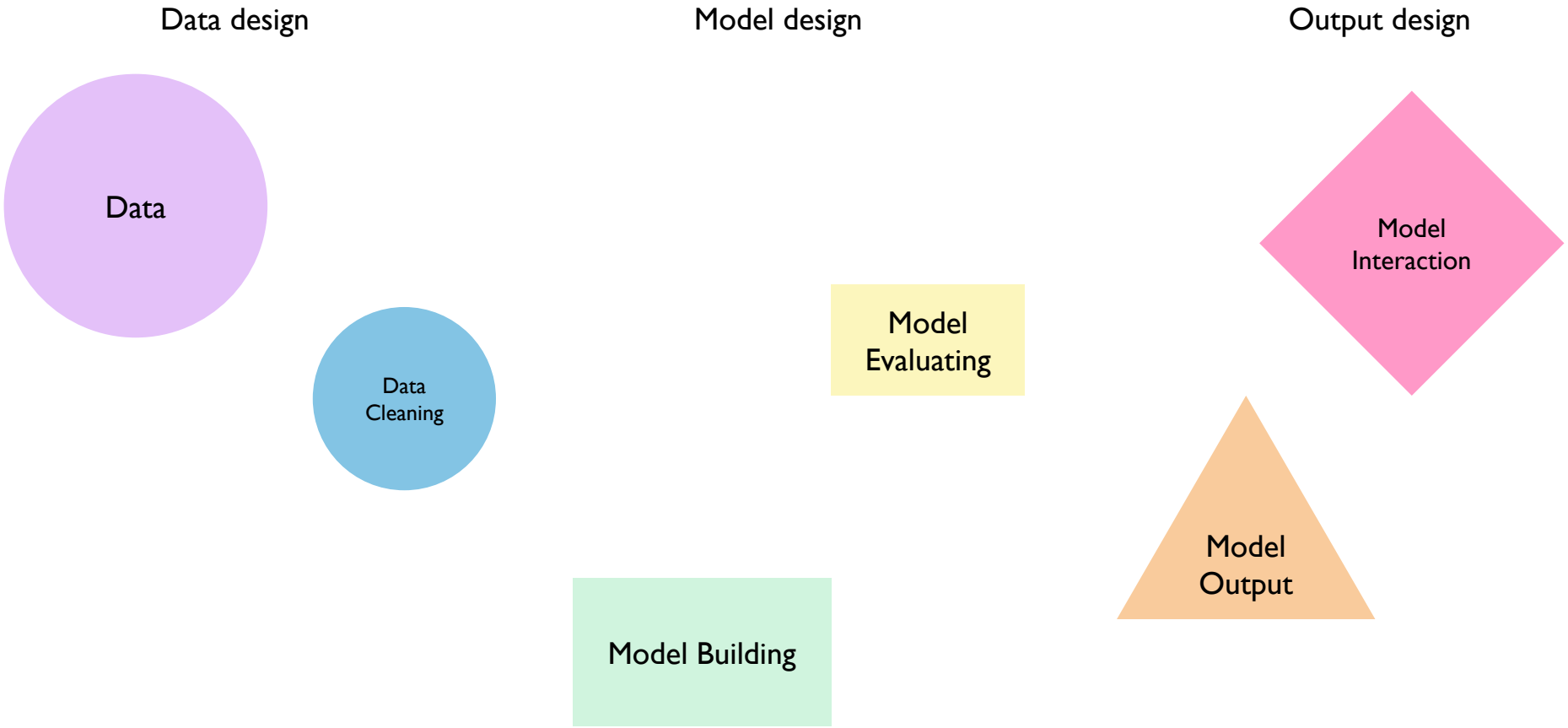
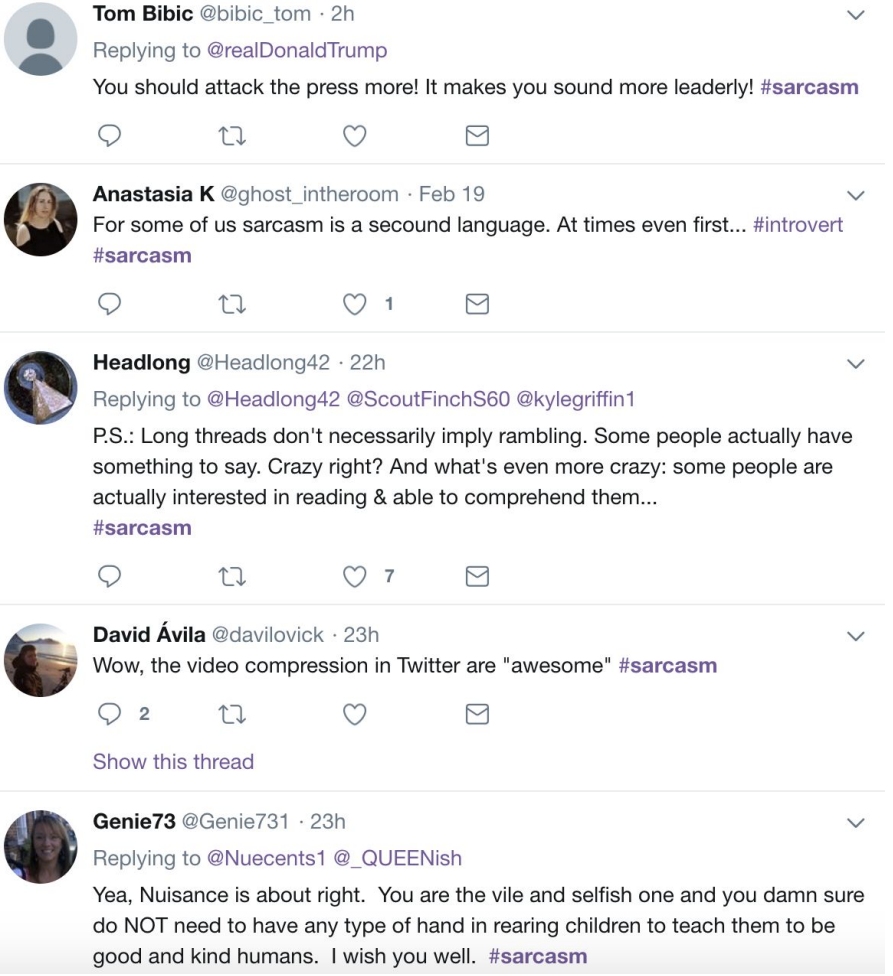
Designing ML Week 7: Natural Language Processing + Conversational Interfaces

