM) e um conjunto de scripts do PowerShell (suporte entre plataformas), junto com modelos de compreensão de linguagem para domínios de base comum, bases de conhecimento do QnA Maker (perguntas frequentes de personalidade e exemplo) e as ferramentas do operador.

4 Além disso, você pode seguir a [documentação de início rápido do Bot Framework](#) para criar uma experiência mais simples que pode ser estendida para cenários adicionais.

5 A arquitetura e as funcionalidades do modelo são descritas na [documentação online para o modelo de Assistente Virtual](#).

Tutorial online: crie um assistente virtual

Siga o tutorial online (em **C#** ou **TypeScript**) para criar sua primeira aplicação de assistente virtual que cumprimenta um novo usuário e trata intenções conversacionais básicas. Durante este tutorial, você irá:

1. Verificar se tem uma assinatura do Azure (ou **obter uma conta gratuita do Azure** se ainda não tiver uma).
2. Fazer o download e instalar pré-requisitos de desenvolvimento do Bot Framework e o modelo de Assistente Virtual.
3. Criar seu projeto do Virtual Studio usando o modelo de assistente virtual.
4. Provisionar o assistente usando o modelo do ARM fornecido e um script do PowerShell.
5. Executar e testar seu assistente.

Tutorial online: personalize seu assistente

Agora que tem um assistente, você pode opcionalmente personalizá-lo para personalizar a experiência da sua marca e dos usuários. Os diálogos poderão ser adicionados diretamente ao assistente ou por meio de habilidades se você quiser criar uma solução mais complexa. Isso é abordado abaixo. Siga este tutorial online (em **C#** ou **TypeScript**) para saber como executar as seguintes tarefas:

- Edite a saudação personalizando o cartão adaptável.
- Edite respostas personalizando arquivos de geração de linguagem (.lg).
- Edite modelos cognitivos, por exemplo, atualizando bases de conhecimento (perguntas frequentes e/ou bate-papo), adicionando uma base de conhecimento adicional e atualizando arquivos locais LU para Compreensão de Linguagem e QnA Maker.

Conectar assistentes a clientes e canais

Clientes e canais são as maneiras como os usuários podem interagir com uma aplicação de IA conversacional. A plataforma de IA do Azure faz a maior parte do trabalho necessário para conectar bots a diversos canais e dispositivos.

Configurado por meio do portal do Azure, você pode conectar bots ao Facebook Messenger, Slack, Microsoft Teams, Cortana, email, Telegram, Twilio, LINE e outros canais. Você pode usar widgets de Web Chat para incorporar bots em sites ou aplicações móveis.⁶

Tutorial online: habilite seu assistente para fala

O **Direct Line Speech** é uma solução robusta e completa para a criação de um assistente de voz flexível e extensível que é otimizado para a interação de fala com bots. O Direct Line Speech oferece um alto nível de personalização e sofisticação para assistentes de voz.

Siga o **tutorial online** para executar as tarefas a seguir para conectar seu assistente ao canal do Direct Line Speech e criar uma aplicação simples integrada com o SDK de fala para demonstrar interações de fala.

1. Crie um recurso Serviço de Fala.
2. Adicione o canal do Direct Line Speech.
3. Use o Emulador do Bot Framework ou a aplicação cliente de amostra de fala e conecte-se ao seu assistente.
4. Mude a voz.

Tutorial online: evidencie o assistente virtual para Microsoft Teams

Se você estiver analisando a criação de um assistente corporativo, poderá seguir o **tutorial online** para conectar o assistente ao Microsoft Teams e criar o manifesto de aplicação necessário para instalar o assistente no Teams.

1. Adicione o canal do Microsoft Teams.
2. Instale o Teams App Studio.
3. Crie o manifesto de aplicação para o Teams.
4. Teste no Teams.
5. Adicione comandos.

⁶ Você pode encontrar detalhes sobre como se conectar a canais na **documentação de canais do Bot Framework**, com o sumário à esquerda com links para instruções adicionais específicas de canal. Além disso, você tem a opção de conectar o assistente ao Amazon Alexa, ao Google Home e a outros por meio da integração feita pela **comunidade Open Source do Bot Builder**.

Opcional: adicionar inteligência ao assistente com habilidades

Uma Habilidade do Bot Framework fornece um modelo de componente conversacional que permite a desenvolvedores separar a experiência de assistente em um conjunto de blocos de construção conversacionais, que podem ser desenvolvidos sem depender uns dos outros e reunidos em uma experiência unificada. Esse é um padrão comum para experiências conversacionais maiores, em que há um "bot pai" com o qual os usuários interagem, transmitindo a várias habilidades de "filho" para lidar com certas tarefas.

