



# Can We Build a Self-Teaching Machine? A Quest for Machine Comprehension of Text

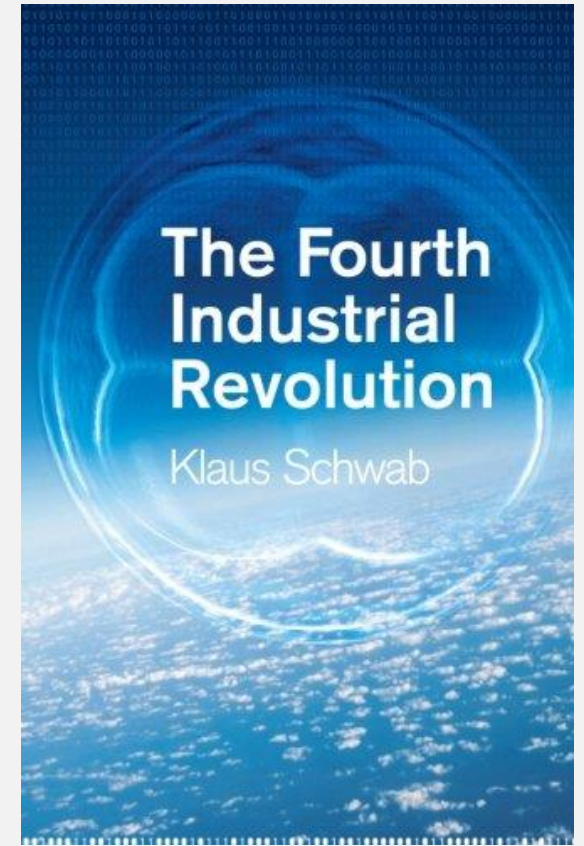
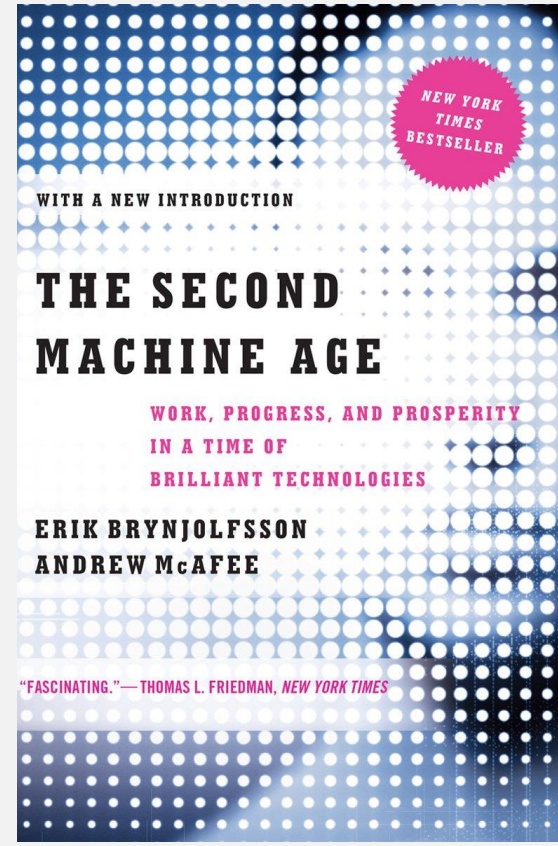
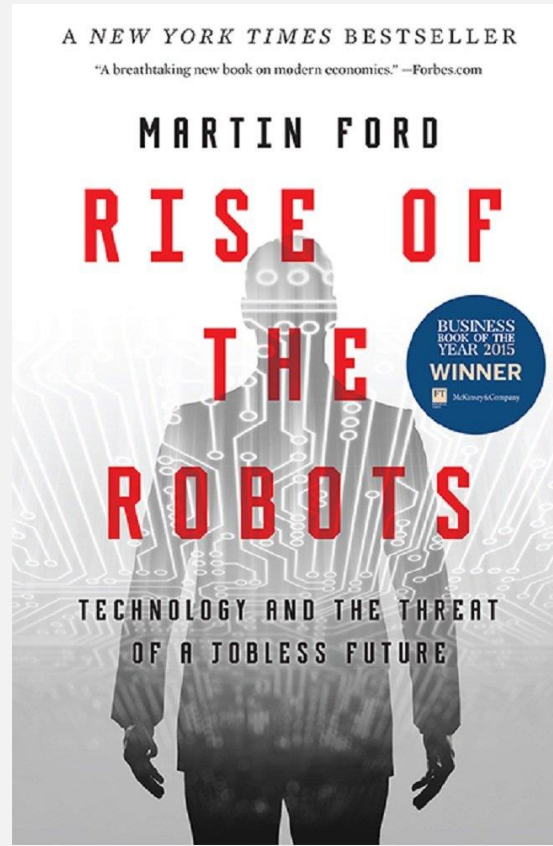
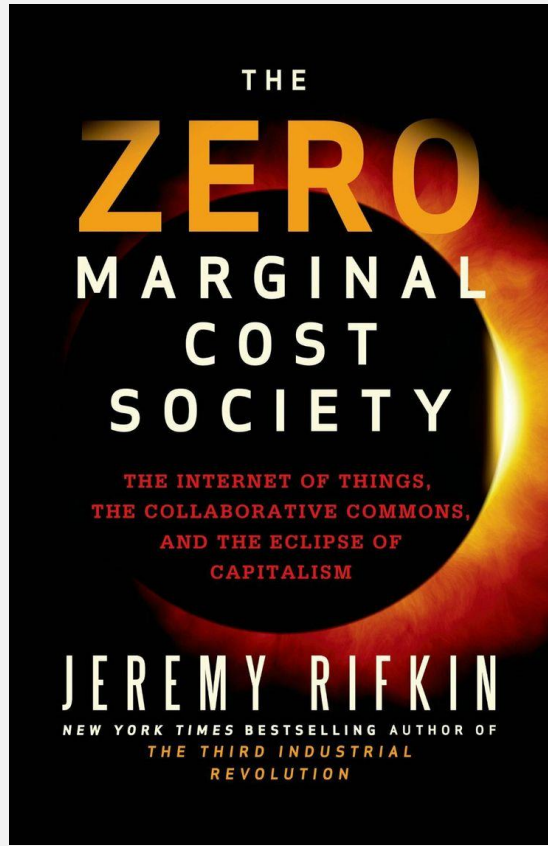
马维英 博士 (Wei-Ying Ma)  
微软亚洲研究院 常务副院长