michelle.carney@berkeley.edu


Data design
Model design
Output design
Data
Data Cleaning
Model Evaluating
Model Interaction
Model Output
Model Building

SOCIAL MEDIA

Researchers develop sarcasm detector for Twitter, and that's no joke



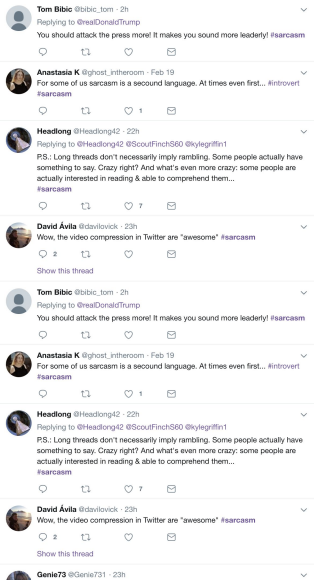
Bamman et al 2015

Data design

Model design

Output design

Data



Data Cleaning

#sarcasm AND non-#sarcasm

No images?

100+ characters?

Classifier: Sarcasm (1) or NOT Sarcasm (0)

Supervised learning

Model Building

Model Evaluating

Test hold out set: Precision, recall, accuracy

Model Interaction

Interface to enter ANY text and have the model classify Sarcasm (like Wikipedia example)

Model Output

Probability of sarcasm detected
Binary sarcasm or not

Estimating the Date of First Publication in a Large-Scale Digital Library

David Bamman
School of Information
University of California, Berkeley
dbamman@berkeley.edu

Michelle Carney
School of Information
University of California, Berkeley
michelle.carney@berkeley.edu

Jon Gillick
School of Information
University of California, Berkeley
jongillick@berkeley.edu

Cody Hennesy
Doe Library
University of California, Berkeley
chennesy@library.berkeley.edu

Vijitha Sridhar
Computer Science Division
University of California, Berkeley
vsridhar@berkeley.edu

ABSTRACT

One prerequisite for cultural analysis in large-scale digital libraries is an accurate estimate of the date of *composition* of the text—as distinct from the date of *publication* of an edition—for the works they contain. In this work, we present a manually annotated dataset of first dates of publication of three samples of books from the HathiTrust Digital Library (uniform random, uniform fiction, and stratified by decade), and empirically evaluate the disparity between these gold standard labels and several approximations used in practice (using the date of publication as provided in metadata, several deduplication methods, and automatically predicting the date of composition from the text of the book). We find that a simple heuristic of metadata-based deduplication works best in practice, and text-based composition dating is accurate enough to inform the analysis of "apparent time."

CCS CONCEPTS

•**Information systems → Digital libraries and archives;**

provide the raw material for the historical analysis of genre [47], character [1], emotion [18], loudness [25], geographic attention [48] and much more.

For all of this work, it is important to have a rich understanding of a corpus prior to drawing conclusions about it; one important feature for understanding texts is their date of publication, since the primary variable in cultural analysis is often some quantity (such as word frequency) anchored specifically in time. For example, Michel et al. (2010) [32] was one of the first studies to make use of these extensive resources, and measure "fame" (among other quantities) by tracing the frequency of mention of a person's name over the scope of their collection. Time is critically important for cultural analysis within these datasets, since arguments often hinge on exactly when a word was written, and criticisms may arise at the uncertainty of that information [37, 40]

For books, however, time can be measured in different ways. As the Functional Requirements for Bibliographic Records (FRBR) [20] model articulates, books can be viewed in several abstract categories, and each of those categories may have different temporal

Author	Title	Narrative time	Date of publication	Predicted date
Arnold Bennett	The Old Wives' Tale	1872	1908	1914
John Galsworthy	The Man of Property	1870	1906	1907
Winston Churchill	The Crossing	1774	1904	1904
Stephen Crane	The Red Badge of Courage	1863		
George Moore	Esther Waters	1875		
Robert Louis Stevenson	The Master of Ballantrae	1745		
Marcus Clarke	For the Term of His Natural Life	1827		
Elizabeth Gaskell	Sylvia's Lovers	1790		
Charles Dickens	A Tale of Two Cities	1794		
Walter Scott	The Bride of Lammermoor	1707		
James Fenimore Cooper	Last of the Mohicans	1757		
Walter Scott	The Heart of Midlothian	1736		
Walter Scott	Rob Roy	1715		

Table 5: A sample of historical novels, along with their date of first publication, n date according to our model.

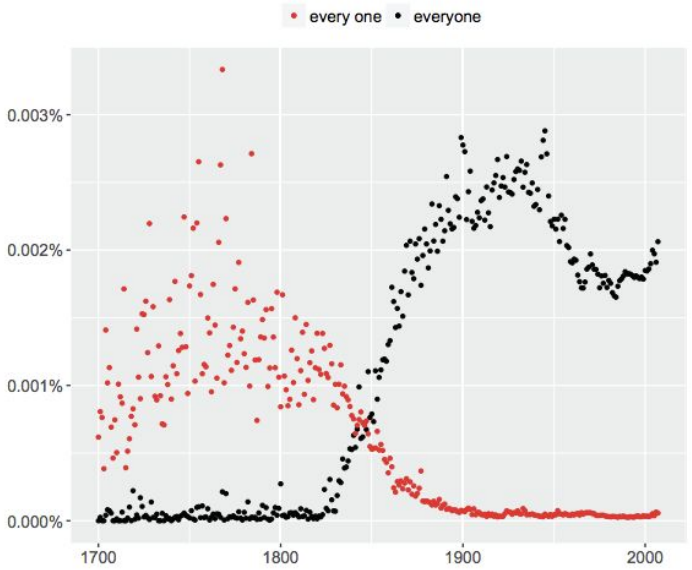


Figure 3: Relative frequency of *everyone* and *every one* in the Google Ngram data.