Pense no amplo conjunto de recursos comuns e diálogos que os desenvolvedores tradicionalmente criavam. Os cenários de produtividade são um bom exemplo, em que cada organização precisa criar seus próprios modelos de linguagem, diálogos, integração de API e respostas. O trabalho é ainda mais complicado pela necessidade de dar suporte a vários idiomas, resultando em uma grande quantidade de trabalho necessária para qualquer organização que cria sua própria experiência de assistente.

O Bot Framework fornece diversas **habilidades de conversação Open Source** em várias linguagens, incluindo Calendar, Email, To Do, e Point of Interest, para reduzir esse esforço. A estrutura também oferece várias habilidades de experimentação, incluindo o gerenciamento de serviços de telefone, notícias, clima, música e TI. Essas habilidades de conversação são bots e incorporam modelos de linguagem, diálogos e código de integração. Elas são criadas da mesma forma que qualquer bot, mas podem ser incorporadas por meio da fácil configuração em uma experiência conversacional existente para estender recursos. Todos os aspectos de cada habilidade são totalmente personalizáveis por desenvolvedores, e o **código-fonte completo é fornecido no GitHub** junto com o Assistente Virtual.

As organizações também podem criar habilidades para uso privado ou para compartilhar com outras organizações para compor suas próprias experiências. Por exemplo, uma aplicação conversacional desenvolvida por um serviço de entrega de refeições para seus próprios canais (aplicações móveis, sites e telas de conversação) também pode ser exposta como uma habilidade para dispositivos IoT domésticos e carros para integrar conforme apropriado. Isso destaca uma capacidade essencial do Bot Framework e do Serviço de Bot do Azure: eles permitem escrever uma habilidade uma vez e, depois, fornecê-la por meio de canais diferentes (incluindo Alexa e Assistente do Google) com uma única base de código para reduzir a duplicação em diferentes ecossistemas.

Tutorial online: conecte-se a uma habilidade de exemplo

Conforme necessário, você pode adicionar uma das **habilidades** de exemplo oferecidas no repositório de soluções do Bot Framework ao seu assistente. As habilidades estão disponíveis apenas no C# neste momento, mas podem ser adicionadas a um assistente baseado em JavaScript ou Python.

Siga o **tutorial online** para executar as seguintes tarefas:

1. Implante um projeto de habilidade de exemplo.
2. Adicione a habilidade de exemplo ao seu assistente e republique o assistente para o Azure.
3. Teste a habilidade de exemplo.

Tutorial online: crie uma habilidade personalizada

Se desejar criar sua própria habilidade personalizada para estender o assistente, siga o **tutorial online**, durante o qual você executará as seguintes tarefas:

1. Crie seu projeto de habilidade usando o modelo de habilidade no Visual Studio.
2. Provisione recursos do Azure usando o modelo do ARM fornecido e um script do PowerShell.
3. Execute sua habilidade.
4. Adicione sua habilidade a um assistente.
5. invoque sua habilidade.

Habilitar ferramentas analíticas para o assistente virtual

Os desenvolvedores podem obter insights importantes sobre a integridade e o comportamento do assistente com as soluções de análise do Bot Framework, que incluem consultas de aplicações de exemplo e um painel (aproveitando o Microsoft Power BI) para entender a amplitude das conversas do assistente com os usuários. Os dados subjacentes capturados podem ser usados por uma ampla gama de ferramentas de análise de dados, conforme necessário.

Tutorial online: Exiba análises com o Power BI

Siga o **tutorial online** aqui para conectar o assistente aos painéis fornecidos no modelo de análise do Assistente Virtual do Power BI:

1. Configure o log de telemetria do assistente com a aplicação Application Insights.
2. Abra o modelo de análise do Assistente Virtual (um modelo do Power BI) e conecte-o ao log de telemetria.

