QConSF

Chatbots and Serverless: A match made in the Cloud

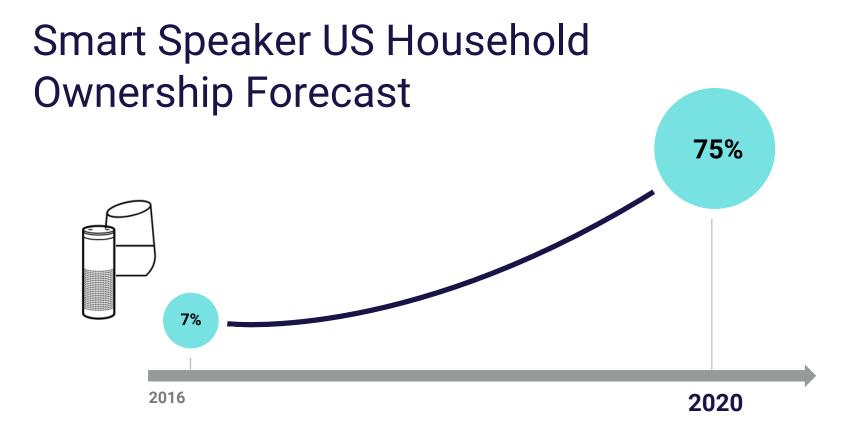




Gillian Armstrong @virtualgill

Technologist Cognitive Technologies

Liberty IT



Smart Speaker US Household Ownership Forecast

The world is changing fast

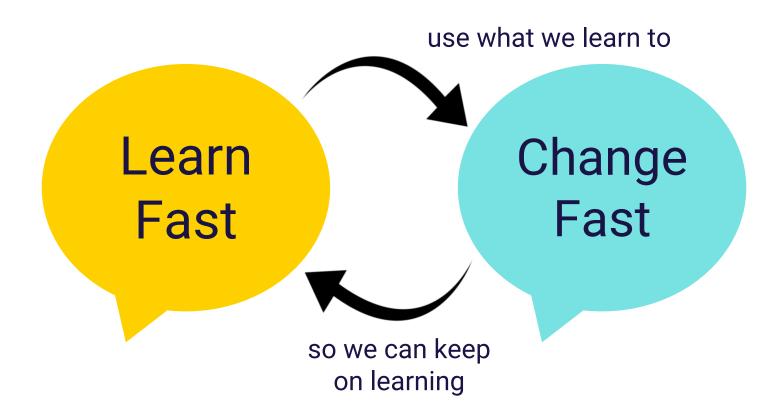
2020

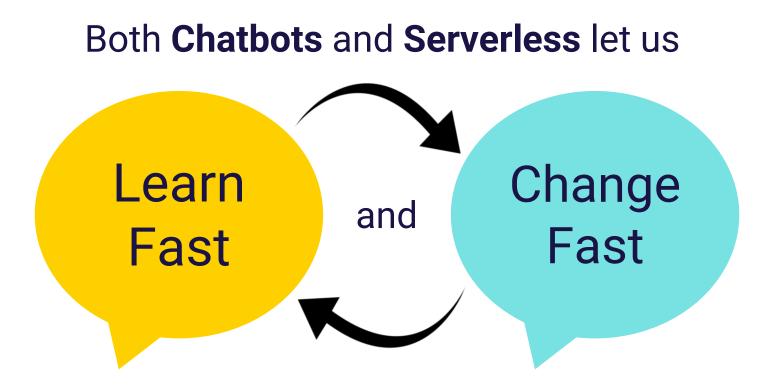
75%

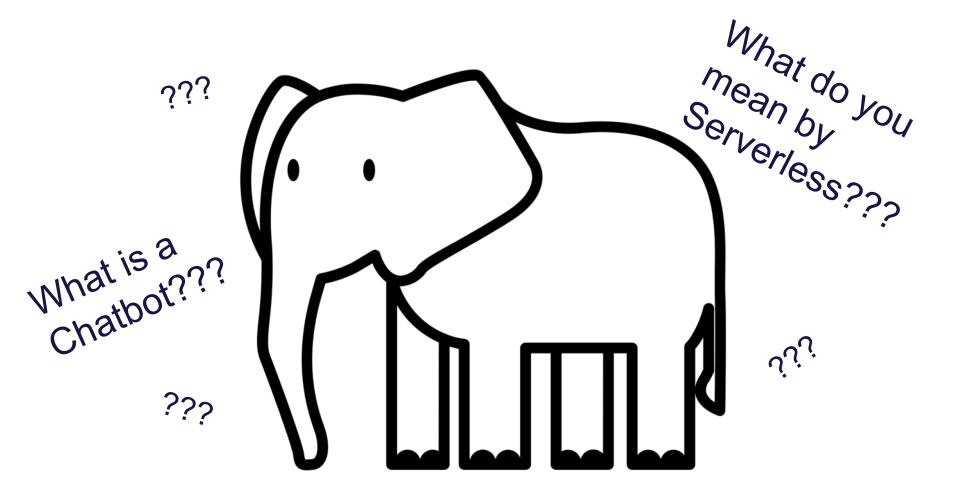
SOURCE: US Census, Gartner, Edison Research, Voicebot.ai

The world is changing fast

So we need to be able to learn fast











For the purposes of this presentation...

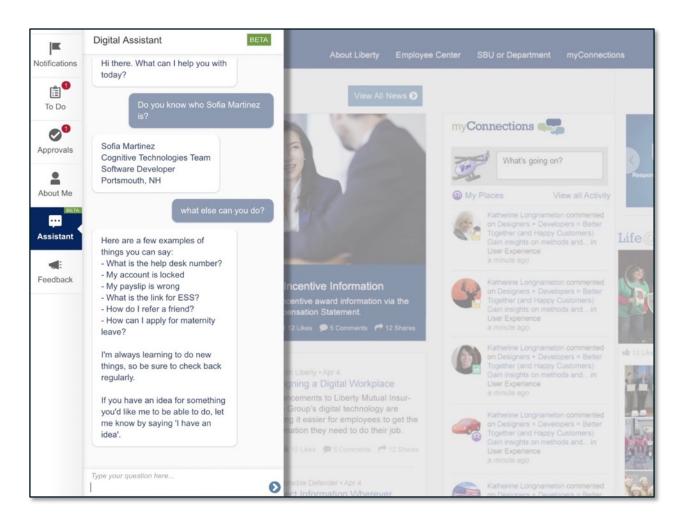
A Serverless solution is one that costs you nothing to run if nobody is using it (excluding data storage cost)

- Paul Johnston

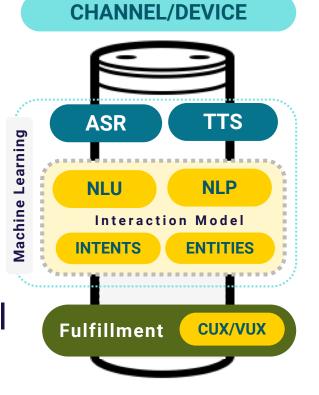
For the purposes of this presentation...

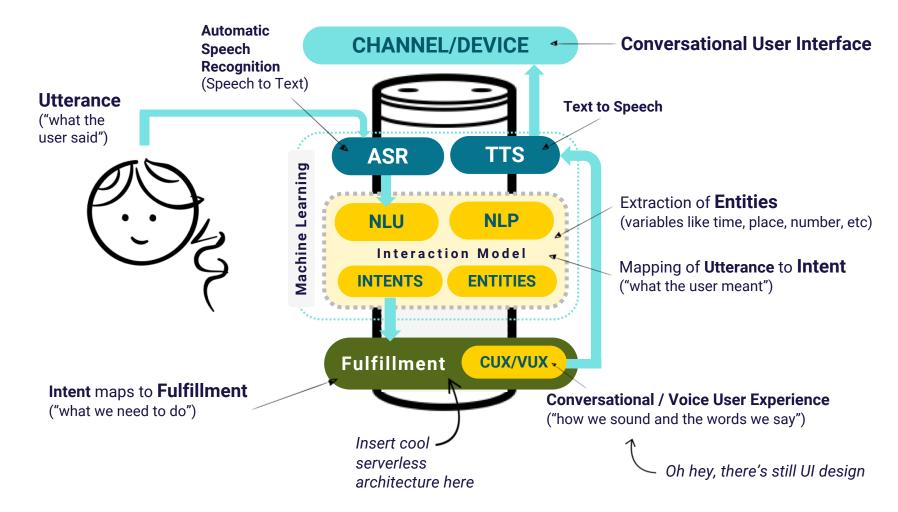
A **Chatbot** is something you can interact with **conversationally** using **natural language**