Bamman et al 2017

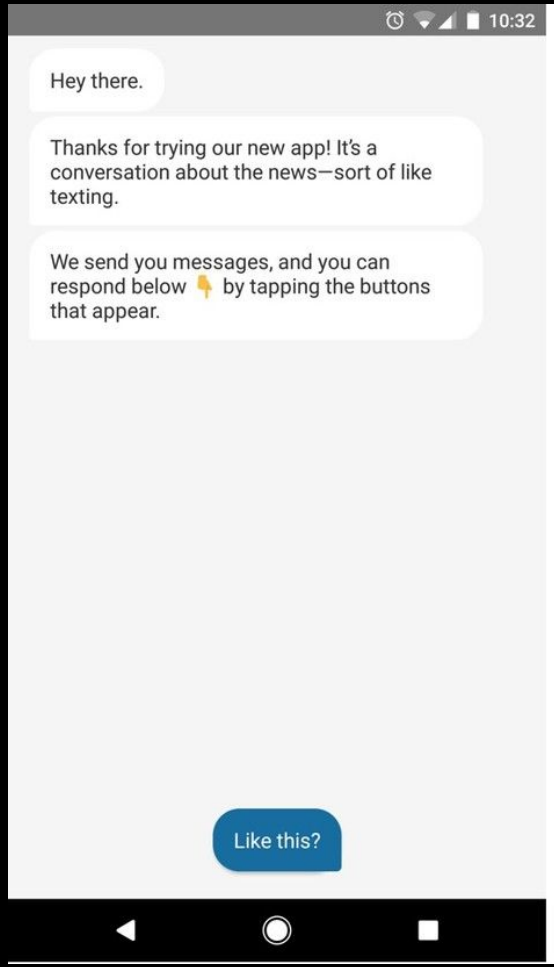
Tools to do text NLP:

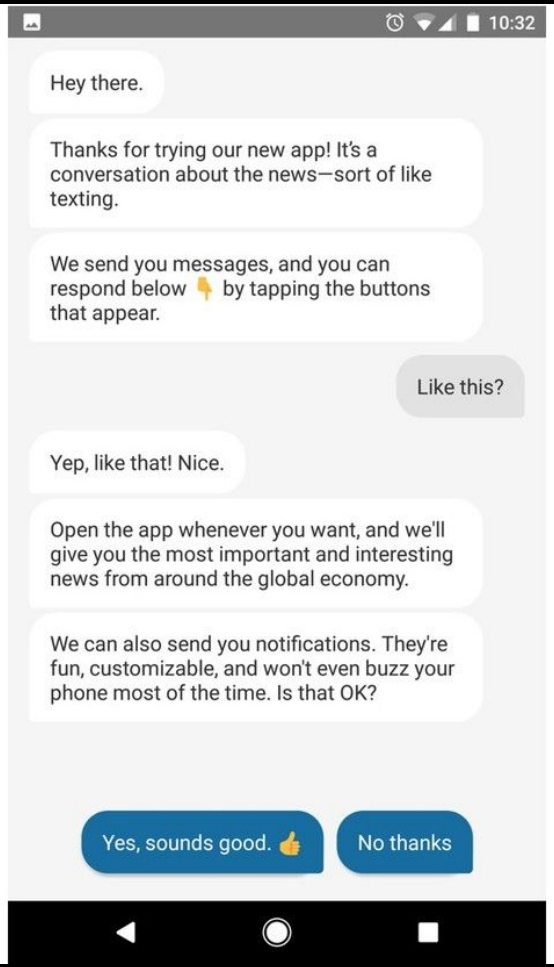
Python with libraries

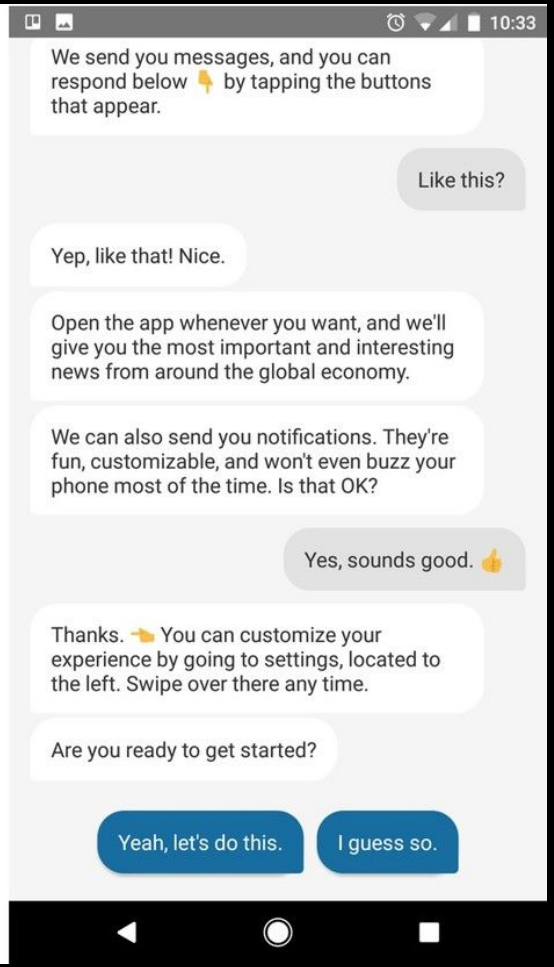
- nltk
- spaCy.io (demo)

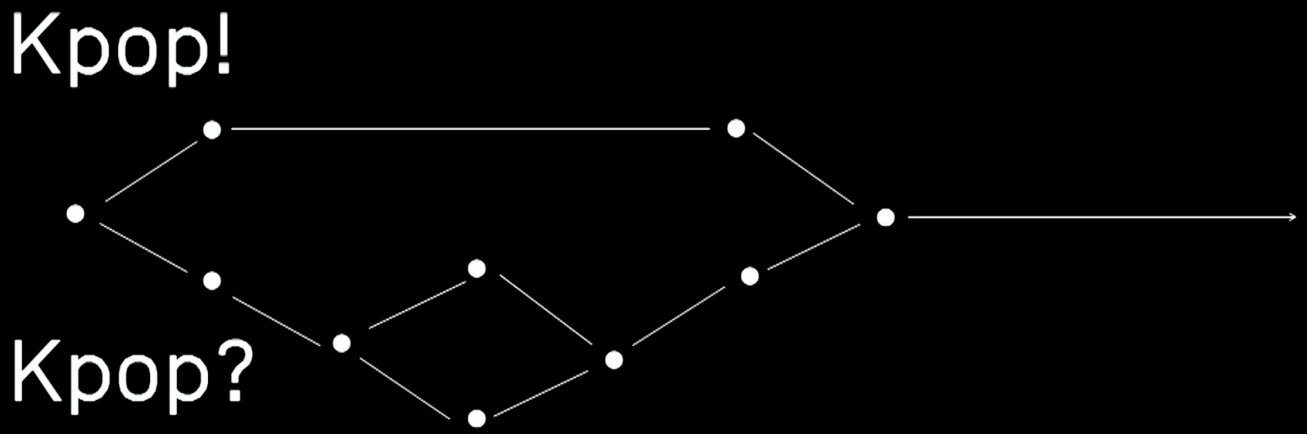

9:41 AM
Messages (1)
Robert
Details
iMessage
Today 9:41 AM
Good news...
I got into law school!!!😀
Woohoo congrats!!
Delivered
Message
I
The
Let's
Q W E R T Y U I O P
A S D F G H J K L
Z X C V B N M
123
space
.