O modelo do Power BI fornece um painel abrangente para você obter insights valiosos sobre a performance do assistente e as áreas que precisam de aperfeiçoamento. Ele fornece insights, incluindo:

Uso geral

Entenda métricas principais, como usuários, mensagens e conversas exclusivas por dia e canais (**Figura 4**)

Diálogos

Revise a popularidade de todos os diálogos, bem como os resultados (abandonado, cancelado, concluído ou iniciado)

Compreensão de Linguagem

Obtenha insights sobre as intenções de Compreensão de Linguagem, que são úteis para fins de monitoramento do que os usuários falam

Conversas

Exiba dados sobre conversas por usuário por dia e duração média, com a capacidade de detalhar por conversa para exibir todas as expressões relacionadas do usuário, respostas de bot e diálogos acionados durante a mesma sessão (**Figura 5**)

Transcrições

Mostre interações, sessões e transcrições de conversas entre o assistente e usuários

Insights do QnA Maker

Revise insights sobre consultas de usuários correspondentes com o QnA Maker, que é útil para identificar as lacunas da base de dados de conhecimento do usuário

Insights de comentários do usuário

Revise comentários explícitos do usuário (positivo, negativo) e expressões correspondentes do usuário e do bot se o bot tiver ativado o middleware de comentários na telemetria

Agora que você aprendeu a ver uma amostra de análise do Assistente Virtual, também pode:

- **Adicionar a telemetria ao seu bot:** saiba quais componentes de código específicos são necessários para a telemetria pronta para uso.
- **Analisar dados de telemetria do seu bot .**
- Trabalhar com **eventos gerados pela telemetria do Bot Framework Service.**

Roteiro e mais recursos

A Microsoft continua a investir na plataforma de IA do Azure para facilitar para organizações e desenvolvedores a criação de soluções conversacionais robustas e implantá-las onde quer que os clientes estejam.

Por exemplo, a Microsoft continua a melhorar o processo para criar experiências conversacionais por meio de lançamentos regulares dos SDKs e ferramentas do Bot Framework. À medida que as práticas recomendadas evoluem, elas são empacotadas no modelo de assistente virtual e nas habilidades de exemplo.

A Microsoft está empenhada em facilitar a criação de bots com o **Bot Framework Composer**, que fornece uma interface visual Low Code para criar, editar, testar e refinar bots. A Microsoft também está conectando bots a mais usuários, adicionando mais canais e suporte à transferência humana e comunicações por telefone.

Ao mesmo tempo, a Microsoft permite conversas mais naturais, dinâmicas e sofisticadas com **Diálogos Adaptáveis**. Esses diálogos permitem uma interação mais natural, com o usuário conseguindo se mover sem problemas em todas as fases de um diálogo, mudando de ideia sobre uma resposta anterior ou fornecendo informações adicionais sobre perguntas posteriores automaticamente.

Além disso, a Microsoft está melhorando e aprimorando os recursos de Compreensão de Linguagem e fornecendo suporte para a compreensão de documentos.

Para obter recursos adicionais, confira:

- [Documentação de SDK do Bot Framework do GitHub](#)
- [Documentação do Serviço de Bot do Azure da Microsoft](#)
- [Documentação do Bot Framework da Microsoft](#)
- [Notícias do Bot Framework do GitHub](#)

Criar IA responsável

O Capgemini Research Institute, no [relatório de julho de 2019](#), identificou que quase nove em cada dez organizações tiveram consequências não intencionais resultantes do uso da IA. Os autores deste relatório identificaram suas principais preocupações, que incluíram:

- Dependência excessiva de decisões de máquina sem divulgação
- Coleta e processamento de dados pessoais em algoritmos de IA sem consentimento ou para fins diferentes do objetivo da coleta
- Recomendações tendenciosas e pouco claras, resultando em acesso e preços de produtos ou serviços discriminatórios
- Cidadãos que se opõem à vigilância em massa, à coleta e ao uso de dados pessoais, incluindo biometria
- Clientes exigindo análise e clareza por trás de uma decisão tomada por um algoritmo de IA

Como descrevemos neste livro, avanços decisivos em tecnologia de IA nos últimos cinco anos estão começando a transformar produtos e serviços, o que afeta a vida de todos. Embora muitas dessas mudanças sejam boas, elas também suscitam preocupações sobre consequências não intencionais resultantes de preconceitos, perda da privacidade, desinformação e automação.

Projetar a IA para ser confiável exige a criação de soluções que reflitam princípios éticos profundamente enraizados em valores importantes e atemporais. Na Microsoft, seguimos [princípios de imparcialidade, confiabilidade e segurança, privacidade e segurança, inclusão, transparência e responsabilidade para criar sistemas de IA responsáveis](#).