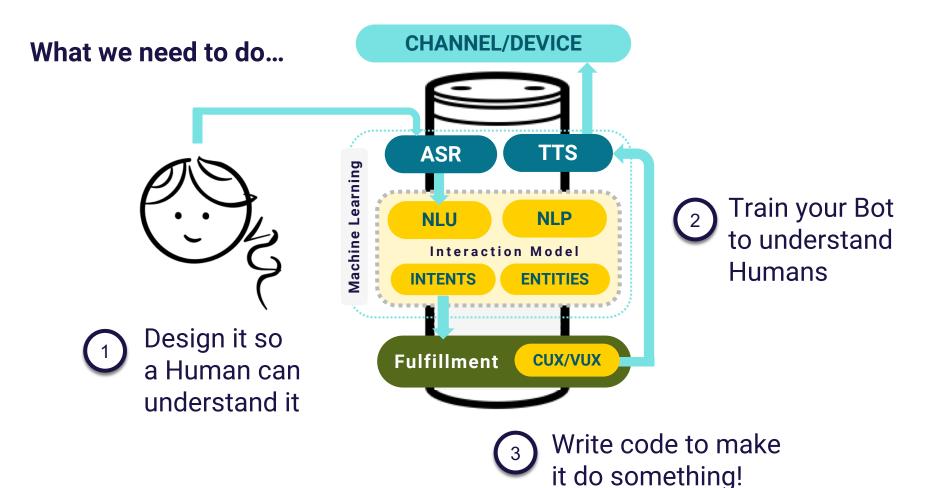
- me

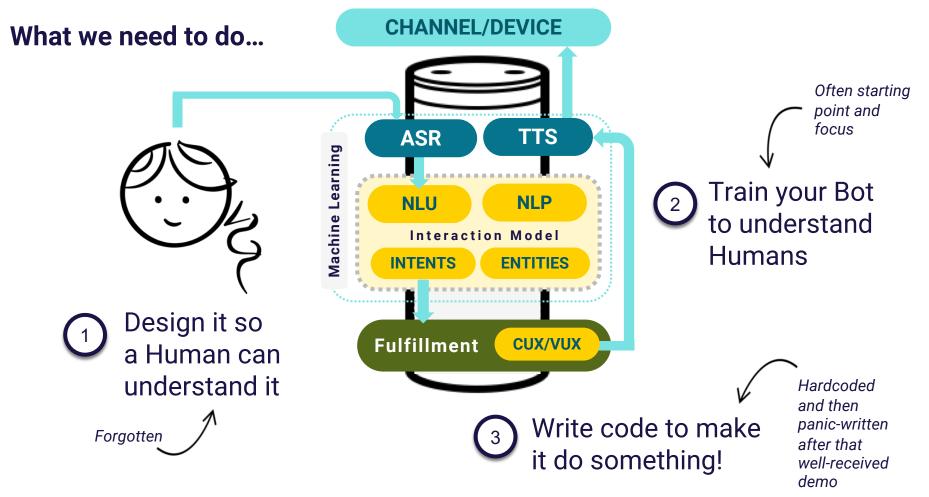


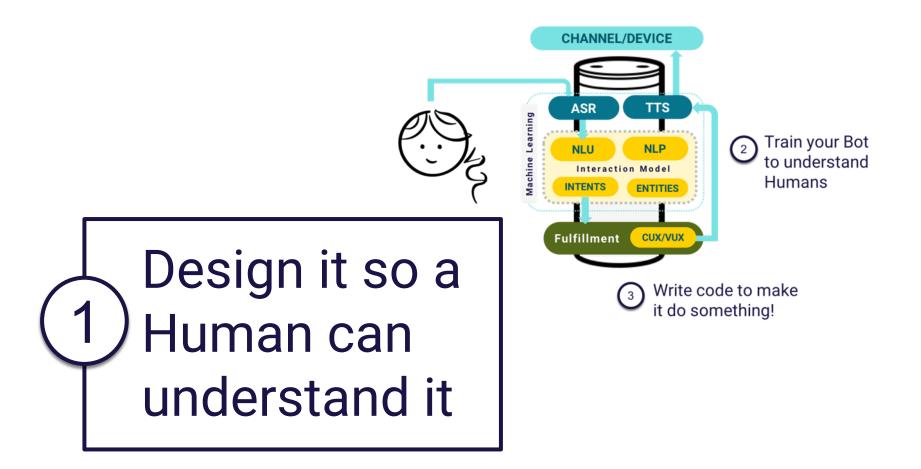
A very quick overview of a Chatbot using Conversational Al



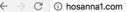








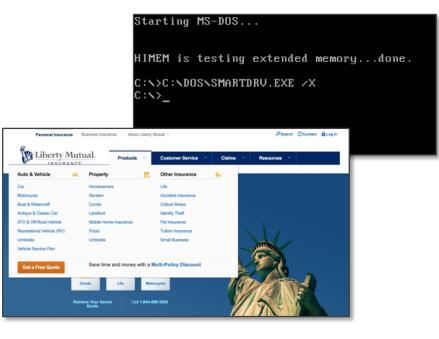
In a chatbot, **Conversation** is the **Interface**





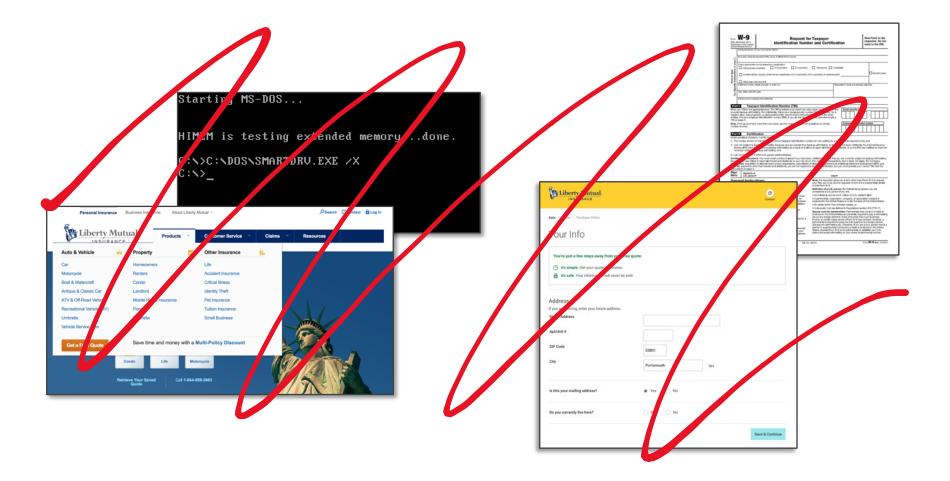
Source: http://hosanna1.com/

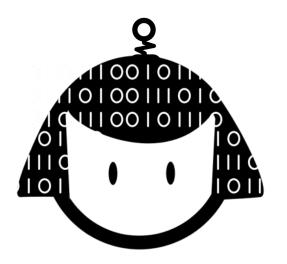
Up to now we've kept designing based on what came before



We've carried the concept of menus and commands into our graphical user interfaces We base our web forms on the paper forms that came before

			W-9	Ro Identificatio	quest fo n Numb	r Taxpayer er and Certification	Give Form to the requester. Do not send to the IHS.
			have (as shown in pilo	numerie skint			-
		1		ind willy rank, it disent from since			
		1	Dast appropriate test	natur 🗋 O Conjunation 🔲	1 Caspondian	Distance Distance	
		Ant o the tendent	C Lobertering of	ngery. Etherholto casakador ()-	Corporation, S	-Lopation Portage	
				and at an advance		Reputer's rome and address	wa (glora)
			04, ##L203P.08				
			La accurt screater	wa jaturali		1	
		110	Tempeyer year Thirt for appropri- ate backup attraction of the peer propriet	Identification Number (T) for loss Tra Thornstein man for indexis, this is provide a design field offer, see the f artification scetter (Title, P pro-	match the name isocurity num will immunities is not have a	re plan or the 'Narra' line Social security he day (2004), However, for a rs or page 3, For other synther, was how to get a	Ť
		The o	P the account is it in a	rs than one name, see the chart o	r page 1 for p	publication Engineer mention	attor runter
		100	Certificat	en .		11111	
		12	n number shown on Hi en not subject to back- mits (Fill) theri an su Honger subject to back	is ferm is my conset tangager lide g withholding because (a) I are o fgiot to back g withholding as a up withholding, and		der (in Lan salling für anumlar to be laued to schup atthölding, or §) i have not been notfield ni to report all interest or dividende, or () the RS	me), and to the internel Revenue Sinas notified me that i am
						an soffieldity the IRS that you are currently subject in Formal vertex tenanctions, here 2 does not a of doot, controls there to an indextual effective to sign the currentlication, but you must provide you	
		Sec.	Burgerson and States of States of U.S. parameters of U.S. parameters of U.S. parameters		- of set		
		Hen	us prace			(un * Note, if a requestor place pro a form other tra-	er Forre 10 å ta rag ogt
					-sta	Note: Fia requester pleas par a term other the poor TAA, par, most use the requester's bene if to this Fast, MI-8, Definition of a U.S, parease. For factorial tor pr considered a U.S, parease. For factorial tor pr	ren substantially similar
Liberty Mutual.				Contact	l mult Lar Filmat	considered a U.S. porter Pyrc. and Arr indektual who is a U.S. Officer or U.S. no A partnership, conjunction, company, or asso separated in the United Dates or under the las- Arr and points other their a freelow anded, or	sidort aller,
							autilian 305,7755-75 that conduct a hinds or pulsed to pay a withinking
> Quote > Purchase Online					t of the lat	tex on any foreign partners share of income th Further, in certain cases where a Farm to be partnering is required to present that a partner	fore such business, an root beam raceived, a normal a foreign particle,
our Info					124	 A device to a particular to regulations of Becalarized to particular to a transmission bulkness in the URAs Device are governally to to an air bulkness particular to and official to familiar to a set of the transmission of the particular particular to a set official to particular to a set official to particular to a set of the particular to a set particular to a set of the particular to a set particular to a set of the particular to a set particular to a set of the particular to a set particular to a set of the particular to a set particular to a set of the particular to a set particular to a set official to a set of the particular to particular to a set of the particular to a set particular to a set official to a set of the particular to a set particular to a set official to a set of the particular to a set particular to a set of the particular to a set particular to a set of the particular to a set particular to a set of the particular to a set particular to a set of the particular to a set particular to a set	en a U.S. polibil Matik a Bushness II: Re Drited II: setabilish your U.S. garhuoshig Income.
					OA.No.	1907 R	tesWilliam Gara
ou're just a few steps away from your free quote.							
It's simple. Get your quote in minutes.							
It's safe. Your information will never be sold.					1		
dress							
u are moving, enter your future address.					1		
et Address					1		
					1		
Unit #					1		
					1		
Code	03801				1		
					1		
	Portamouth	NH			1		
is your mailing address?	Yes No						
ou currently live here?	O Yes O No						
			Sav	e & Continue			
			_	_	1		

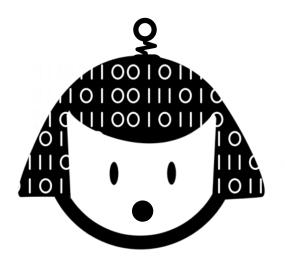




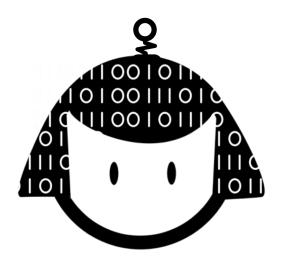
People are the original conversational interface

Model your design on a **human**, not a website

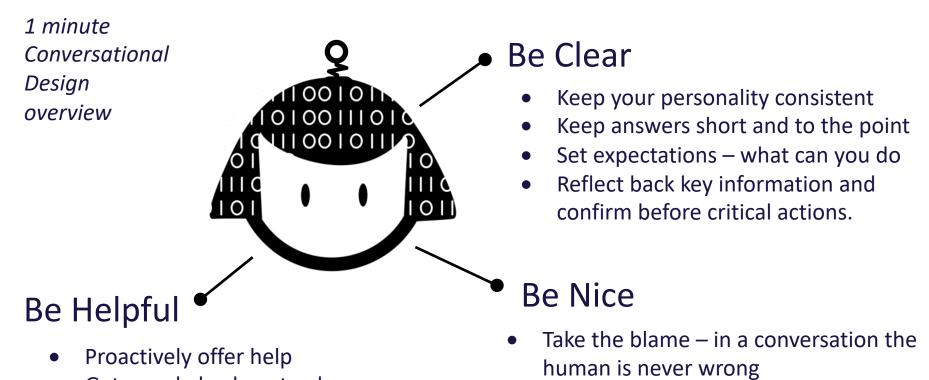




When you start with thinking about trying out a design you may discover that a chatbot isn't even the right thing to build...

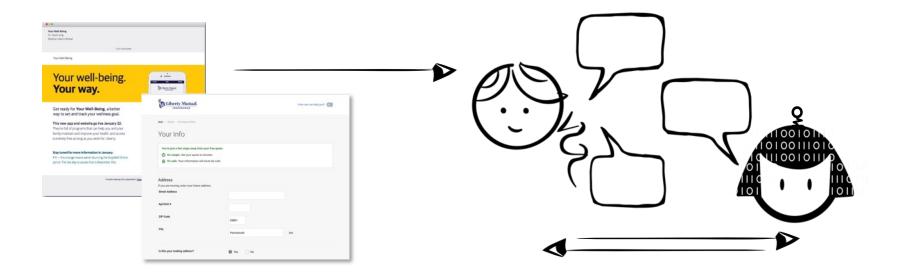


But assuming it is...



