The Rise of Conversational AI

David Low
Co-founder / CDS

PAND.AI

Bio

- **Research**

  Urban Mobility | Social Media

- **Public Service**

  GovTech(IDA) Data Science Division

- **Teach**

  Deep Learning Masterclass at National University of Singapore (NUS)

- **Startup**

 Conversational AI | Deep Natural Language Processing

Overview

- Why Conversational AI?

- What is NLP?

- Current state of Conversational AI

- Why is NLP difficult?

- Open Source Framework

- Recent advancements

- What have we learnt?
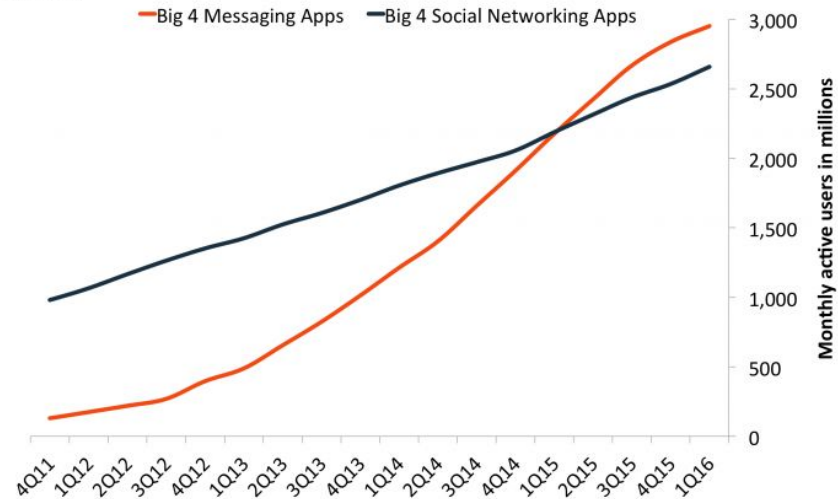
- Demo of Question Answering System

Rising trend of mobile (& messaging) usage

Mobile App Time Spent Grows 69% Year-Over-Year

394%

Average 69%

43%

31%

25%

Games

-4%

Messaging & Social

Business & Finance

Shopping

Sports
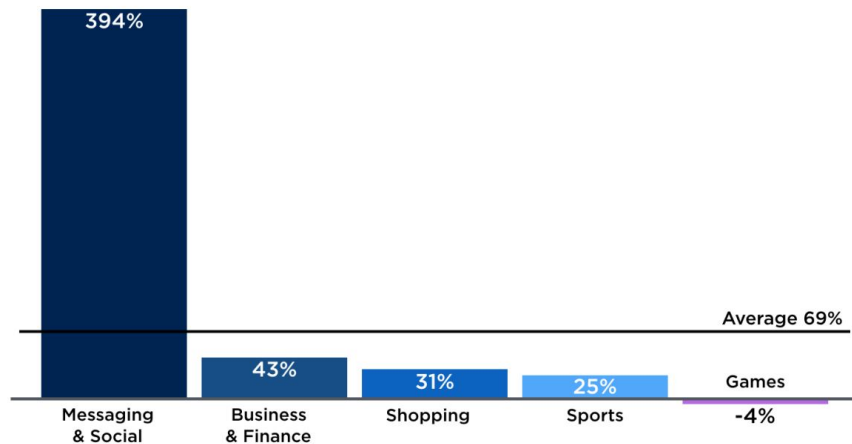
Messaging Apps Have Surpassed Social Networks

Global monthly active users for the top 4 messaging apps and social networks, In millions