Como ocorre em qualquer tecnologia, a confiança depende, em última análise, da capacidade dos sistemas baseados em IA de serem operados de forma confiável, segura e consistente, não apenas em circunstâncias normais, mas também em condições inesperadas ou quando estão sob ataque.

— presidente da Microsoft Brad Smith, *The Future Computed*

É essencial criar aplicações de IA responsáveis, mas, em última análise, cabe aos desenvolvedores de IA e às organizações que implantam tecnologias de IA garantir um comportamento responsável. Para ajudar os desenvolvedores e as organizações a criar soluções responsáveis, a Microsoft propôs **18 diretrizes de design para a interação entre humanos e IA**. Para sistemas de IA conversacional, a Microsoft também lançou **um conjunto de 10 diretrizes** que abrangem impacto, transparência, inclusividade, confiabilidade, privacidade, segurança e detecção de preconceitos ou exclusão injusta.

Para adotar uma abordagem responsável com a IA, é essencial considerar as necessidades das pessoas que usarão a solução durante o design inicial. Por exemplo, um agente conversacional projetado para ajudar cidadãos a obter acesso a serviços de saúde pública deve considerar e atender a toda a gama de habilidades desses indivíduos. Projetar uma interface conversacional com suporte à entrada de texto e áudio ajuda aqueles com deficiências visuais, bem como indivíduos que operem um veículo.

Como usuários e desenvolvedores de soluções de IA, você deve definir e seguir critérios que reflitam princípios de IA. Por exemplo, com imparcialidade, é importante considerar os possíveis malefícios da alocação, representação ou associação que podem resultar de tendência injusta em dados e modelos. Uma aplicação de IA utilizada para selecionar candidatos a emprego pode associar o sexo a certos hobbies, como atividades esportivas, com base em preconceitos em dados de treinamento e modelos de incorporação de palavras, o que pode resultar em seleção injusta. É importante **avaliar e monitorar a imparcialidade de uma solução de IA** e mitigar a iniquidade observada usando ferramentas como o kit de ferramentas Fair Learn.

Também é importante entender o contexto em que a solução opera. Por exemplo, com bots que têm perfis humanizados, é fundamental que eles interajam com respeito e segurança com os usuários, com proteções incorporadas para lidar com uso indevido e abuso. Um bot que ajuda os consumidores a reservar o entretenimento é propenso a oferecer suporte a tons e termos não apropriados a um bot focado em serviços de saúde.

Para criar confiança, é essencial que as pessoas entendam o que uma solução pode fazer e quais informações essa solução coleta. Algumas perguntas críticas que os desenvolvedores devem fazer incluem:

- Os usuários devem estar cientes de que um bot é a base do diálogo em que estão envolvidos?
- Como o sistema lida com falhas, por exemplo, quando há um erro na transcrição de fala para texto?
- Quais dados ele coleta e como os usuários controlam essa coleta de dados?
- O sistema é suscetível a novas formas de ataque?

Por fim, para que as pessoas sejam responsáveis e de fato usem e tomem decisões com base nos resultados de uma solução de IA, elas devem entender como a solução opera e ter a capacidade de explicar os resultados. Isso é fundamental em situações em que o impacto é alto, por exemplo, em uma solução de serviços de saúde que pode afetar os tratamentos que um paciente recebe. Historicamente, os desenvolvedores precisavam compensar a precisão e a transparência, mas as técnicas mais recentes, incluindo **seleção de modelos e ferramentas de interpretação de modelos**, associadas ao maior rigor na **documentação de dados e sistemas**, podem eliminar essa compensação. É importante projetar soluções de IA para precisão e transparência.

O objetivo de considerar os princípios de IA responsável é gerar confiança na solução e, em última análise, nas pessoas, no serviço e na empresa representados pela solução.