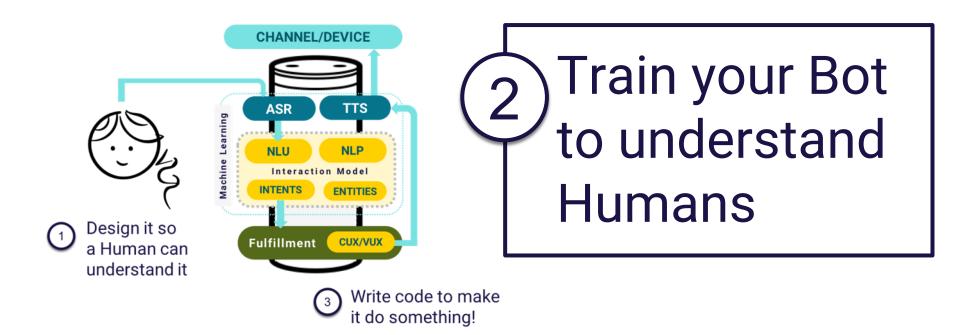
- Get people back on track
- Repetition is annoying
- Repetition is annoying

- Chatbots aren't smarter than a human
- Be understanding if people don't want to talk to you offer alternatives.



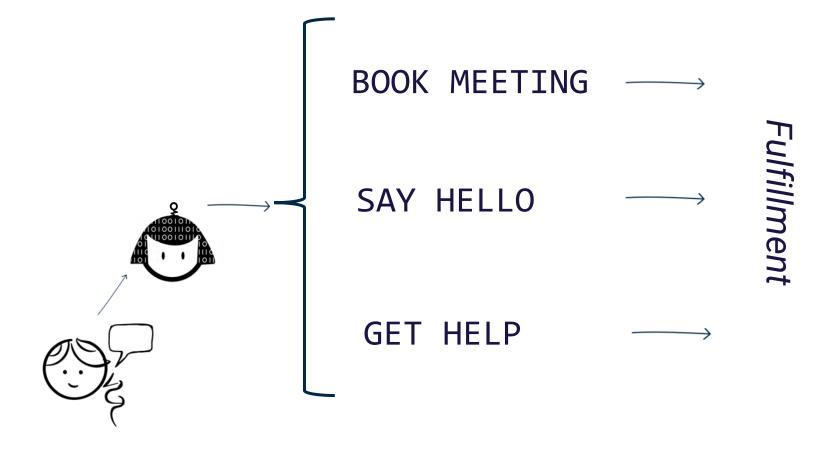
The Web is about *Telling* the user what to do

Conversational UI is about **Listening** to what the user wants to do



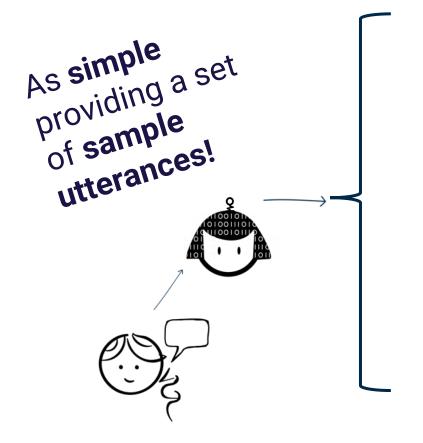
Amazon Lex Console

aws	Service	15 ~	Resource Groups ~	*			û	e-sami-developer/a0 N	. Virginia 👻 Support 👻
	stantBot La						Build Publish @	> Test bot (Latest)	Ready. Build complete.
Editor	Settings	Channe	•						
Intents	0		demo_BookMeeting	gIntent 2 - Edit					
demo_BookMeetingInte demo_CalendarAvailabilit demo_GenericHelpIntent	-		- Sample utterances	0					
			please book a meeting at	(time)					testing. Type an utterance
demo_GoodbyeIntent demo_GreetingIntent			i want to book a meeting					below to begin conv	ersation with your chatbot.
demo_Orde	-		i need a meeting booked						
Slot type		٨/	hat t	ho M	ah	Tuto	riala	told	mo
No slots cre	atari	V	ιαιι		GD	IULU	11012	ιυίu	
NU SIDIS CIE				on with (nerson) at (time) on fr					
Error Handli	ng		ildin		ha	tbot l	ooko		chat history
		JU		iy a C	IIA		UUNE		て
			book time in my calendar	with person}				Inspect response	
			book me a meeting {time						Hide
			book me a meeting with	(person)					
			book me a meeting with	(person) at (time) on (date)					
			book me a meeting on [d	late) at (time)					
			book me a meeting on [d	late)				When you chat with	your bot, you can see the
			book me a meeting						intent and the response here.
			book a meeting on (date)						
			book a meeting with (per	son) at <mark>{time}</mark>					
			book a meeting with (per	son) (date)					
			book a meeting with {per	son)					



Utterance

Intent



Book Meeting at <TIME> on <DATE> BOOK MEETING Set up Appointment Create event on <DATE> Set up Meeting Book me a meeting at <TIME> Hello

- HELLO SAY
- Help HELP What can you do Help me GET

Hey

I need help

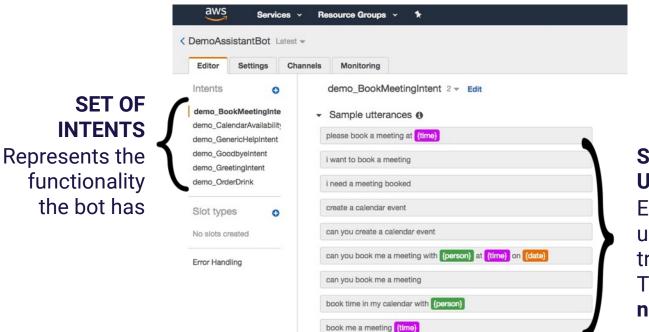
Hi there

Good Morning

Utterance

Intent

Amazon Lex Console



SAMPLE UTTERANCES

Examples of what a user might say to trigger this Intent. This is **training data not a regex** match.

Amazon Lex Console

Lambda initialization and validation
 Initialization and validation code hook

book a meeting with {person} at {time}

book a meeting with {person}

book a meeting with {person}

SLOTS (ENTITIES) are defined showing type (place, name, time, custom, etc.)



The sample utterances show where to expect the data to be supplied in an utterance

O

Ο

O

If the user does not supply this data the bot knows to prompt for it.

FULFILLMENT location defined.

Intent and **Slot** matches will be sent to this lambda or returned directly to client.

	uired Name	Slot type
	e.g. Location	e.g. AMAZON.US_CITY
1. 🗸 🗹	numberOfAttendee	AMAZON.NUMBER
2. ^ ~ 🗹	person	AMAZON.Person
3. ^ ~ 🗹	date	AMAZON.DATE
4. ^ 🗸	time	AMAZON.TIME
Confirmation pr	ompt 0	
Confirmation pr Fulfillment () AWS Lambda fr		s to client
Fulfillment 0		s to client
Fulfillment () AWS Lambda fr	unction C Return parameter	s to client

- Slots O

Publish to make **Build** to create the the functionality Interaction Model live to users. Services - Resource Groups - * ↓* Im-federate-sami-teveloper/a0... * N. Virginia * Support * aws Build Publish < DemoAssistantBot Latest + Test bot (Latest) Ready. Build complete Editor Settings Channels Monitoring demo_BookMeetingIntent Latest + Intents 0 demo BookMeetingIntent Sample utterances 0 demo_CalendarAvailabilityIntent e.g. I would like to book a flight. 0 demo GenericHelpIntent You're now ready for testing. Type an utterance demo_Goodbyeintent below to begin conversation with your chatbot 0 please book a meeting at (time) demo_GreetingIntent demo_OrderDrink i want to book a meeting 0 Slot types 0 i need a meeting booked 0 No slots created 0 create a calendar event Error Handling can you create a calendar event 0 **Clear chat history** can you book me a meeting with (person) at (time) on (date) 0 Chat with your bot... 0 can you book me a meeting Inspect response book time in my calendar with {person} 0 book me a meeting (time) 0 book me a meeting with forces 0

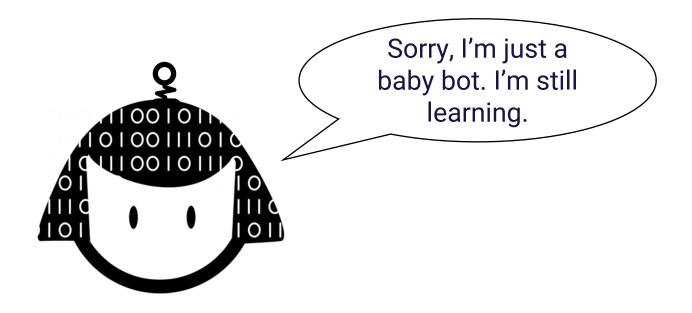
Built-in **Test Console** allows you to test directly here after building, but before publishing.

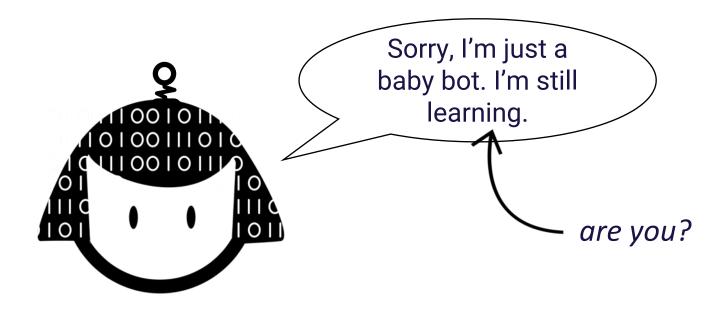


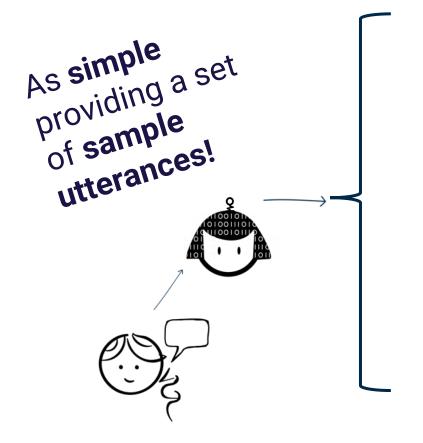
Easy to **connect** your bot to popular **channels** through console.

