

Benedict Evans

HOME NEWSLETTER CONTACT TWITTER

Chat bots, conversation and AI as an interface

March 30, 2016 In AI

Chat bots tap into two very current preoccupations. On one hand, the hope that they can actually work is a reflection of the ongoing explosion of AI, and on the other, they offer a way to reach users without having to get them to install an app.

One way to look at any given AI problem, in the crudest terms, is to ask whether solving this needs 'general AI' or whether your domain is narrow enough and your solution broad and flexible enough that you can deal with a wide enough percentage of potential scenarios without having HAL 9000. That is, we're clearly pretty close to making a car that can handle highway driving with no human input, since the range of possible events is pretty small, but is driving in central Rome or Moscow an entirely different matter that needs an entirely different level of decision-making? What does that hand gesture mean?

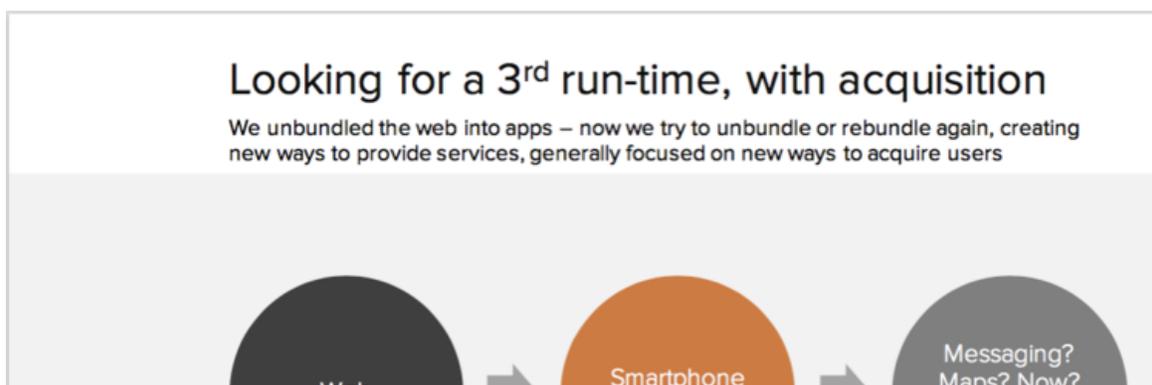
Hence, the challenge in plugging an AI into a 'conversational' chat bot interface is that you don't have HAL 9000 but are in some sense pretending to the user that you do. You speak to it, it speaks back, and it uses natural language (either in voice or text), but it isn't general AI at all. So, how likely is it that the users can ask something that breaks it? How badly? How narrow is the domain of questions that they might ask, and how can you define the users' expectations such that they understand the domain? Perhaps more importantly, can you get people to accept the domain?

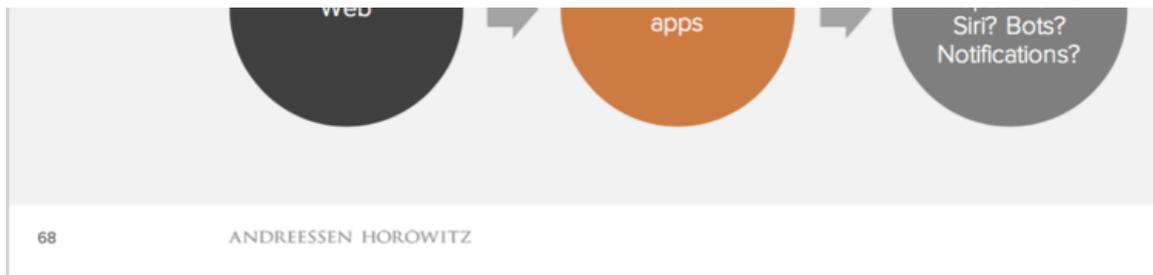
A good way to see this problem in action is to compare Siri and Google Now, both of which are of course bots *avant la lettre*. Google Now is push-based - it only says anything if it thinks it has something for you. In contrast, Siri has to cope with being asked anything, and of course it can't always understand. Google Now covers the gaps by keeping quiet, whereas Siri covers them with canned jokes, or by giving you lists of what you can ask. The actual intelligence might (for the sake of argument) be identical, but you see Siri failing.