Sobre os autores

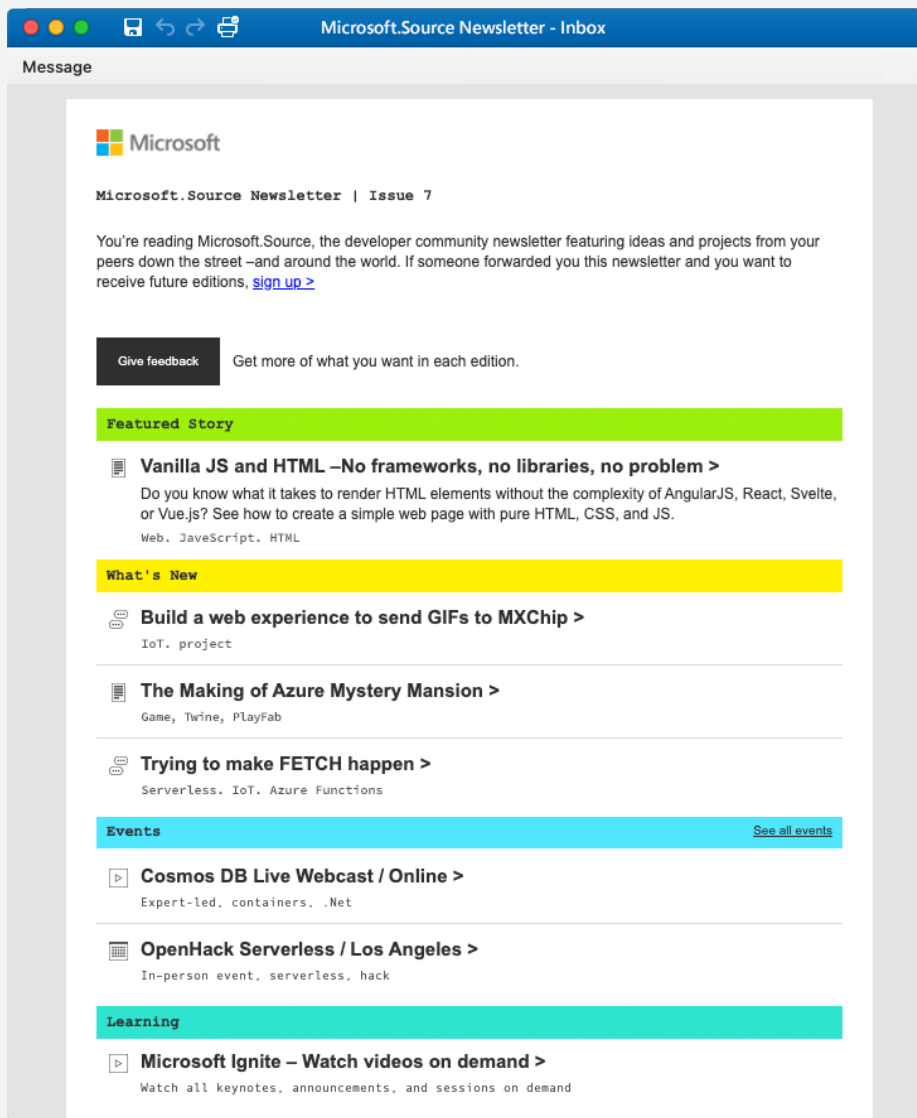
Elaine Chang é líder de desenvolvimento de produtos e sucesso de clientes para IA conversacional na Microsoft, em que ela se concentra em soluções, incluindo o acelerador de soluções e habilidades do assistente virtual. Ela tem sido uma das principais líderes de produtos do Microsoft Bot Framework e colocou o Serviço de Bot do Azure em nível de disponibilidade geral e conformidade empresarial.

Elaine é palestrante em destaque na conferência do Microsoft Build, na conferência do Microsoft Ignite, no Microsoft MVP Summit, no Microsoft AI Innovate, entre outros. Elaine também é uma inovadora estratégica, treinadora profissional certificada e líder de negócios que defende a inovação por meio da diversidade e da inclusão.

Darren Jeford tem mais de 20 anos de experiência em engenharia e arquitetura em uma variedade de indústrias. Na Microsoft, ele exerceu funções de alto impacto e voltadas para o cliente para arquitetar e fornecer soluções altamente complexas usando uma ampla gama de tecnologias. Nos últimos anos, ele liderou alguns dos primeiros projetos de IA conversacional para diversas organizações.

No momento, Darren é arquiteto-chefe na equipe do Bot Framework na Microsoft, em que lidera a equipe de assistente virtual para permitir experiências conversacionais complexas com os principais clientes e o ecossistema de desenvolvedores mais amplo.

Darren é palestrante regular em eventos da Microsoft e também é autor de dois livros sobre o Visual Studio e o BizTalk Server.



Por desenvol- vedores, para desenvol- vedores

Microsoft.Source newsletter

Receba artigos técnicos, código de exemplo e informações sobre os próximos eventos no Microsoft.Source, o boletim informativo mensal da comunidade de desenvolvedores.

- Acompanhe as tecnologias mais recentes
- Conecte-se com colegas em eventos da comunidade
- Aprenda com recursos práticos



Inscrição