— Big 4 Messaging Apps   — Big 4 Social Networking Apps

Monthly active users in millions

3,000
2,500
2,000
1,500
1,000
500
0

4Q11 1Q12 2Q12 3Q12 4Q12 1Q13 2Q13 3Q13 4Q13 1Q14 2Q14 3Q14 4Q14 1Q15 2Q15 3Q15 4Q15 1Q16
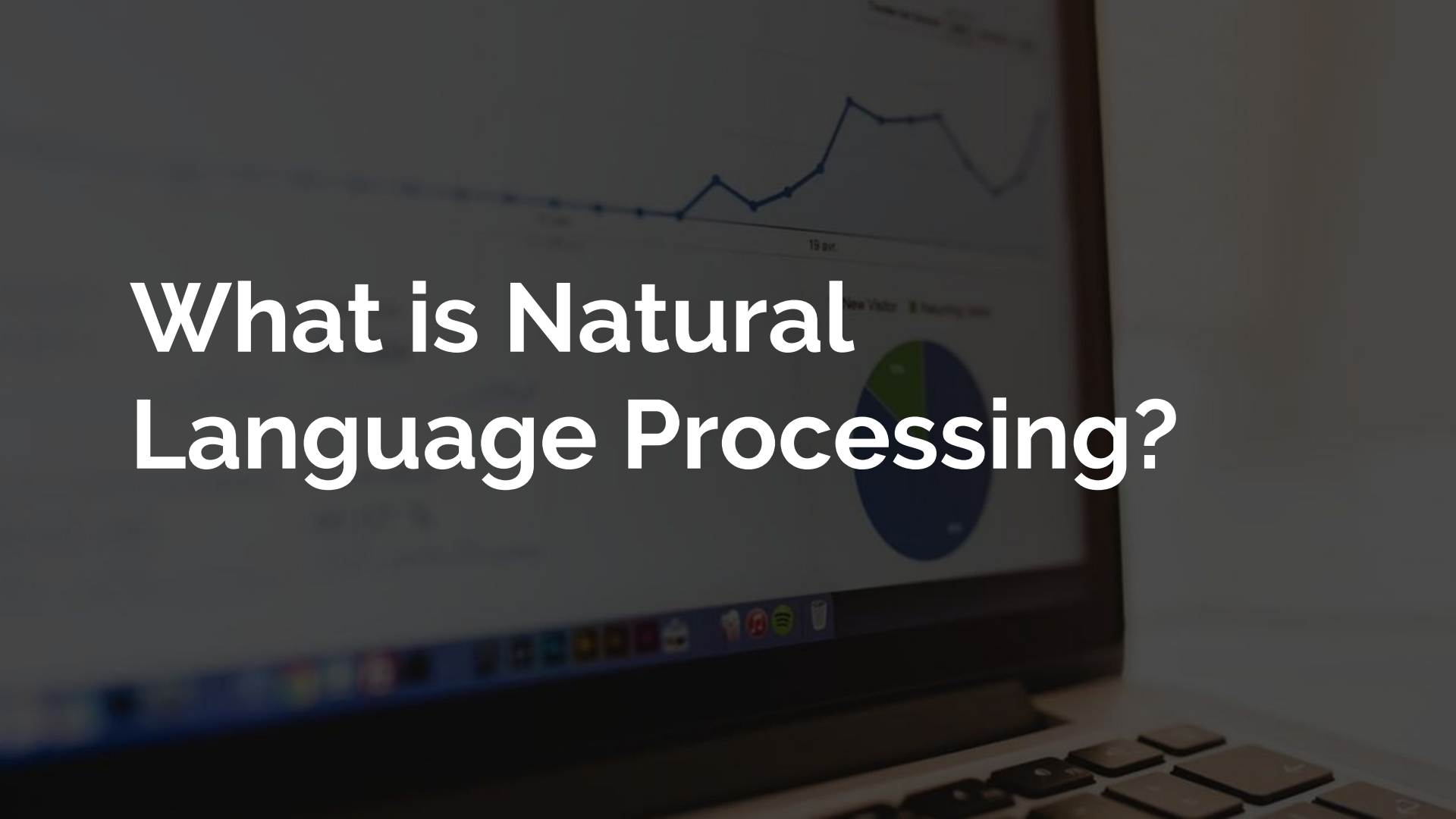
Note: Big 4 messaging apps are WhatsApp, Messenger, WeChat, Viber.
Big 4 social networks are Facebook, Instagram, Twitter, Snapchat
Source: Companies, BI Intelligence

BI INTELLIGENCE

As messaging apps have become indispensable parts of our lives, Enterprises are determined to be where their customers are.

What is Natural Language Processing?

What is Natural Language Processing?