To invert this, the Siri 'this is what you can ask' screen is essentially a command line help prompt, whereas the whole point of a GUI was that you didn't have to know what you could type anymore. Though natural language processing means you don't need to know a specific syntax for Siri or a bot, there's still a basic discovery issue - what can I ask, given I can't ask anything?

All of this means that for now, it seems that a bot or conversational UI might work best for something very specific - where the user knows what they can ask, and where those are the only things that they will ask. However, when it does work, it becomes very interesting indeed, particularly now because it happens to align pretty well with the second preoccupation - getting around the app-installation problem.

Instead of going through all the pain of persuading someone to install your app, what if you could have an immediate, fluid and easy way to engage through a conversational interface? Could you make chat bots a third runtime, after the web and native apps, that gives you a new interaction model? And, of course, could that have a new way to acquire users that works better than the App Store? (In other words, 'never mind the dream of artificial intelligence - will this let me stop giving so much money to Facebook?')





That is, the immediate promise of a conversational UI is less something that you do within your own app than that it might make it much easier to interact with your users without having to get an app installed in the first place.

The practical question here, of course, is, well, what runtime? We had the smartphone OS with its APIs, and we have the web browser, but what's the local target for a bot? If a bot is a command line UI, what's the terminal? How does it get onto the phone? And does it have its own meta-mechanisms for user acquisition and discovery of services as well as the runtime itself? Where are those?

One pretty obvious smartphone trend has been for interaction models not just to proliferate but to move up and down the stack - up the stack into new apps that themselves act as platforms (Facebook, WeChat, maps etc) and down into the OS (Siri, Now, deep links etc). So too with bots: Apple has not opened up Siri to developers (yet), but Google is selectively opening Now, and it would certainly seem to be strategically desirable for the OS provider to be the one creating any new runtimes.

However, the challenge for both Apple and Google here is the other end. They can put a bot terminal on the phone, but for anything in the cloud to talk to you on your phone, that service has to know where to look. Both companies have some form of cloud identity platform, but Apple's privacy strategy precludes many uses of this, and neither Apple nor Google have a platform today that Gap or Instacart might ask you to log into when you place your order, such that they can talk to you about it afterwards. Of course, this could change - in particular, the rumour that Apple will extend Apple Pay (and by implication Touch ID) to the web opens up lots of possibilities in this direction.

Moving up the stack, Facebook has both ends - the terminal, in Messenger (which now has 1bn mobile MAUs), and the web identity platform. Facebook's challenge is in a different place: would publishers want to give it so much control, and would they trust it to continue the platform into the future, given its track record?

Meanwhile, I found Microsoft's heavy focus on bots and associated tools at 'Build' pretty interesting, because of course Microsoft has neither end - it doesn't have a mass-market smartphone platform or a web identity platform that could be the connective glue. (It does have Skype, with 300m MAUs, but I'm not at all clear that that has the right market positioning to be a viable terminal in this sense.) But a lot of Microsoft's narrative is about building cloud platforms to enable other people to create bots that could then run between any two other end-points - between FB Messenger or Slack at one end and Gap at the other, logged in through a Facebook identity, but with the actual capabilities coming from Microsoft. That is, Microsoft as, well, a development platform. Naturally, this is Google's self-image too - what is Google if not a platform for understanding and communicating data at massive scale?

I've deliberately been using the rather retro term 'terminal' to describe the place a bot UI lives in on the phone because that seems to be the way many people think about it - there's an AI in the cloud and you talk to it through a dumb text or voice UI on the phone or PC. That also reflects the way Now works - the device is mostly dumb glass. This is an extension of the shifting of interaction layers up and down the stack I mentioned earlier - here you're in some senses shifting the interaction up off the device entirely. Coincidentally, 'terminal' is the word telcos used to use to refer to mobile phones before smart happened at all, and indeed many of the sample use cases I've seen for chat bots could be done with SMS, or perhaps USSD.