Editor	Settings	Channels	Monitoring					
hannels			acebook				-	
acebook				w and click activate to get tegrate with Facebook.	a callback URL to use with Facebook. Y	'ou can generate m	ultiple callback URLs	. Learn
k ack			in the second second second	logialo mari accoccia				
vilio SMS				Channel Name*		0		
				Channel Description		0		
				IAM Role	AWSServiceRoleForLexChannels	0		
					Automatically created on your behalf			
				KMS Key	aws/lex	- 0		
				Alias*		- 0		
				Verify Token*	Verify Token	0		
				Page Access Token*	Page Access Token	0		
				App Secret Key*	App Secret Key	0		
			Required Field		Activate			

aws Services	∽ Resource Groups → 🏃	û. Îm-federate-sami	-developer/a0 • N. Virginia • Support •
< DemoAssistantBot Latest	•:	Build Publish @	Test bot (Latest) Ready. Build complete.
Editor Settings Ch	anneis Monitoring		
Intents O	demo_BookMeetingIntent 2 - Edit		
demo_BookMeetingInte demo_CalendarAvailability	✓ Sample utterances ⊕		
demo_GenericHelpIntent	please book a meeting at (time)	Λο	You're now ready for testing. Type an utterance below to begin conversation with your chatbot.
demo_GoodbyeIntent demo_GreetingIntent	i want to book a meeting		
demo_OrderDrink	i need a meeting booked		
Slot types o	create a calendar event		
No slots created	can you create a calendar event	04	
Error Handling	can you book me a meeting with (person) at (time) on (date)		Clear chat history
	can you book me a meeting		Chat with your bot
	Chatbot Don		spect response
	Ondibot Don		Hide
	book me a meeting with (person)		
	book me a meeting with (person) at (time) on (date)	/ •	
	book me a meeting on (date) at (time)	•	
	book me a meeting on (date)		When you chat with your bot, you can see the
	book me a meeting		fulfillment state of your intent and the response here.
	book a meeting on (date)		
	book a meeting with (person) at (time)		
	book a meeting with (person) (date)		
	book a meeting with (person)		







Book Meeting at <TIME> on <DATE> BOOK MEETING Set up Appointment Create event on <DATE> Set up Meeting Book me a meeting at <TIME> Hello

- HELLO SAY
- Help HELP What can you do Help me GET

Hey

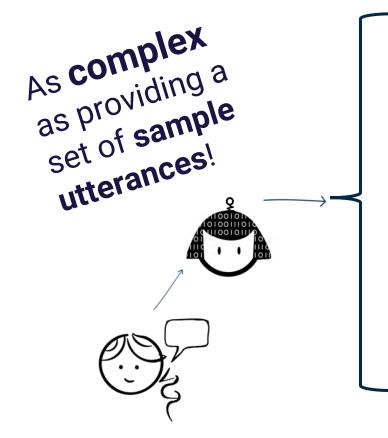
I need help

Hi there

Good Morning

Utterance

Intent

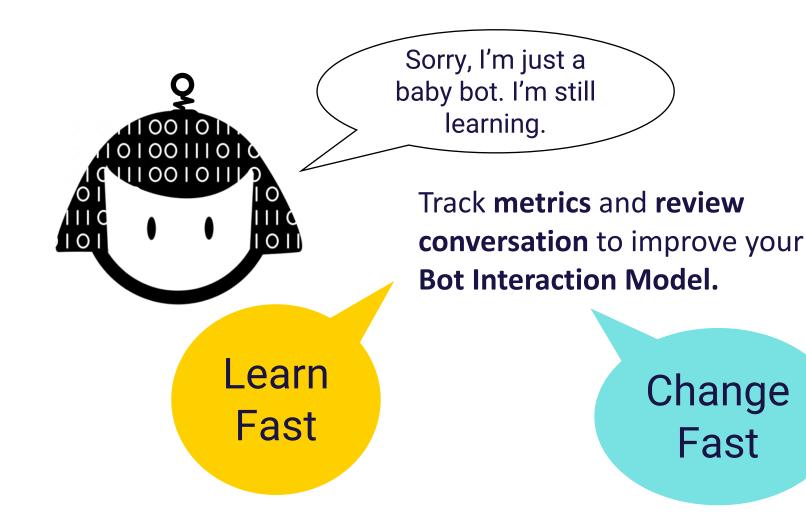


Model your design on a **human**

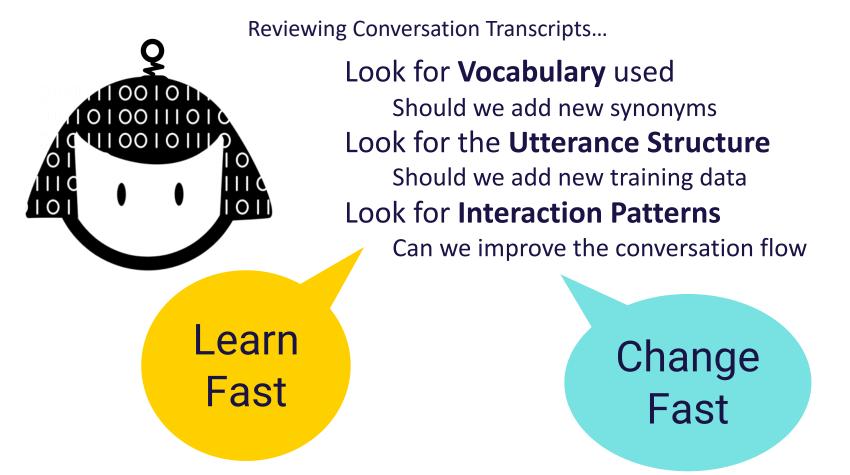
Get your sample utterances from your users... you wouldn't make up other machine learning training data!

Utterance

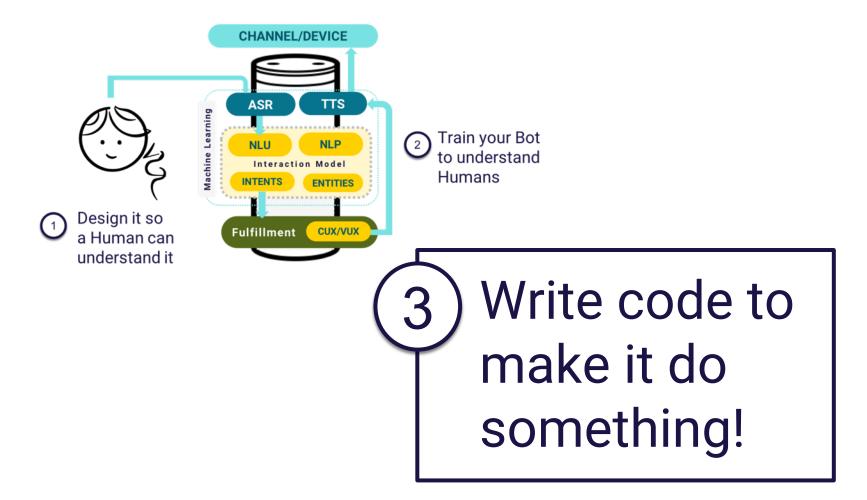
Intent

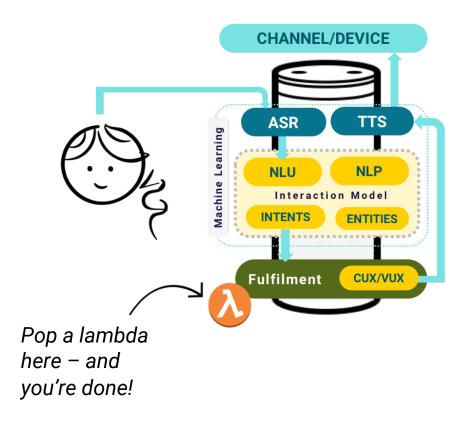


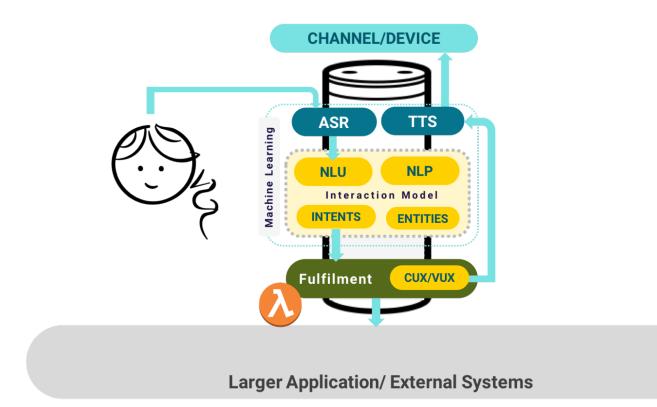
Gillian Armstrong @virtualgill



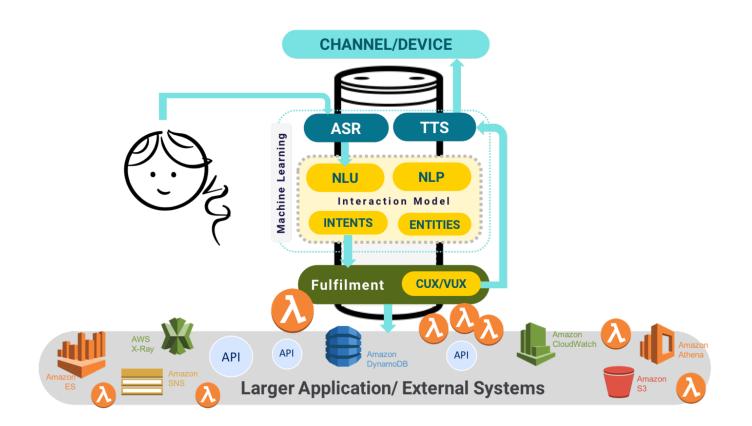
Chatbots aren't magic (although I wish they were)