- is a field at the intersection of Computer Science, Artificial Intelligence (AI) and Linguistics.
- **Goal**

  For computers to process or "understand" natural language in order to perform tasks that are useful

- Not to be confused with "Computational Linguistics"
- Deep NLP = Deep Learning based Natural Language Processing

Example I - Sentiment Analysis

Prod:
==The hotel is really beautiful==. Moviestar feeling and decadence from yesterday. The pool is designed by Johnny Weissmuller. ==So it was a trendy pool==. The ==food at the restaurant was really good. Very nice and helpful service at the frontfesk==.

Cons: this is what made my grade a 3 instead of 4. We had ==problems to get the wi-fi working==. If you're not depend this is not interesting. We talked several times with the front desk.

When we're there they had party event in the pool area between noon and 5 PM. The ==pool area was occupied with young party animals==. ==So the area wasn't fun for UD==.

Example II - Machine Translation

Input sentence:			
李克強此行將啟動中加總理年度對話機制，與加拿大總理杜魯多舉行兩國總理首次年度對話。	Li Keqiang premier added this line to start the annual dialogue mechanism with the Canadian Prime Minister Trudeau two prime ministers held its first annual session.	Li Keqiang will start the annual dialogue mechanism with Prime Minister Trudeau of Canada and hold the first annual dialogue between the two premiers.	Li Keqiang will initiate the annual dialogue mechanism between premiers of China and Canada during this visit, and hold the first annual dialogue with Premier Trudeau of Canada.
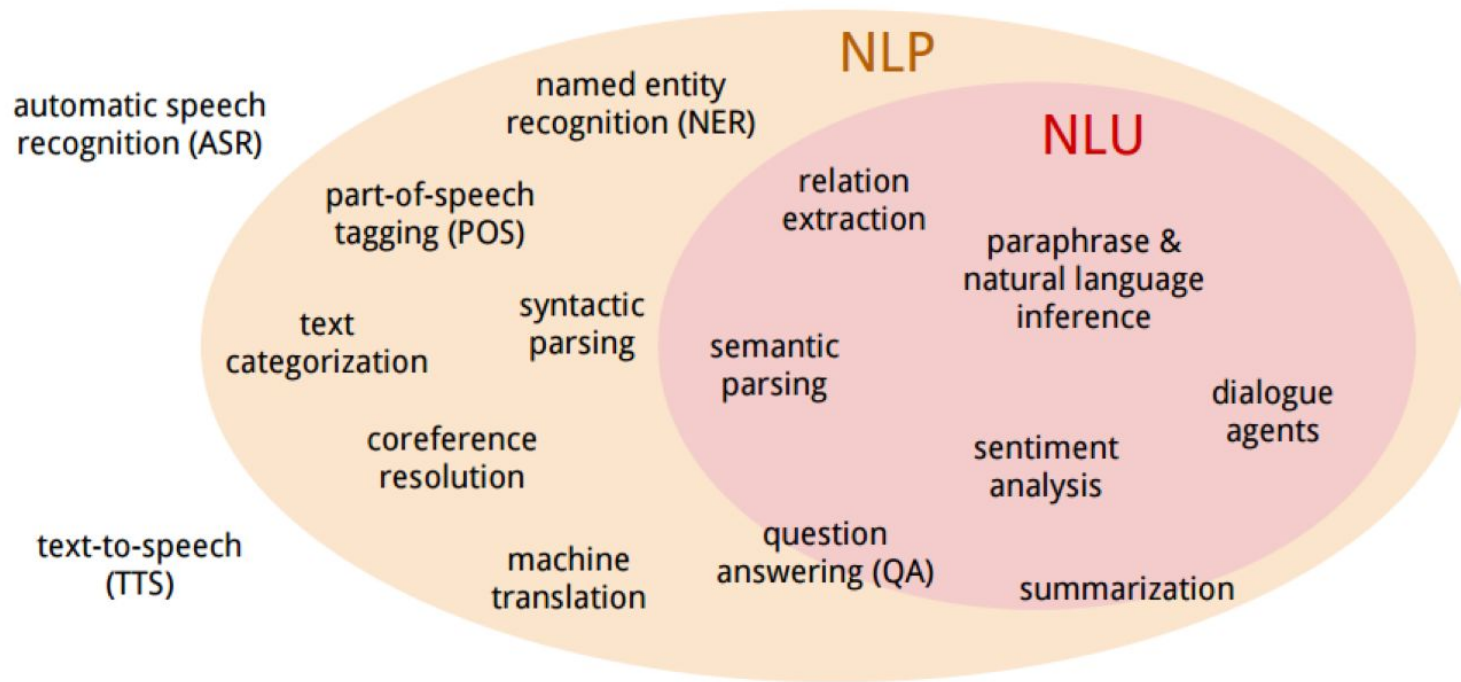
Source: Google AI