But that's not the only potential model. WeChat uses messaging as a platform but doesn't push interaction down to the level of 'dumb text' - it provides its own social and acquisition model but much of the actual interaction with third party services happens in rich UIs (mostly built as web views, as it happens). And after all, if you're offering a user a question to which there are only two options, should you tell them 'you can reply 'red' or 'green'', or should you give them two buttons within the chat? What if there are five options? Should you perhaps construct some

sort of on-screen interface for your users that lays out, graphically, the options? You could call, it, perhaps, a 'GUI'. You could have 'links' that you tap on, that load new 'pages'... And indeed, if you've got your chat bot working, does that need to be in Facebook, or could it be on your own website too? It depends what kind of interactions you're looking for, and maybe whether you're solving your own problems or your users'.

There's an old computer science idea that a computer should never ask a question that it should be able to work out the answer to. One promise of AI is that you can get, not so much a computer that you can talk to like a person, but one that you don't need to talk to at all, or much less - that you can remove the mental load and friction and maintenance of engaging with a computer or online service. The computer can ask fewer questions and work out more by itself. (This is somewhat analogous to the reduction in mental load involved in moving to iOS or Android from Windows or Mac). Instead of going back to the website and logging into your order and editing it, you can just send a message: 'cancel the shirt and get me the green bag instead of the grey one'. A computer ought to be able to work that out, and will probably be able to. You ought to be able to use AI, and chat, in some way, to achieve the same aims with much less work. But it doesn't follow that that's a great way to do everything. If it takes you more time to work out what you can ask the AI assistant than to drag the meeting to a new slot on your calendar, you're doing it wrong. An AI shouldn't be more mental load than just tapping the damn button - an IVR with better buzzwords. That might be a better user acquisition model for you, but that's tough. And so this comes back to my opening point - is this a domain where you can just tell it what to do and move on, or one where the 'general intelligence' still needs to be between the chair and the screen?

 AI, Messaging

 85 Likes  Share

 Newer Older 

Comments for this thread are now closed.



28 Comments

Benedict Evans

1 Login ▾

♥ Recommend 16

🔗 Share

Sort by Oldest ▾

**kevinmarks** • 5 months ago

Why do you think that Google Now is passive? You can say "ok google" or press the mic button and ask it things or give commands. Some come back as voice, some as onebox UI, some as search results.

2 ^ | ▾ • Share ›

**d_n** → kevinmarks • 5 months ago

You're confusing Google Now with Google Voice Search.

"Google Now Is Not Google Voice Search":

<http://googlesystem.blogspot.c...>

^ | ▾ • Share ›

**kevinmarks** → d_n • 5 months ago

Maybe true in 2013, they're integrated now as you can give commands that set things up later in Now. Or if they are formally distinct, Ben should be discussing Google Voice Search as well.

3 ^ | ▾ • Share ›

**d_n** → kevinmarks • 5 months ago

Ah yes, fair point. Google's branding has always been a bit confusing (Google Wallet comes to mind).

^ | ▾ • Share ›

**pbreit** • 5 months ago

Would be interested to hear a take on Amazon's efforts.

2 ^ | ▾ • Share ›

**GJSmyth** • 5 months ago

Microsoft, in providing the tools and capabilities to developers puts it's platforms, and integrated connectivity to it's own services, into focus of developers building these new interactions.

It also, no doubt, sees bots as a means of breaking down the dominance of Android and iOS ecosystems by facilitating the proliferation of this '3rd runtime' - this reduces the 'app gap' not by building out duplicate apps in another app store but by making some of those existing iOS/Android apps redundant. They appear to be looking past the existing mobile landscape to position themselves for what is next.

Im puzzled by Benedict's comment that Microsoft doesn't have a web identity platform to aid as connective tissue - Microsoft Accounts are used by all MSFT online properties (Outlook.com/Hotmail, Xbox live, Skype, Office 365,

OneDrive, Windows 8 and Windows Phone) as well as by some third party services. Many of these are overlapping users but the sum total of unique users would be a large number.

^ | v • Share ›



Divraj Jain • 5 months ago

The most valuable bot will be preemptive, platforms with scale who first leverage private will probably build inexorable winner take all advantage as they build in a habit + network effects. While narrow specific bots seem like a step on the ladder we will quickly shift to monopoly.