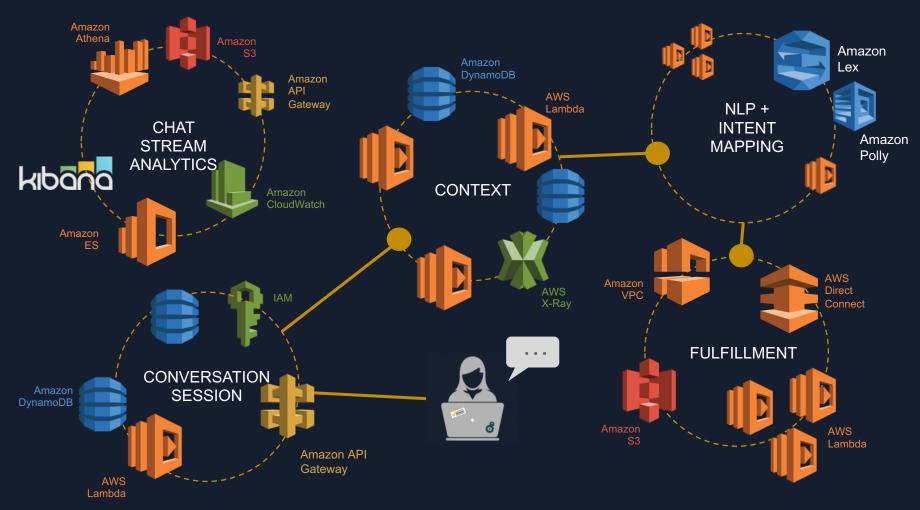






Gillian Armstrong @virtualgill

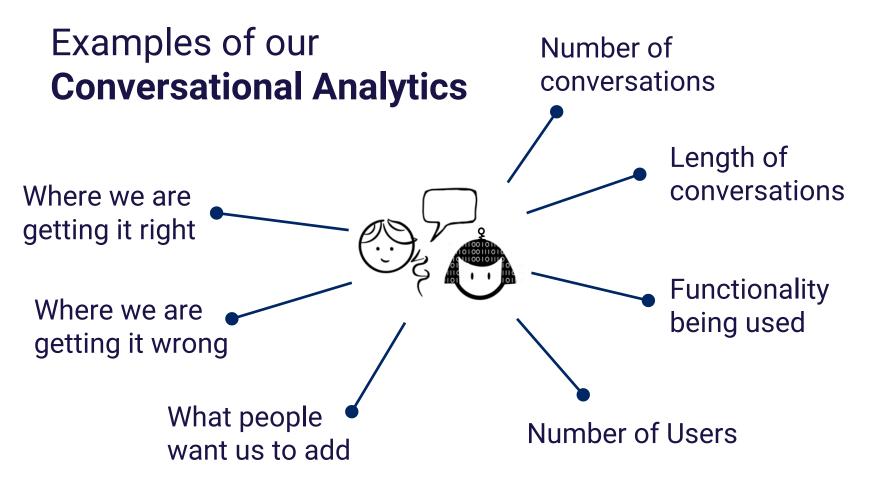




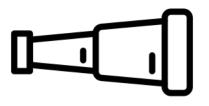
THE DIGITAL ASSISTANT BOT PLATFORM ARCHITECTURE

Sessions & Utterances

kibana 229 Number of Unique Sessions	166 Unique Users		 MobileDeviceManage OutOfDomainIntent CafeMenuIntent NPSIntent ExamplesIntent QuickLinkIntent PasswordIntent HelpDeskIntent WhoIsIntent ThanksIntent BrowserTroubleshoo
Intent Name 🗢	Count ≑		NextOfficeHolidayInt
MobileDeviceManagementIntent	237		 HrServiceCenterGene DefinitionService
OutOfDomainIntent	56		 GreetingIntent Other
CafeMenuIntent	43		- Oulei
NPSIntent	35		
ExamplesIntent	32	Intent hits (short timeline)	
QuickLinkIntent	22	25	📃 🔊 🔵 OutOfDomainIntent
PasswordIntent	19	20	— MobileDeviceManage
HelpDeskIntent	16	¥ 15-	ExamplesIntent
WhoIsIntent	16		CafeMenuIntent CafeMenuIntent HrExecutiveCompInt
ThanksIntent	13		O NextOfficeHolidayInt O NotHelpfulIntent
BrowserTroubleshootingIntent			PasswordIntent
NextOfficeHolidayIntent	9	2018-04-30 01:00 2018-05-02 01:00 2018-05-04 01:00 date.S per hour	 OfficeHolidayListIntent TaxFormIntent
HrServiceCenterGenericIntent	8	uate.s per nour	WholsIntent
DefinitionService	6	Hosting - Completed Conversations By A	
GreetingIntent	6	Hosting - completed conversations by A	
	584	Action 🗢 Count 🗢	



Observability is critical



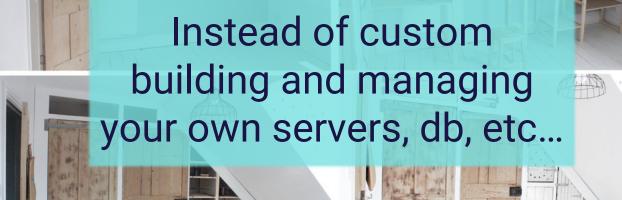
This lets you **quickly learn** where there might be problems with your **system**.



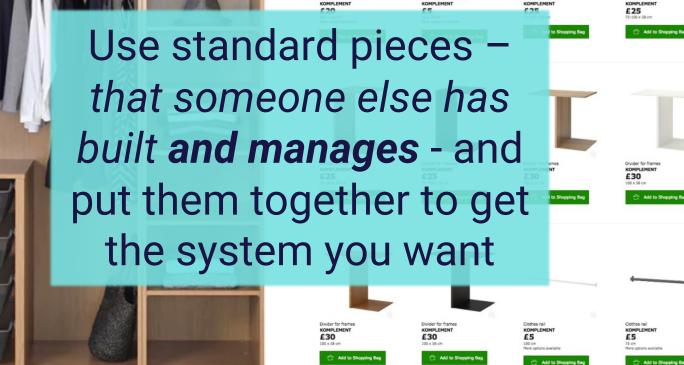


Off-loading Maintenance lets us Fast

Gillian Armstrong @virtualgill



wardrobe by www.hestershandmadehome.com



Glass shell

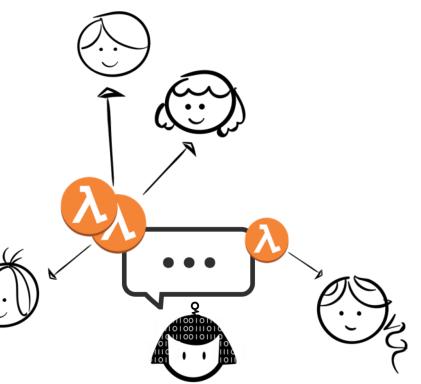
Divider for pull-out trav

Divider for frames

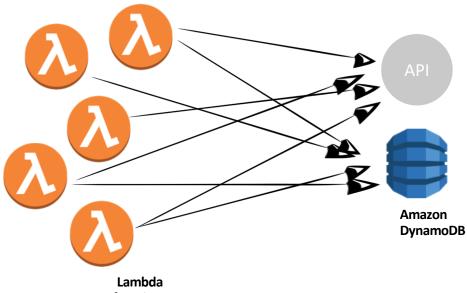




Lambdas are **quick** to develop, **isolate** small pieces of functionality and allow you to **scale** different **parts of the conversation** independently.



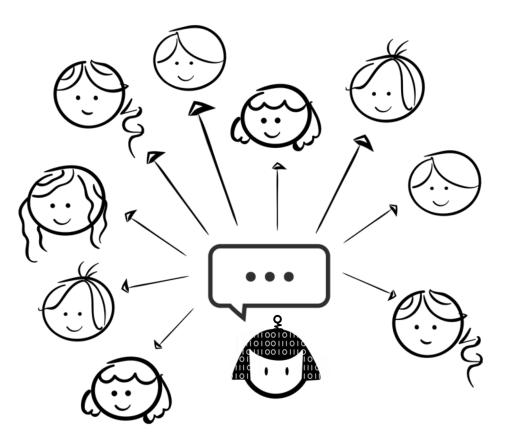
Ensure that everything scales - or at least know your limits.

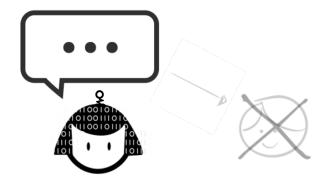


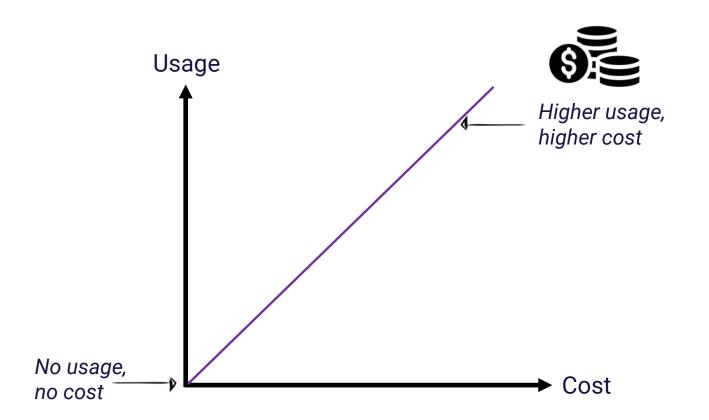
functions

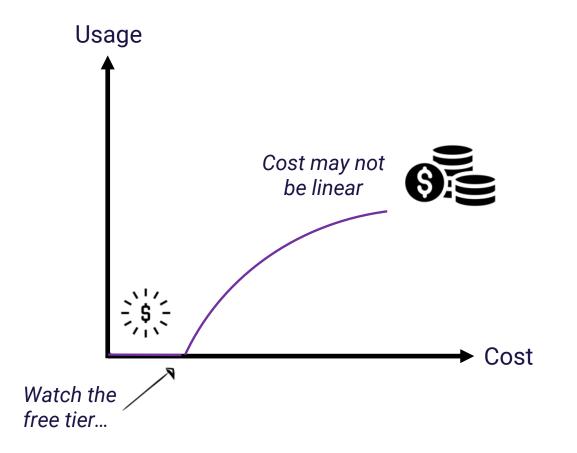
Serverless means you don't pay if **no-one** talks to your bot...

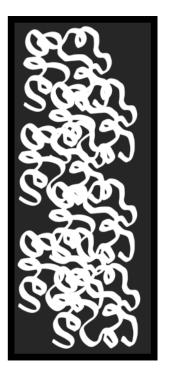
...and lets you scale seamlessly if **everyone** wants to talk to your bot!





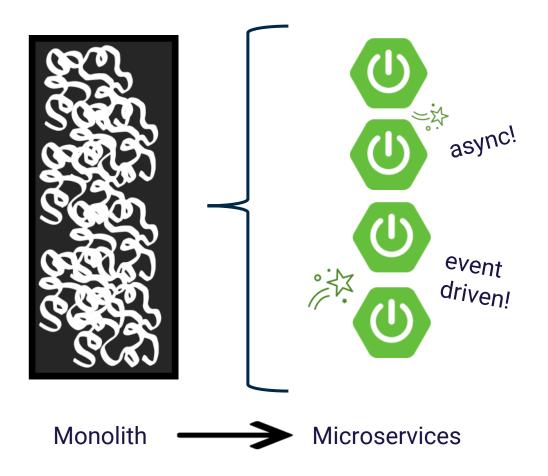


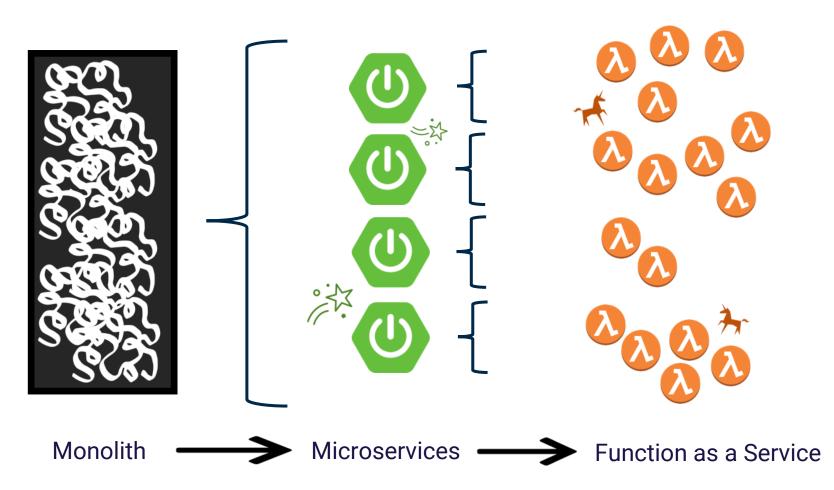




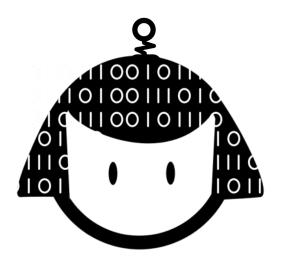
Some notes for those with existing systems