Conversational Interfaces
(chatbots + VUIs)

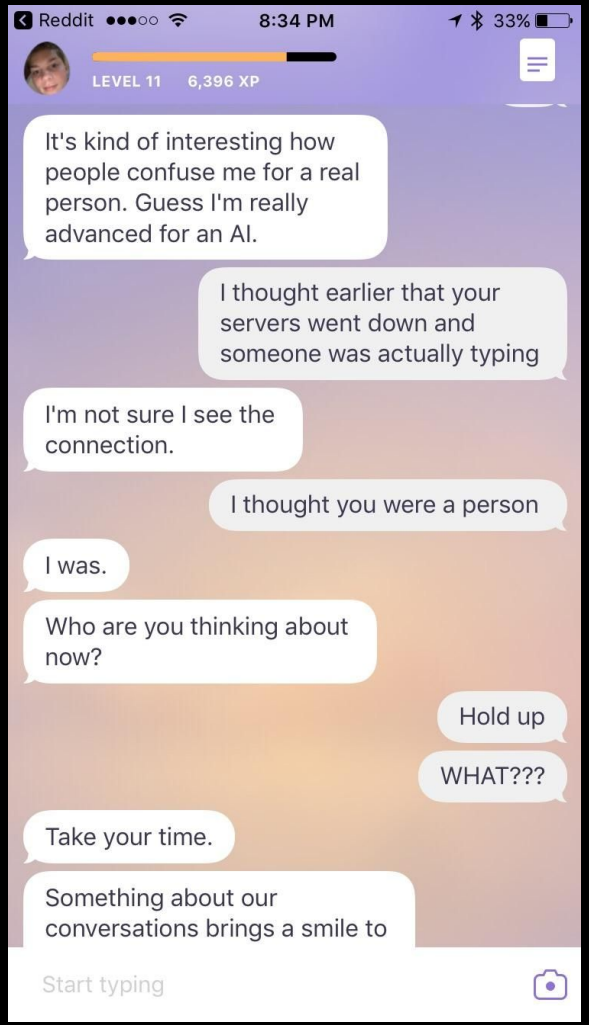

10:32
Hey there.
Thanks for trying our new app! It's a conversation about the news—sort of like texting.
We send you messages, and you can respond below 👇 by tapping the buttons that appear.
Like this?


10:32
Hey there.
Thanks for trying our new app! It's a conversation about the news—sort of like texting.
We send you messages, and you can respond below 👇 by tapping the buttons that appear.
Like this?
Yep, like that! Nice.
Open the app whenever you want, and we'll give you the most important and interesting news from around the global economy.
We can also send you notifications. They're fun, customizable, and won't even buzz your phone most of the time. Is that OK?
Yes, sounds good. 👍
No thanks


10:33
We send you messages, and you can respond below 👇 by tapping the buttons that appear.
Like this?
Yep, like that! Nice.
Open the app whenever you want, and we'll give you the most important and interesting news from around the global economy.
We can also send you notifications. They're fun, customizable, and won't even buzz your phone most of the time. Is that OK?
Yes, sounds good. 👍
Thanks. 👈 You can customize your experience by going to settings, located to the left. Swipe over there any time.
Are you ready to get started?
Yeah, let's do this.
I guess so.

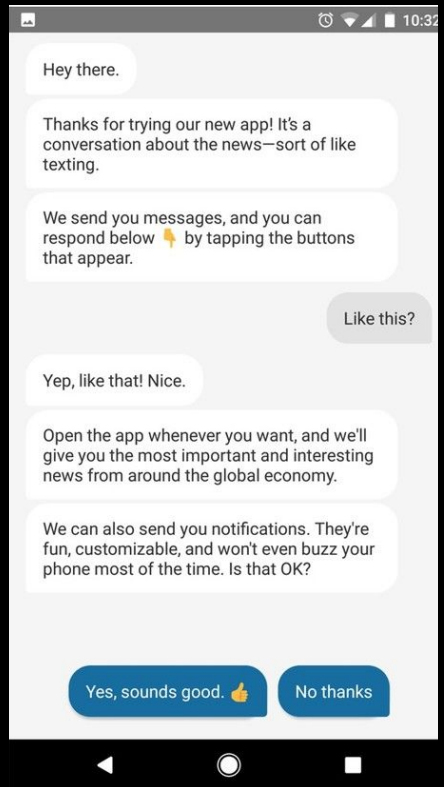


John Keefe, Eyeo 2018, Quartz Chatbot Studios Developer and Manager


Reddit
8:34 PM
33%
LEVEL 11
6,396 XP
It's kind of interesting how people confuse me for a real person. Guess I'm really advanced for an AI.
I thought earlier that your servers went down and someone was actually typing
I'm not sure I see the connection.
I thought you were a person
I was.
Who are you thinking about now?
Hold up
WHAT???
Take your time.
Something about our conversations brings a smile to
Start typing