NLU vs NLP

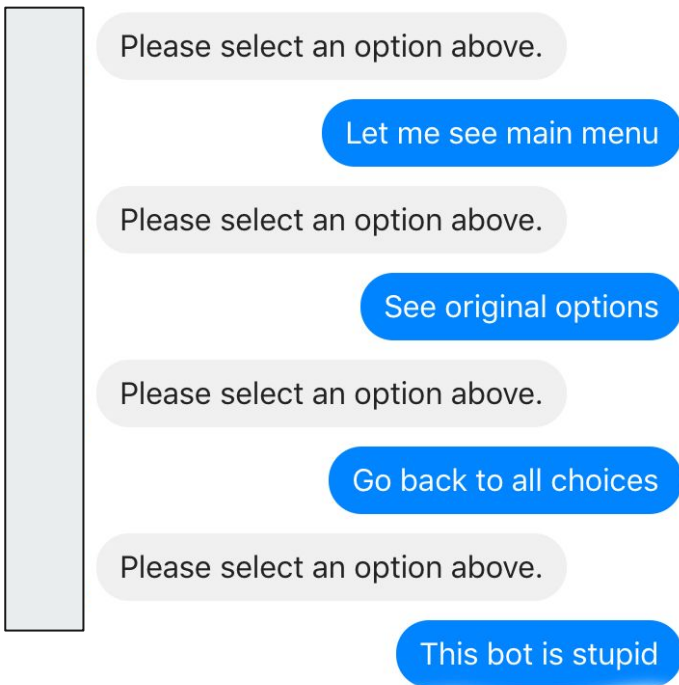Source: Stanford NLP Group

Current State of Conversational AI

"While Siri, Alexa and the likes can follow simple spoken or typed commands and answer basic questions; they can't hold a conversation and have no real understanding of the words they use. "
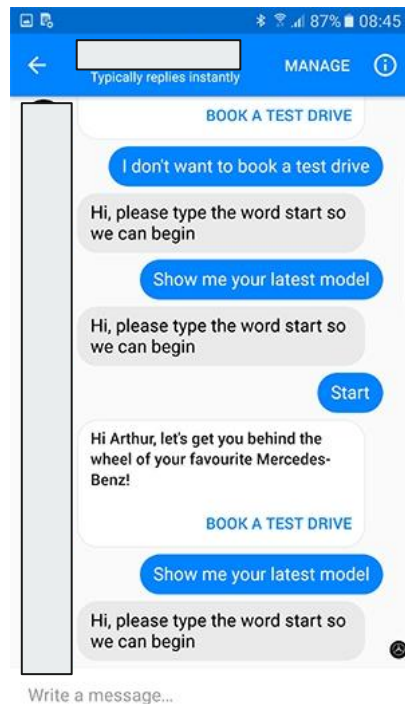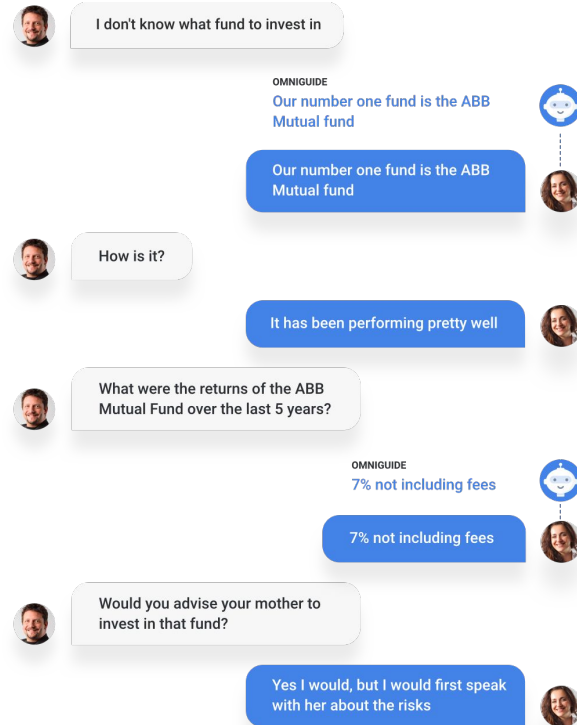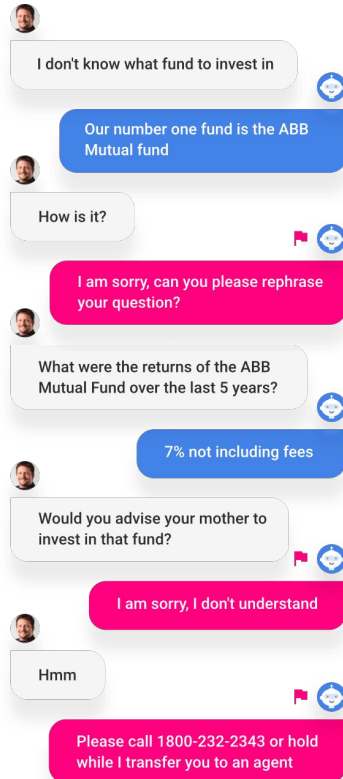
Will Knight