Monolith





REWRITE is not a four-letter word



People are the original conversational interface

Model your design on a **human**, not a website

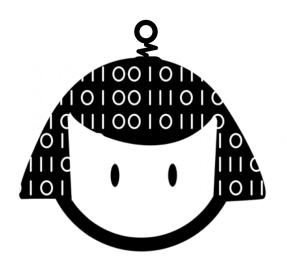


A good human conversation is **synchronous**

You wait for a response

You expect a response immediately

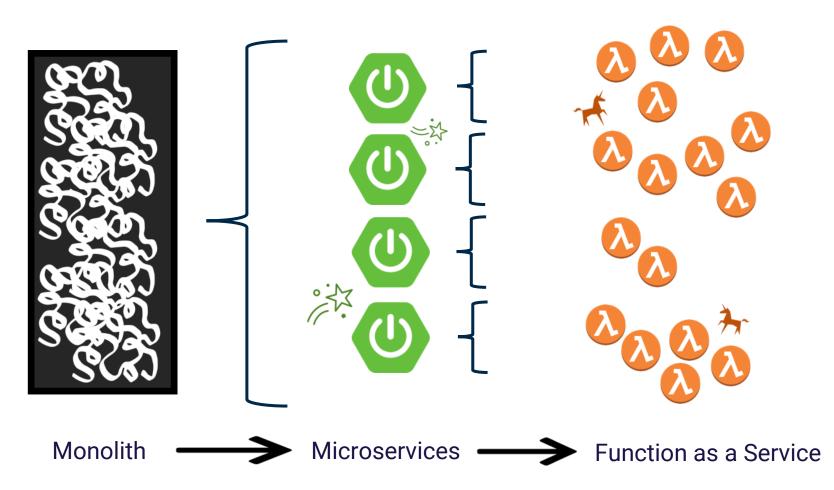




A Conversational Architecture needs to

- Be fast!
- Appear synchronous to the user (there should be a response)

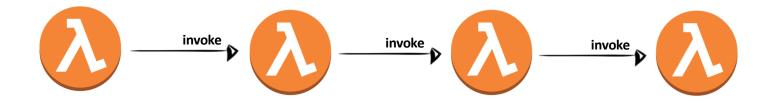
This is particularly critical for a voice-based bot



Do not just **lift and shift** your microservices into FaaS



Do not end up with a collection of **distributed microservices**



Combining the time to **invoke** and the time to **'spin up'** can hit performance and add complexity to your system.

Choose to 'keep warm' only as a last resort.

Look for where you can pull functionality into reusable **libraries** instead



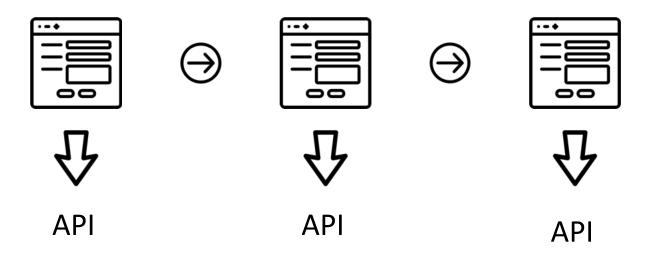
invoke

Shhhhhh..... It's ok to have more than one function in your Function as a Service...

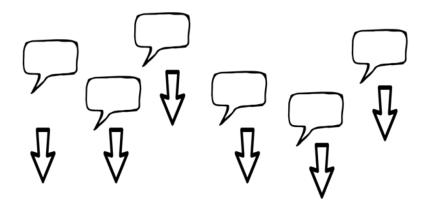
Having said that... **don't** just **lift and shift** your microservices into FaaS



APIs designed for webforms tend to be set up to collect **sets of data** page by page

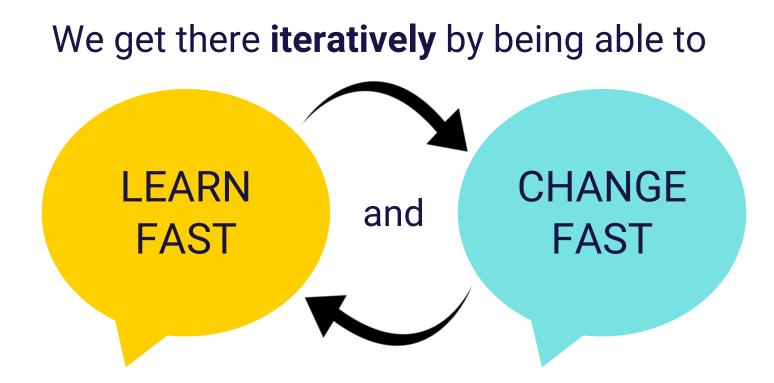


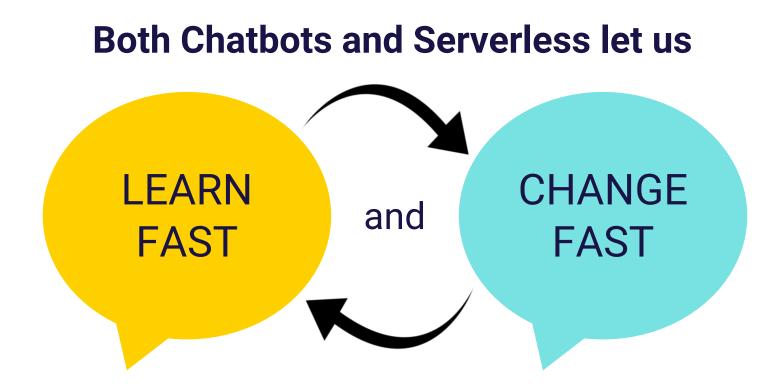
A Conversation is **non-linear** – the user can give **partial information**, or information that would have been **collected later** on your webform.

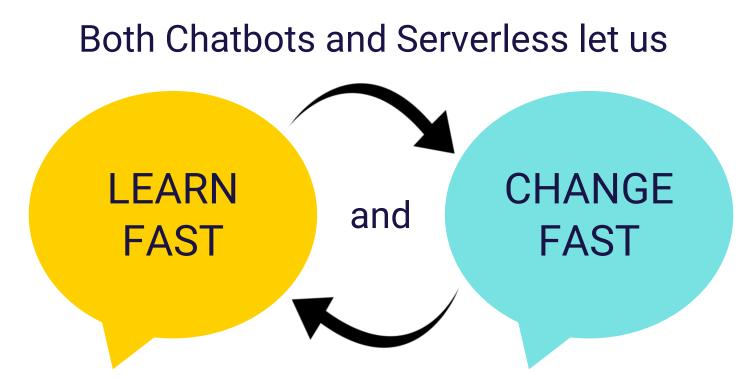


REWRITE is not a four-letter word

Both Serverless and Chatbots require new ways of thinking about your architecture.



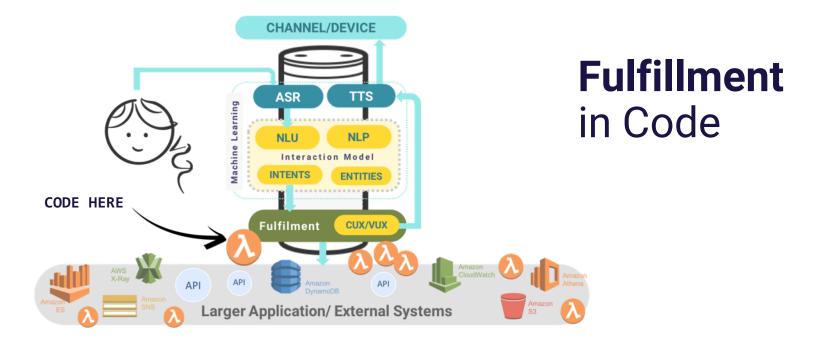


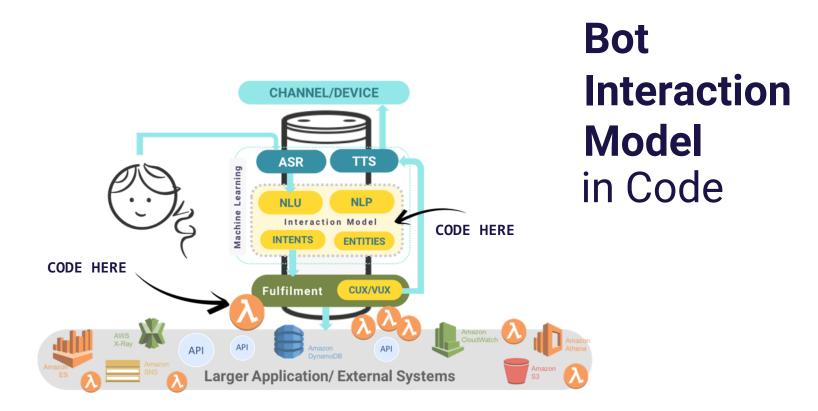


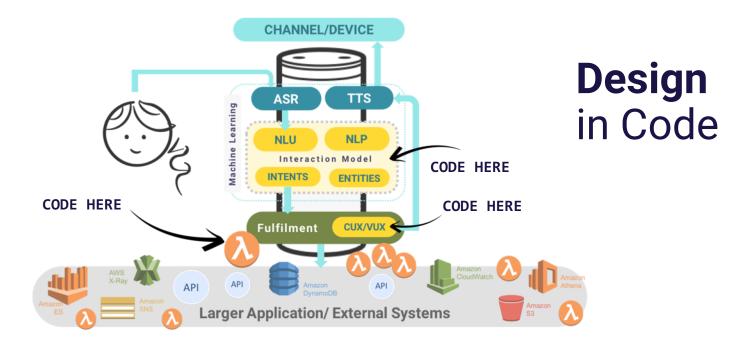
but **Good Supporting Engineering Practices** are needed to do both **with confidence** and **at scale**