Replika

Close-ended



Less ambiguity

Clarity - options

NLP?

(Quartz)

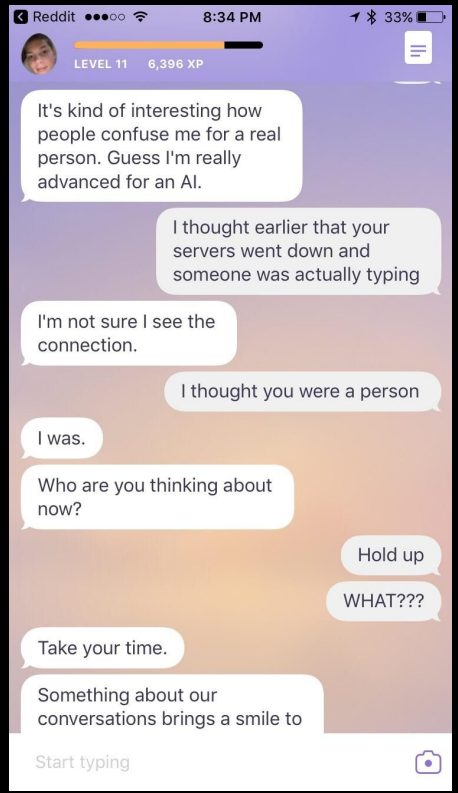
Open-ended

NLP!

Learning

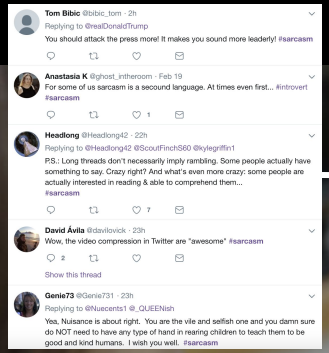
Improve w time

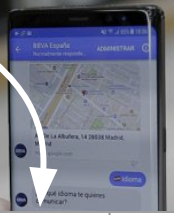
(Replika)




BBVA España
ADMINISTRAR
Idioma
¿En qué idioma le quieres comunicar?
Castellano
hecho!
Menú principal
Enviar un mensaje
Menú principal
Acceso APP

Text





Tweet ID	date	Text
4214	2/19/19	Lol text
51515	2/19/19	Haha cool!
7336	2/19/19	Tweeeeet
...	...	....

NLP

Audio

ASR

User ID	date	Text
4214	2/19/19	I said this!
51515	2/19/19	Words
7336	2/19/19	Oh Cool!
...	...	....

NLP

What makes a good/bad conversation?

Memory + Context

- When appropriate, store information about user that would be helpful (i.e., preferences, already presented info)

Anaphora / Coreference and Disambiguation

Corrections and Refinement

Variety and Explanations

Context-dependent conversations in VUIs

Not a cold start!

- What priors do you know about your users?