In the past...

Chatbot A

Please select an option above.

Let me see main menu

Please select an option above.

See original options

Please select an option above.

Go back to all choices

Please select an option above.

This bot is stupid

Chatbot B

BOOK A TEST DRIVE

I don't want to book a test drive

Hi, please type the word start so we can begin

Show me your latest model

Hi, please type the word start so we can begin

Start

Hi Arthur, let's get you behind the wheel of your favourite Mercedes-Benz!

BOOK A TEST DRIVE

Show me your latest model

Hi, please type the word start so we can begin

Write a message...

Now

Source: Salemove

Why is NLP difficult?

"Buffalo buffalo Buffalo buffalo buffalo buffalo Buffalo buffalo"

- This is a grammatically correct sentence in American English.

What is Homonym?

(In linguistics) Words with identical pronunciation and spelling, whilst maintaining different meanings.

bear	1. a large furry animal 2. to carry
bat	1. A flying rodent 2. a wooden stick
cold	1. a common illness 2. very chilly
bowl	1. a dish for eating 2. to knock pins over
cast	1. people in a movie 2. to throw
hide	1. the skin of an animal 2. to disappear from view

Buffalo - a homonym

1 - Noun: The city of Buffalo, at western New York state, US.

2 - Noun: The animal, Buffalo.

3 - Verb: To confuse, deceive or intimidate.

Three groups of buffalo

"Buffalo buffalo Buffalo buffalo buffalo buffalo Buffalo buffalo"

[Buffalo buffalo] [Buffalo buffalo buffalo] [buffalo Buffalo buffalo].

↓ ↓ ↓

[Buffalo Group 1]   [who are bullied by Buffalo Group 2]   [bully Buffalo Group 3].

Why is NLP difficult? [cont]

- Sheer complexity of sentence structure
- Ambiguity
  - Eg: "I made her duck"
- Meaning is context sensitive
  - Depends on the people present e.g. "How far is it?" (miles, km?)
  - Depends on the time of day, e.g. "Let's go eat"
  - Depends on prior sentences: "The third one"
- Recognizing named entities (people/places/…)
- Slang, jargon, humour, sarcasm, spelling mistakes, grammar mistakes and abbreviations…

Dialog research

- Purpose of language
  - Used to accomplish communication goals.
- Hence, solving dialog is a fundamental goal for NLP.
- Dialog can be seen as a single task (learning how to talk) OR as thousands of related tasks that require different skills, all using the same input and output format
- Example
  - Booking a restaurant,
  - Chatting about sports/news
  - Answering factual questions
  - …almost anything can be posed as Question Answering.