2 ^ | v • Share ›



jacobgio • 5 months ago

interesting post to complement with @AlbertWenger vision where we need "the right to be represented by a bot" => [http://continuations.com/post/...](http://continuations.com/post/)

^ | v • Share ›



Michael Lipman • 5 months ago

I really enjoyed this discussion on the merits and drawbacks of chat based interfaces and bots <https://medium.com/@MattHartma...>

1 ^ | v • Share ›



Andrew McKay • 5 months ago

It's interesting to see apps like Telegram attempt to find a mix between the "terminal" bot interaction and the rich UI interaction of WeChat, through customized keyboards that change based on the context of the conversation. Thoughts on this from anyone?

6 ^ | v • Share ›



Patrick Kane • 5 months ago

Great piece. Particularly like the point that many bot use-cases could be done with SMS... One question - I've read variations of this line in a few places now: "Instead of going through all the pain of persuading someone to install your app..." Is this really an issue? App discovery, sure, but friction on an install? Using my fingerprint to download an app is easier than 1-click on Amazon. I haven't come across this viewpoint from either a consumer or a publisher - what am I missing?

1 ^ | v • Share ›



Fernando → Patrick Kane • 5 months ago

1-click isn't actually all that easy, hence Amazon Dash, Echo and their attempt at a phone. Apps are similar. This is why Apple Maps has a larger use base on iOS despite being an inferior product. There's also this: <http://qz.com/253618/most-smar...>

^ | v • Share ›



Bjorn → Fernando • 5 months ago

"This is why..."

Apple Maps is pretty great now. I use it by choice, and I doubt I'm alone.

^ | v • Share ›



Quicksingle → Bjorn • 5 months ago

I'm not sure why you would choose to use such an obviously inferior map. Apple maps is still the but of jokes amongst all my friends. Siri is close though.

^ | v • Share ›



atimoshenko • 5 months ago

Language is pretty good for discussing ideas, but for giving instructions we tend to show, rather than tell. We demonstrate, gesticulate, draw, send screenshots, etc., if at all possible. "Cancel the shirt and get me the green bag" would work only in the simplest of circumstances, even with a human PA. Or, for all the problems people have with Ikea instructions, imagine one that was just text. Imagine having to read it... or having to write it.

Chat may well end up being useful in some niches, but the third run time it will not be.

2 ^ | v • Share ›



James Doody → atimoshenko • 5 months ago

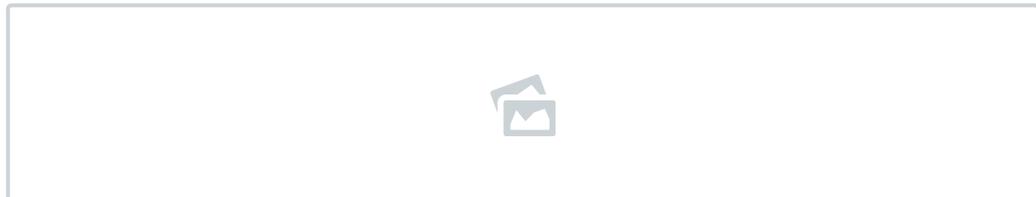
The third run time will be a hybrid of apps and web services. Chat will be a huge part of that new run time. But there are at least four other services to add into the mix.

^ | v • Share ›



srust99 • 5 months ago

totally agree with messaging as the new runtime, Maps as the new browser and notifications as the new apps....voice recognition for dictation and translation, but not too sure about the "chatbots" ...yet...