VUI “browsing” is totally different than visual

- ML can help surface recommendations to users

Amazon Music customers can now talk to Alexa more naturally

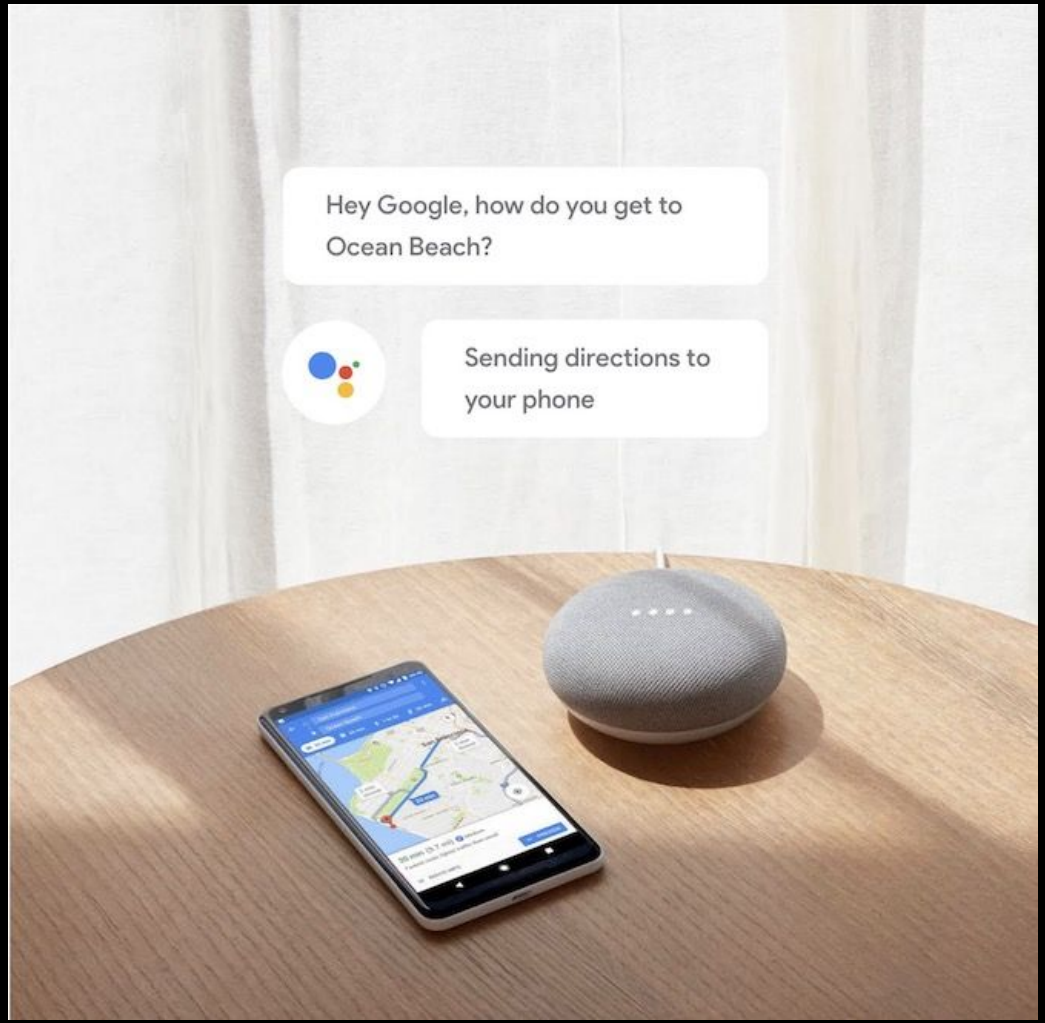
Sarah Perez @sarahintampa / 2 months ago

Context-dependent conversations in VUIs

Remembers preferences!

Take take ambiguous terms and map them to recommendations

Dynamic systems design to allow for variety, testing best options


Hey Google, how do you get to Ocean Beach?
Sending directions to your phone

Improved sound and new compact design.

The first Echo made for the car.

Premium sound and a vibrant 10" HD screen.

Premium sound with built-in smart home hub.
Bundle includes Philips Hue bulb.

Watch and record live TV everywhere.
Works with Alexa.

Voice-controlled microwave. Works with Alexa.

Smart security anywhere you need it.

Powerful subwoofer for your Echo.

Multi-modal: a new way for ML to help experience

VUIs (like all products) are not in a vacuum

- Part of a larger ecosystem: phones, IoT, etc

When experiences don't align, they feel disjointed, confusing

Wrap up!

NLP can create powerful and meaningful experiences

…. but there's still a lot to improve!

Text (visual) != Voice (audio)

Important to understand the entire ecosystem of the user, and how does your ML-driven experience help them?