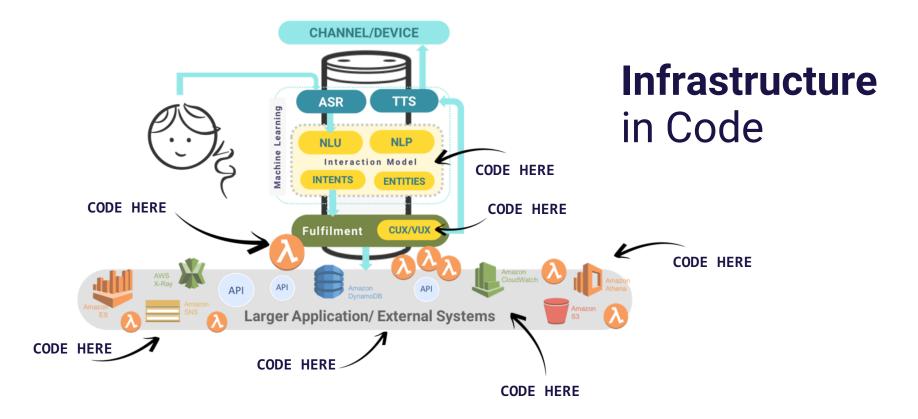
Amazon Lex Console

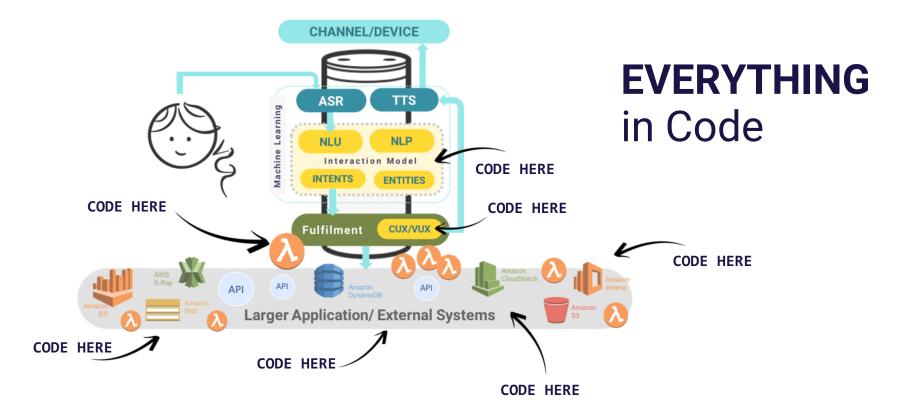
aws	Servic	es v	Resource Groups ~	*	Д [●] Im-federat	e-saml-developer/a0 + N. Virginia + Support +
DemoAssi	stantBot La	itest 🛩			Build Publish Ø	> Test bot (Latest)
Editor	Settings	Channe	als Monitoring			
Intents	0		demo_BookMeeting	gIntent 2 - Edit		
	kMeetingInte		- Sample utterances	0		
-	ndarAvailability ericHelpIntent		please book a meeting at	(time)		You're now ready for testing. Type an utterance
demo_GoodbyeIntent demo_GreetingIntent			i want to book a meeting			below to begin conversation with your chatbot.
demo_Orde	rDrink		i need a meeting booked			
Slot type	s		at a letter even	ha Wat	Tutorials	told mo
No slots cre	ated	V	can you create a calendar		i utoriais	
-			can rou booir ma a meeti	ng with (person) at (time) on (den)		Clear chat history
Error Handl	ing		iildin	a a cha	that looke	dlika
	L		book time in my calendar	iy a che	tbot looke	u iike
			book me a meeting (time			Inspect response
			book me a meeting with			
			-	(person) at (time) on (date)		
			book me a meeting on (o	iate) at (time)		
			book me a meeting on (o	late)		When you chat with your bot, you can see the
			book me a meeting			fulfillment state of your intent and the response here.
			book a meeting on {date			
			book a meeting with {per	rson) at <mark>(time)</mark>		
			book a meeting with (per	rson) (date)		
			book a meeting with [per	rson}		





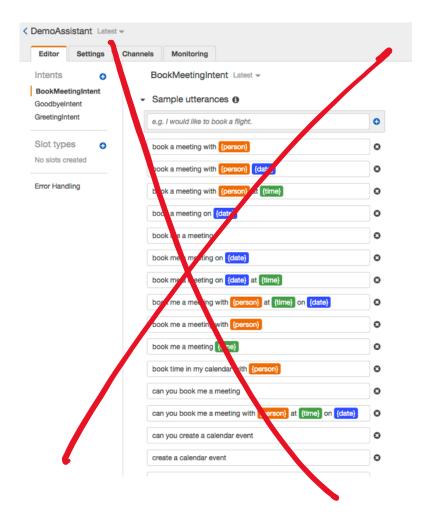




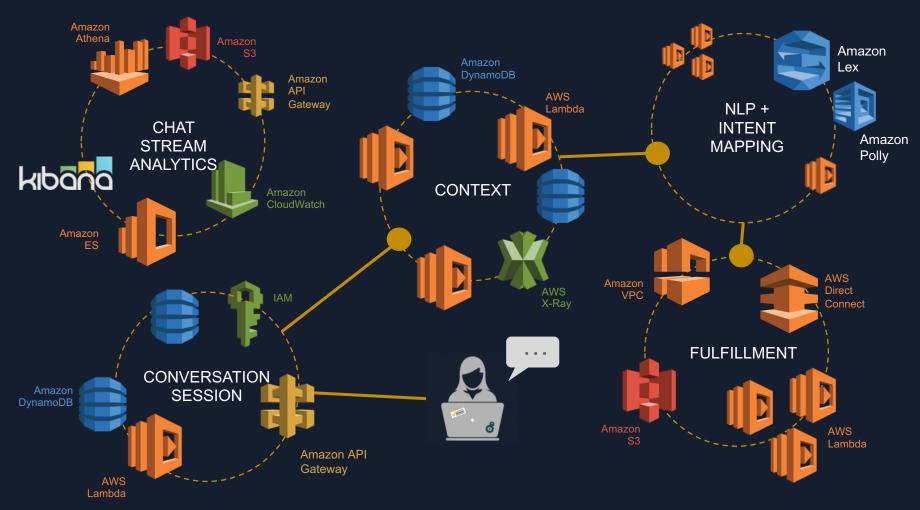


- Testable
- Traceable
- Repeatable
- Observable

EVERYTHING in Code



```
},
"sampleUtterances": [
  "book a meeting with {person}",
  "book a meeting with {person} {date}",
  "book a meeting with {person} at {time}",
  "book a meeting on {date}",
  "book me a meeting".
  "book me a meeting on {date}",
  "book me a meeting on {date} at {time}",
  "book me a meeting with {person} at {time} on {date}",
  "book me a meeting with {person}",
  "book me a meeting {time}",
  "book time in my calendar with {person}".
  "can you book me a meeting",
  "can you book me a meeting with {person} at {time} on {date}",
  "can you create a calendar event",
  "create a calendar event",
  "i need a meeting booked",
  "i want to book a meeting".
  "please book a meeting at {time}"
1,
"slots": [
  -{
    "description": "The person the meeting is with",
    "name": "person",
    "priority": 1,
    "slotConstraint": "Required",
    "slotType": "AMAZON.Person",
    "valueElicitationPrompt": {
      "maxAttempts": 5,
      "messages": [
        {
          "content": "Who would you like to meet with?",
          "contentType": "PlainText"
      1
    3
  },
    "description": "The date of the meeting".
    "name": "date",
    "priority": 2,
    "slotConstraint": "Required",
    "slotType": "AMAZON.DATE",
    "valueElicitationPrompt": {
      "maxAttempts": 2.
      "messages": [
        {
          "content": "What date would you like the meeting on?",
          "contentType": "PlainText"
        },
          "content": "Sure! What date would you like to meet on?",
          "contentType": "PlainText"
        3
     1
    }
  },
    "description": "The time of the meeting".
    "name": "time",
    "priority": 3,
    "slotConstraint": "Required",
    "slotType": "AMAZON.TIME"
```



THE DIGITAL ASSISTANT BOT PLATFORM ARCHITECTURE

Fully Validated and Repeatable Bot Build and Deploy to any AWS Account



Everything written in CFN, with JSON for API calls to Amazon Lex Model Building API as part of deploy

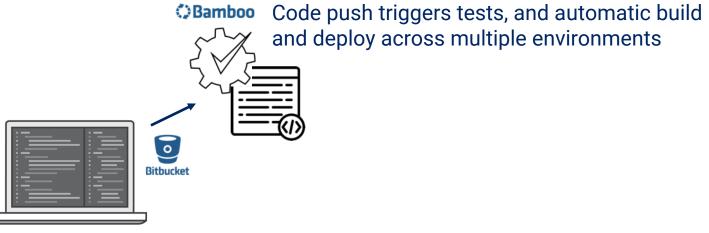
Fully Validated and Repeatable Bot Build and Deploy to any AWS Account



Everything written in CFN, with JSON for API calls to Amazon Lex Model Building API as part of deploy Unit Test
Lambdas
Libraries
Static analysis
Cloud Formation
Code
JSON for API calls

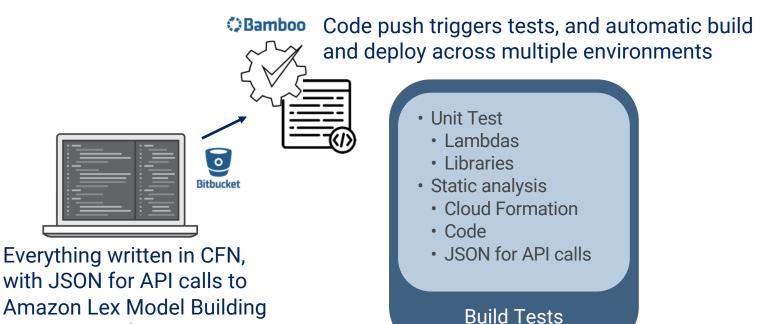
Local and Pre-Commit Hook Tests

Fully Validated and Repeatable Bot Build and Deploy to any AWS Account

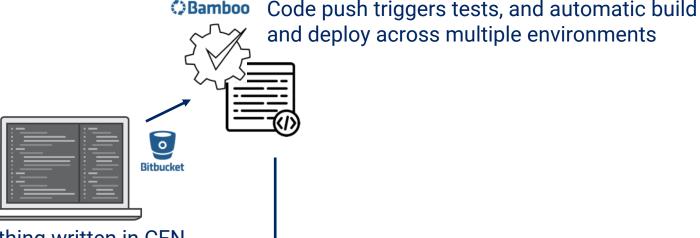


Everything written in CFN, with JSON for API calls to Amazon Lex Model Building API as part of deploy

API as part of deploy



Fully Validated and Repeatable Bot Build and Deploy to any AWS Account



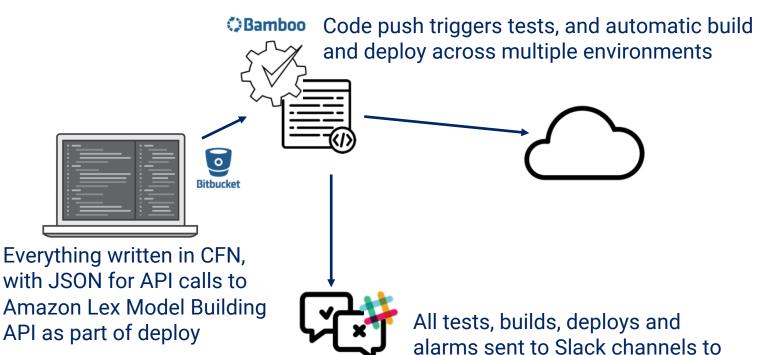
Everything written in CFN, with JSON for API calls to Amazon Lex Model Building API as part of deploy