^ | v • Share ›



Marko Balabanovic • 5 months ago

Some great points about discoverability of syntax and the right situations where a conversational UI is a useful approach. Any case where there's a more complex search with lots of optional fields is a good candidate to look at. For

complex search with lots of optional fields is a good candidate to look at. For example, travel search involves dates, destinations, room types, numbers of people, ages of children, styles of room or seat - all kinds of parameters, many optional, leading to cluttered GUIs. In 2008 what was then lastminute.com labs created the beginnings of a conversational UI for travel and entertainment, reducing 78 separate search boxes across all categories like hotels, flights, restaurants and so on (or 162 if you counted "advanced search") down to one, working in multiple languages. But discoverability of syntax was indeed a problem, that we addressed with a context-sensitive autocomplete to guide people towards valid queries (you can see an old screenshot here <http://www.balabanovic.com/mar...> This works well visually with typing, but not with voice. As it goes beyond a standard chatbot interface, it would be hard to deploy across today's iOS or Android ecosystems, unless combined with a custom keyboard (note Microsoft's recent purchase of SwiftKey, and also Telegram gradually moving in this direction). By giving users the freedom to type in a more conversational interface, you get deeper analytics on users' intentions and potentially a rich set of training data (I remember in testing in France seeing the query "I want to go to the mountains" that in a perfectly reasonable request that even now most travel sites will struggle with).

1 ^ | v • Share >



Nadav Gur → Marko Balabanovic • 5 months ago

Funny. We went through a very similar process at Desti (an AI-driven travel planner) that I described here a couple of years ago - <https://medium.com/@nadavgur/w...>

... and now we're working on chat-as-a-platform and are really debating the same themes. I love to quote a discussion Adam Cheyer of Siri had with Steve Jobs back in the day, where he was asked what he thought speech was good for, and he said something like "if it's on the screen I want to tap it, but if it isn't I want to ask for it". Tapping is easier than asking - if there's something to tap. The clever UI would make sure most of what I can do can be tapped (i.e. there's a widget on the screen), but I can still jump ahead or back a few steps by expressing something else by typing (or voice eventually).

1 ^ | v • Share >



Bjorn → Nadav Gur • 5 months ago

Great quote. Unclear to me from your phrasing whether that was Adam or Steve or said it?

^ | v • Share >



Nadav Gur → Bjorn • 5 months ago

Jobs asked, Cheyer responded.

1 ^ | v • Share >

**matedo berrucchi** • 5 months ago

At Your.MD we built the first vertical AI for healthcare with chat. Ben's is very accurate at listing the main questions we are asking ourselves. I think discovery is one of the biggest challenges: how do you find out that you can add a bot to your chat? Slack has created a bot store but it's visually and logically too far from the app itself and despite being the only bot in the healthcare section the installation numbers are less than impressive. The area of bot marketing needs to develop quickly! Try the bot in our apps or on Telegram and Slack www.your.md

1 ^ | v • Share ›

**benvenuto** • 5 months ago

Aren't we just cycling back to thin client? Last time enabled by middleware, SOA design principles and cheap middle tier hardware; this time enabled by mobile connectivity, NLP, voice recognition and machine learning. And on small screen mobile devices, this may well be a better interaction model, since context frequently changes, and the alternative is a range of apps for every occasion.

^ | v • Share ›



This comment was deleted.

**Benedict Evans** Mod → Guest • 4 months ago

I wrote a paragraph on why it's not quite as easy as that for Apple and Google.

^ | v • Share ›



This comment was deleted.

**Henry Thornton** → Guest • 4 months ago

Connor Hays re: "I don't want to have to open Messenger to use M, then open WhatsApp to interact with some bots for ordering stuff, and then jump over to Slack for work related automation - and no one else wants to either."

I've been grappling to understand this same point - multiple chatbots servicing multiple apps or one chatbot to rule all apps or like messaging services, multiple chatbots supporting all apps.

Also, with chatbots it sounds as if the app disappears to a background process. If so, aren't app vendors going to lose brand value?

Lastly, who is liable when a chatbot offends?

^ | v • Share ›



Mark Beccue • 4 months ago

I think in the first phase, the most likely home for chatbots are on chat app platforms - Facebook Messenger, Kik, Slack, perhaps pivoting enterprise social platforms like Yammer, SAP Jam, IBM Connections, where they have more

Subscribe

I do a newsletter every Sunday: the things that were actually interesting, and why.

It's my notebook for the week. Now 55k subscribers.

[\(Archive and more details\).](#)

SIGN UP

I respect your privacy.



JIRA is over. Get project management developers actually love.

ads via Carbon