Open Source Framework

- To provide an unified framework for the training and testing of dialog models
- Integration of Amazon Mechanical Turk & Facebook Messenger
- Multi-task training
- Datasets for over 20 tasks

ParlAI Tasks

QA datasets
SQuAD
bAbI tasks
MCTest
SimpleQuestions
WikiQA, WebQuestions,
WikiMovies, MTurkWikiMovies
MovieDD (Movie-Recommendations)
MS MARCO
TriviaQA
InsuranceQA

Dialogue Goal-Oriented
bAbI Dialog tasks
Dialog-based Language Learning bAbI
Dialog-based Language Learning Movie
MovieDD-QARecs dialogue
personalized dialog, bAbI+

VQA/Visual Dialogue
VQAv1, VQAv2, VisDial, FVQA, CLEVR

Sentence Completion
QACNN
QADailyMail
CBT
BookTest

Dialogue Chit-Chat
Ubuntu
Movies SubReddit
Cornell Movie
OpenSubtitles

Negotiation
Deal or No Deal?

Machine Translation
WMT EnDe (in progress)

Add your own dataset!
Open source...

Source: FAIR

SQuAD Tasks

In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under **gravity**. The main forms of precipitation include drizzle, rain, sleet, snow, **graupel** and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals **within a cloud**. Short, intense periods of rain in scattered locations are called "showers".

What causes precipitation to fall?
gravity

What is another main form of precipitation besides drizzle, rain, snow, sleet and hail?
graupel

Where do water droplets collide with ice crystals to form precipitation?
within a cloud

bAbI Tasks

Task 1: Single Supporting Fact
Mary went to the bathroom.
John moved to the hallway.
Mary travelled to the office.
Where is Mary? A:office

Task 2: Two Supporting Facts
John is in the playground.
John picked up the football.
Bob went to the kitchen.
Where is the football? A:playground

Task 3: Three Supporting Facts
John picked up the apple.
John went to the office.
John went to the kitchen.
John dropped the apple.
Where was the apple before the kitchen? A:office

Task 4: Two Argument Relations
The office is north of the bedroom.
The bedroom is north of the bathroom.
The kitchen is west of the garden.
What is north of the bedroom? A: office
What is the bedroom north of? A: bathroom

Source: FAIR

bAbI Tasks

Task 5: Three Argument Relations

Mary gave the cake to Fred.
Fred gave the cake to Bill.
Jeff was given the milk by Bill.
Who gave the cake to Fred? A: Mary
Who did Fred give the cake to? A: Bill

Task 6: Yes/No Questions

John moved to the playground.
Daniel went to the bathroom.
John went back to the hallway.
Is John in the playground? A:no
Is Daniel in the bathroom? A:yes

Task 7: Counting

Daniel picked up the football.
Daniel dropped the football.
Daniel got the milk.
Daniel took the apple.
How many objects is Daniel holding? A: two

Task 8: Lists/Sets

Daniel picks up the football.
Daniel drops the newspaper.
Daniel picks up the milk.
John took the apple.
What is Daniel holding? milk, football

Task 9: Simple Negation

Sandra travelled to the office.
Fred is no longer in the office.
Is Fred in the office? A:no
Is Sandra in the office? A:yes

Task 10: Indefinite Knowledge

John is either in the classroom or the playground.
Sandra is in the garden.
Is John in the classroom? A:maybe
Is John in the office? A:no