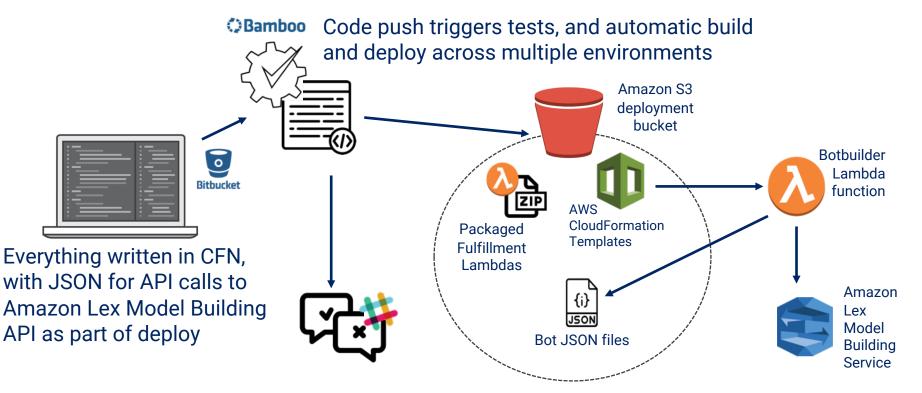
All tests, builds, deploys and alarms sent to Slack channels to alert developers immediately

Fully Validated and Repeatable Bot Build and Deploy to any AWS Account

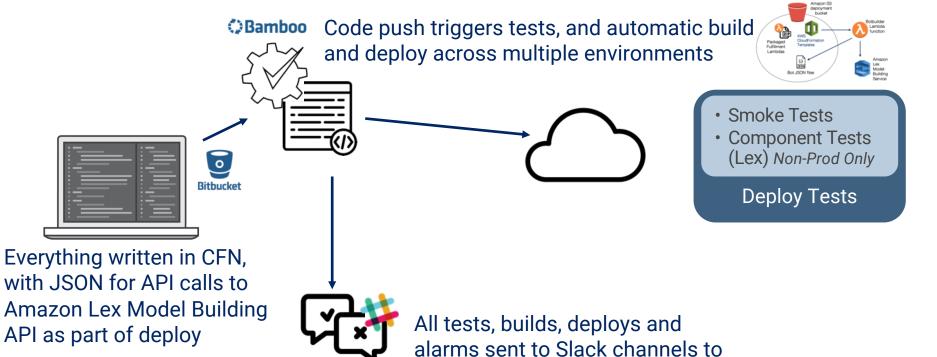
alert developers immediately



Gillian Armstrong @virtualgill

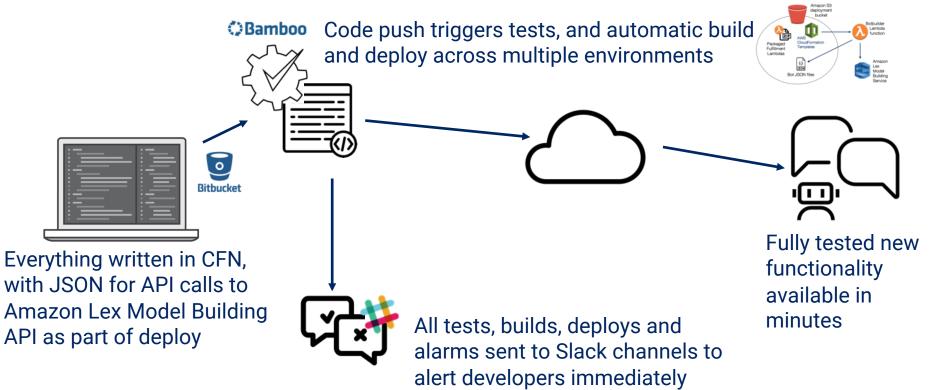


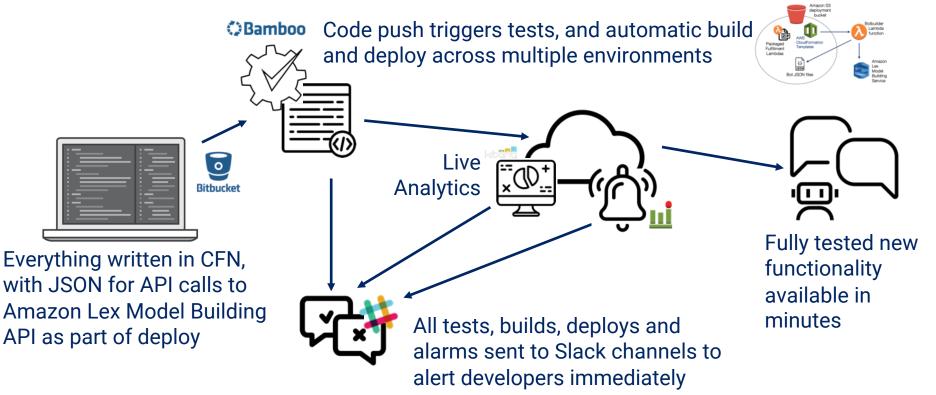
Fully Validated and Repeatable Bot Build and Deploy to any AWS Account

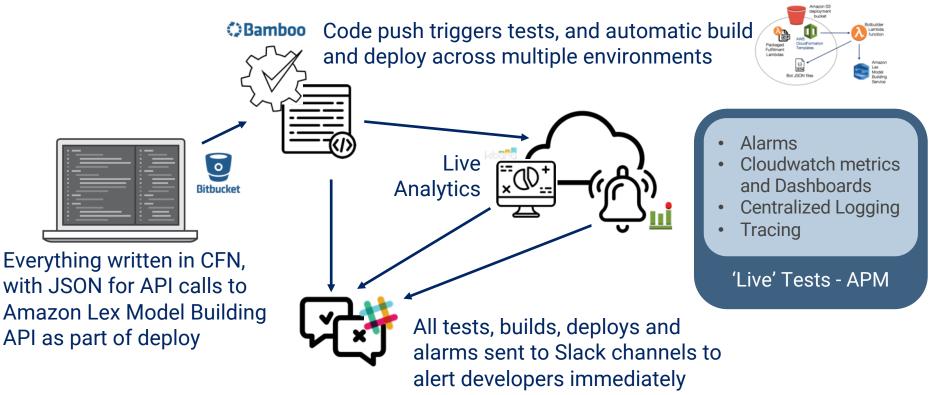


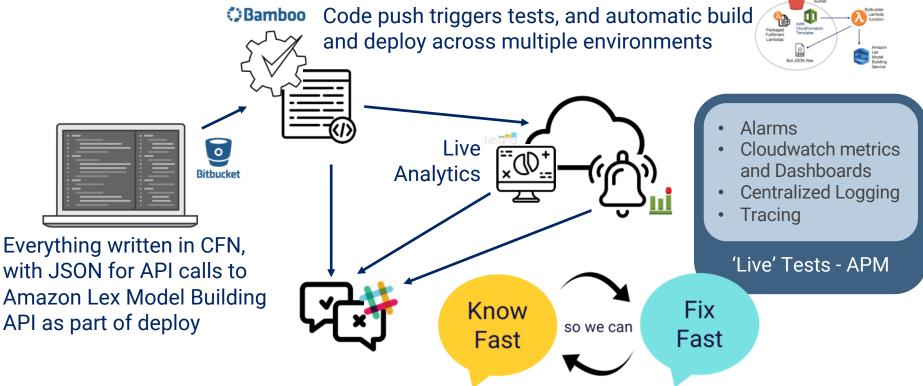
alert developers immediately

Gillian Armstrong @virtualgill









Monitor, no, seriously, monitor

If your **chatbot stopped working**, would you **know**? Don't just wait for user complaints!

Know what isn't working and **fix** it, and keep **adding more** of what is!

Conversational Analytics are vital



Work out what **metrics** you need to track to know how your chatbot is performing.

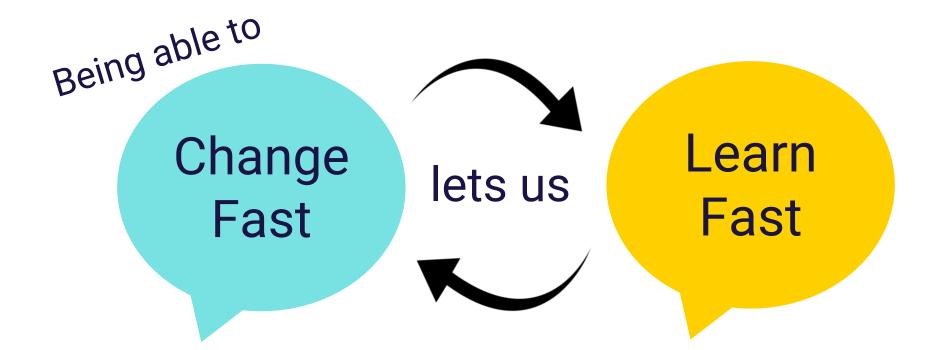
Know what isn't working and **fix** it, and keep **adding more** of what is!

Getting Feedback is vital

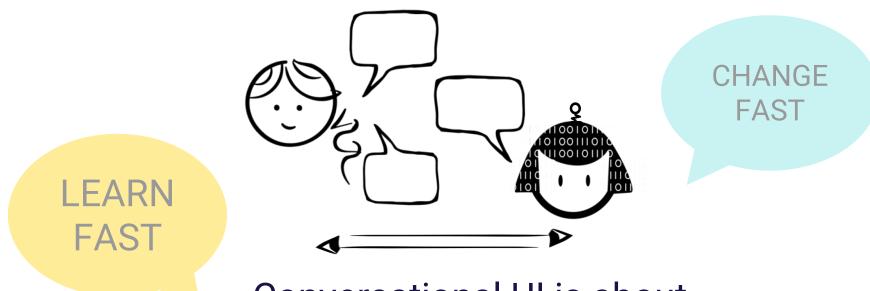


You won't get it right first time – **listen** to your users.

Know what isn't working and **fix** it, and keep **adding more** of what is!



Gillian Armstrong @virtualgill



Conversational UI is about Listening to what the user wants to do...

...and that's the secret to a great chatbot!



- Model your Chatbot UI and Architecture based on observing Humans and real Conversation. Break out of old mindsets!
- New tech still needs the good engineering practices you already know

Got questions? Let's chat!



Appendix – Additional Information on Conversational Design

Starting out:

https://hackernoon.com/new-to-conversational-design-start-here-7f2f3a1b81bb

Error Handling in Chatbot Series:

Part 1 — Voice Recognition Errors: The one where we end up shouting at a computer <u>https://chatbotsmagazine.com/difficult-conversations-1-voice-recognition-error-handling-74d93056ddce</u>

Part 2 – Conversational Errors: The one where it gets really interesting https://chatbotsmagazine.com/helping-your-baby-bot-learn-to-chat-like-a-grown-up-bot-99f5170f1c55

Part 3 – Technical Faults: The one we want to pretend will never happen (or 'Help! My chatbot has fallen over and can't get up') https://chatbotsmagazine.com/good-ux-when-your-chatbot-is-having-a-very-bad-day-e4f22885d7fb

You can find my other Blogs at <u>https://medium.com/@virtualgill</u> Come talk to me on Twitter – I hang out there a lot @virtualgill For my other talks and ways to get in contact with me head to <u>http://virtualgill.io</u>