

# Microsoft introduces guidelines for developing responsible conversational AI

By Lili Cheng — Corporate Vice President, Conversational AI

As more of our partners, clients and customers set out to design conversational interfaces such as chatbots and virtual assistants, they often ask us for advice on how to develop these technologies in a way that will benefit people while also maintaining their trust. Today, I'm excited to share guidelines that we've developed for responsible development of conversational artificial intelligence, based on what we have learned both through our own cross-company work focused on responsible AI and by listening to our customers and partners.

The field of conversational AI isn't new to me or to Microsoft. In fact, I've been working on conversational interfaces since 1995 when we developed Comic Chat, a graphical chat service that was embedded in an early version of Internet Explorer. The lessons we've learned from those experiences, and from our more recent work with tools such as Cortana and Zo, have helped us shape these guidelines, which we follow in our own efforts to develop responsible and trusted bots.

These guidelines are just that – guidelines. They represent the things we've found helpful to think through, especially when designing bots that have the potential to affect people in consequential ways, such as helping them navigate information related to employment, finances, physical health and mental well-being. In these situations, we've learned to pause and ask: Is this a situation in which it's important to make sure there are people involved to provide judgement, expertise and empathy?

In addition to these guidelines, we hope you'll take advantage of other tools we offer, such as the offensive text classifiers in the [Microsoft Bot Framework](#) to protect your bot from abuse and [Microsoft Azure Application Insights](#) to build traceability capabilities into your bot, which are helpful in determining the cause of errors and maintaining reliability.

In general, the guidelines emphasize the development of conversational AI that is responsible and trustworthy from the very beginning of the design process. They encourage companies and organizations to stop and think about how their bot will be used and take the steps necessary to prevent abuse. At the end of the day, the guidelines are all about trust, because if people don't trust the technology, they aren't going to use it.

We think earning that trust begins with transparency about your organization's use of conversational AI. Make sure users understand they may be interacting with a bot instead of – or in addition to – a person, and that they know bots, like people, are fallible. Acknowledge the limitations of your bot, and make sure your bot sticks to what it is designed to do. A bot designed to take pizza orders, for example, should avoid engaging on sensitive topics such as race, gender, religion and politics.

Think of conversational AI as an extension of your brand, a service that interacts with your customers and clients using natural language on behalf of your organization. Remember that when a person interacts with a bot that represents your organization, your organization's trust is on the line. If your bot violates your customer's trust, then their trust in your organization may in fact be violated. That's why

the first and foremost goal of these guidelines is to help the designers and developers of conversational AI build responsible bots that represent the trust in the organization that they represent.

We also encourage you to use your best judgement when considering and applying these guidelines, and to also use the appropriate channels in your organization to ensure you're in compliance with fast-changing privacy, security and accessibility regulations.

Finally, it's important to note that these guidelines are just our current thoughts; they are a work in progress. We have more questions than we have answers today. We know we'll learn more as we design, build and deploy more bots in the real world. We look forward to your feedback on these guidelines and working with you as we work toward a future where conversational AI help us all achieve more.

Related:

Read: [Responsible Bots: Ten Guidelines for Developers of Conversational AI](#)

Learn more about [Microsoft's approach to AI](#) and take a look at our book, [The Future Computed](#)

About [Lili Cheng](#), Corporate Vice President, Conversational AI