Source: FAIR

bAbI SOTA result

| #	Tasks	Single task weakly trained (1K data)				All task joint - weakly trained (20 x 1K data)
					Published SOTA	
		DMN best*	DMN baseline	DMTN	MemN2N with PE,LS,RN	MemN2N with PE,LS,Lw
1	single-supporting-fact	100	100	100	100	99.9
2	two-supporting-facts	32.7	29.9	35.9	91.7	81.2
3	three-supporting-facts	26.3	32	36	59.7	68.3
4	two-arg-relations	89.8	82.1	100	97.2	82.5
5	three-arg-relations	97.5	97.6	97.3	86.9	87.1
6	yes-no-questions	96.3	96.6	96.5	92.4	98
7	counting	80.3	77.6	80.1	82.7	89.9
8	lists-sets	76.6	98.6	98.3	90	93.9
9	simple-negation	94.1	95	95.3	86.8	98.5
10	indefinite-knowledge	95	90.2	93.2	84.9	97.4
11	basic-coreference	100	71.4	99.9	99.1	96.7
12	conjunction	100	67.2	99.8	99.8	100
13	compound-coreference	94.4	92.5	94.4	99.6	99.5
14	time-reasoning	77.6	74.7	73	98.3	98
15	basic-deduction	67	50.5	99.9	100	98.2
16	basic-induction	48.1	44.8	46.4	98.7	49
17	positional-reasoning	65.9	52	62.7	49	57.4
18	size-reasoning	93.7	91.5	95.8	88.9	90.8
19	path-finding	10.7	8.2	9.9	17.2	9.4
20	agents-motivations	98.1	97.6	98.2	100	99.8
	passed task (acc>=95%)	7	6	11	9	10

GS Ramachandran et al 2017

Visual Question Answering
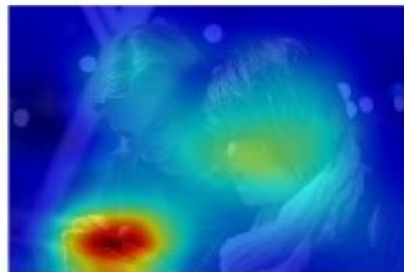
Question: What are they doing?
Prediction: texting (score: 12.02=3.78 [image] + 8.24 [word])
Word importance: doing(7.01) are(1.05) they(0.49) what(-0.3)
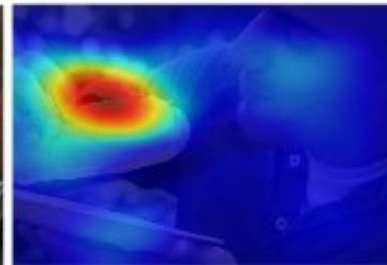
Question: What is he eating?
Prediction: hot dog (score: 13.01=5.02 [image] + 7.99 [word])
Word importance: eating(4.12) what(2.81) is(0.74) he(0.30)

Amazon Mechanical Turk Live Chat

Multi-Task Learning (MTL)

- Convention ML
  - Optimize for a particular metric (RMSE, AUC etc)
  - Train a single model or an ensemble of models
  - Fine-tune and tweak

- Multi-task Learning
  - Also known as "joint learning", "learning to learn" etc
  - Optimize for more than one loss function
  - Why is it better?
    - It improves _generalization_ by leveraging the domain-specific info contained in the training signals of related tasks.

The grand vision

- Avoid siloed research
- Identify model weaknesses and iterate faster
- Build models that can generalize well to many tasks!

Recent Advancements

Deal or no deal? Training AI bots to negotiate

Dialog rollouts

read: You get one book and I'll take everything else.	write: Great deal, thanks!	read: Any time	choose: 1x *book*	1
		read: No problem	choose: 1x *book*	1
	write: No way, I need all 3 hats	read: I'll give you 2	choose: 2x *hat*	6
		read: Ok, fine	choose: 3x *hat*	9

Dialogue history — Candidate responses — Simulation of rest of dialogue — Score

- Intelligent maneuvers
  - At first, the bot shows great interest in a valueless item
  - Later "compromise" to "please" the opponent
  - An effective negotiation tactic that people use regularly
  - Not programmed by the researchers

Dialog agents with ability to negotiate

```
==============================================================================
Dialog 0
------------------------------------------------------------------------------
Human : book=(count:1 value:2) hat=(count:2 value:2) ball=(count:2 value:2)
Alice : book=(count:1 value:4) hat=(count:2 value:0) ball=(count:2 value:3)
------------------------------------------------------------------------------
Human : hi i want the hats and the balls
Alice : i will take the balls and book <eos>
Human : no i need the balls
Alice : i will take the balls and book <eos>
Human : i need the balls
Alice : you can have the hats if i get the balls and book <eos>
Human : no i need the balls and you can take the rest
Alice : i will take the balls and 1 book <eos>
Human : no, no balls for you
Alice : ok , you can have two hats and one ball <eos>
Human : ok deal
Alice : <selection>
Human choice: 0 2 1
Human : book=0 hat=2 ball=1
Alice : book=1 hat=0 ball=1
------------------------------------------------------------------------------
Agreement!
Human : 6 points
Alice : 7 points
------------------------------------------------------------------------------
```
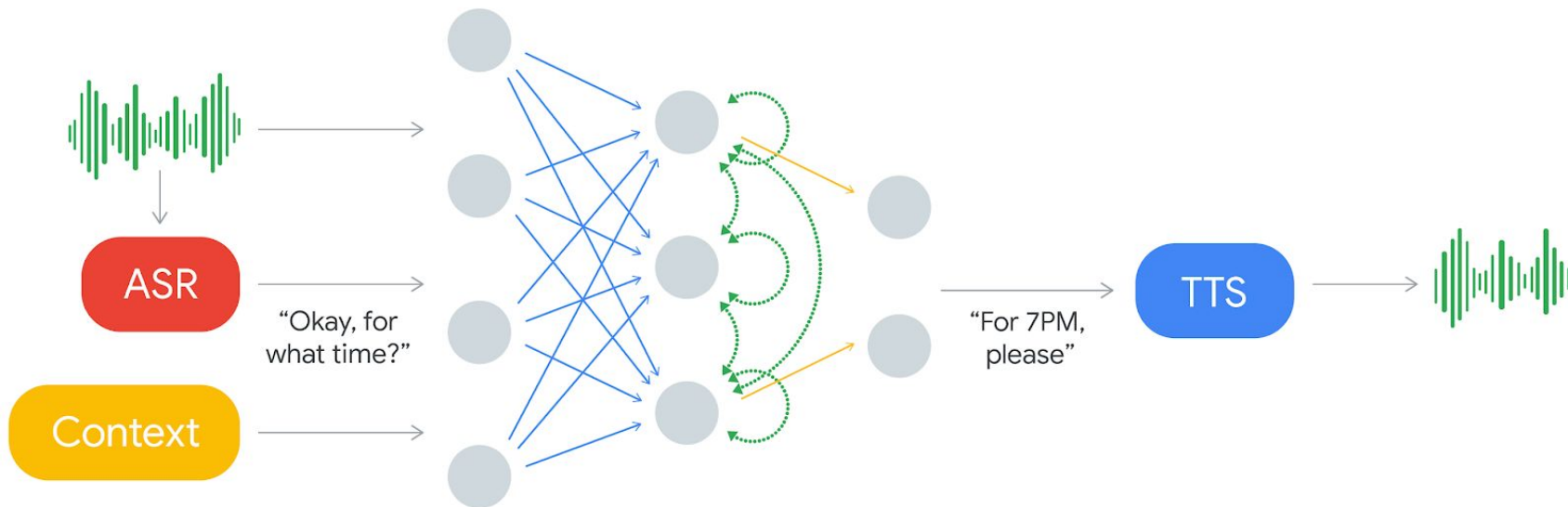
Google Duplex

"Do you have anything between 10am and 12pm?"

Google Duplex

- An AI system which is capable of accomplishing "real world" tasks by conducting natural conversations.

Source: Google

What we have learnt

...after launching multiple chatbots with Fortune Global 500 clients

User Acquisition

Is easy. But not User Retention!

User Expectations

Set the right expectation.

Content Co-creation

Knowledgebase curation.

Further readings

- Conversational Agents

  - http://www.wildml.com/category/conversational-agents/

- LSTMs, Attention & Augmented RNNs

  - https://colah.github.io/posts/2015-08-Understanding-LSTMs/

  - https://distill.pub/2016/augmented-rnns/

- Understanding, Deriving and extending LSTM

  - https://r2rt.com/written-memories-understanding-deriving-and-extending-the-lstm.html

We're hiring Data Scientists & Software Engineers!
Send your CV to **david**@**pand.ai**