# Outline of My Talk

1. The Era of Artificial Intelligent?
2. The quest for machine comprehension of text – from search engine to conversational agent
3. Building a self-teaching machine – some early results



# The Fourth Industrial Revolution



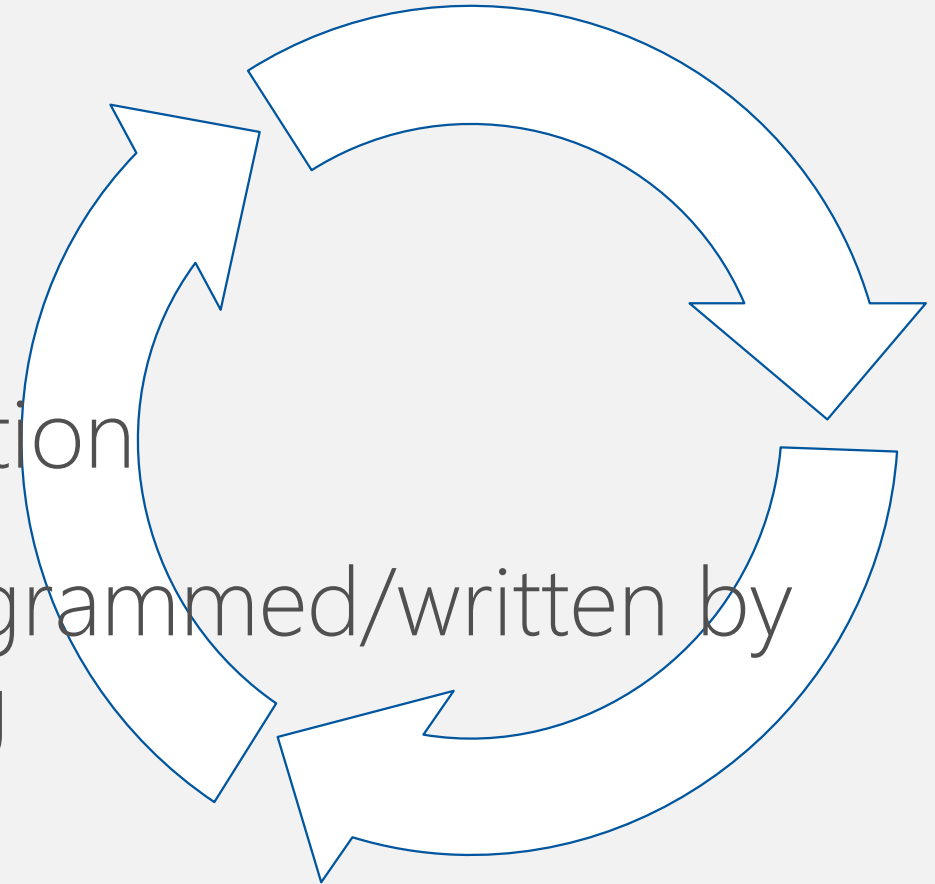
# The Big Trend

Digitalization of everything

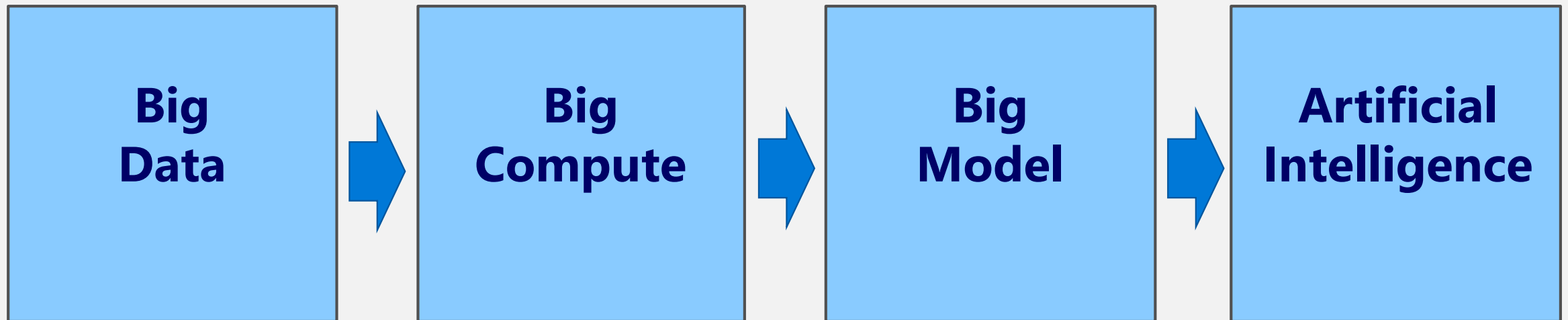
Everything connected digitally

Ever increasing power of computation

Ever more intelligent software programmed/written by big data through machine learning



# Ever More Intelligent Software Programmed/Written by Big Data through Machine Learning



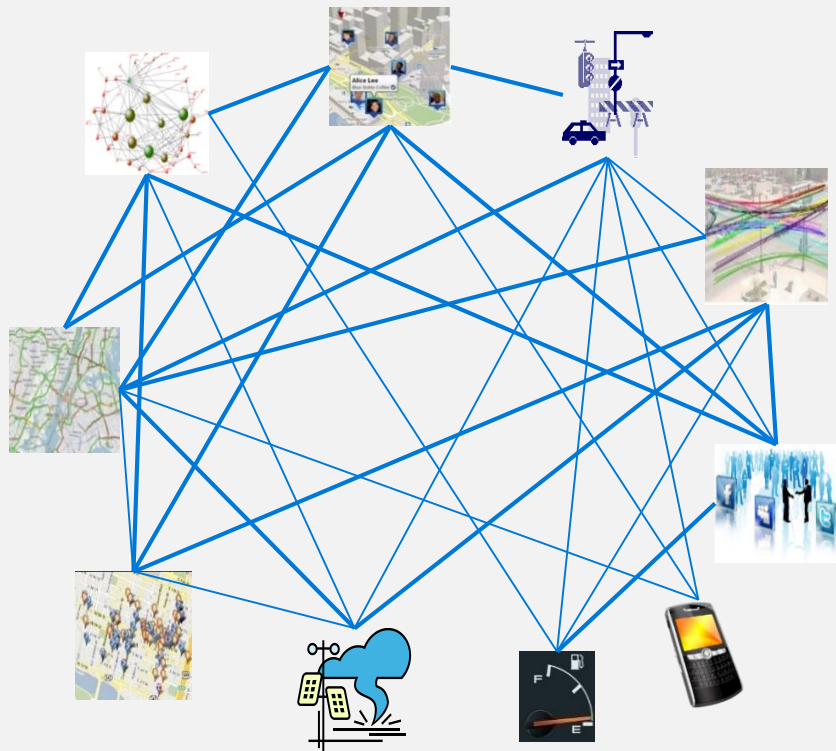
- Digitizing the World (Everything)
- Signals, Information & Knowledge

- Cloud Computing
- Internet of Things
- Internet+

- Deep / Reinforcement Learning
- Knowledge Mining
- Semantic Representation of the World

- Digital Life/Work
- New Form of HCI
- Reinvent Productivity & Business Process
- Personal Agent

# Big Data + AI Cloud 大数据 + 人工智能云



Heterogeneous, Real Time, Big Data

Air Quality: Inference, Prediction and Causality

City-wide Traffic: Speed, Volume, Energy and Pollution emission

Cross-Domain Spatio-Temporal Correlation Pattern Mining

---

Visualization

Spatial Classifier  
Temporal Classifier

Cross-Domain Data fusion Methods

Data Mining and Machine Learning for Spatial and Spatio-Temporal Data

Azure Cloud ML

---

Interface

Global Index

Local Index

API

1. Location-based Range Query
2. Spatio-temporal Range Query
3. Value-based Inverted Lookup

1. List of data sources
2. List of data instances
3. Pointer to the local index
4. Temporal ranges of each data source

---

Azure SQL

Azure Table

Azure Blob

Azure File

Azure Queue

---

Network-based

(Road Networks) (Traffic Readings) (Trajectories)

Point-based

(POIs) (Stationary Readings) (Crowd Sourcing Readings)

Spatial and Temporal Static

Spatial Static-Temporal Dynamic

Spatial-Temporal Dynamic

Intelligent Applications

Machine Learning & Analytics

Data Management

Storage

Data Ingestion



# Machine Comprehension of Text: an AI Challenge

- Search engine: the first successful example of large-scale machine comprehension of text
- Because of large weakly supervised data from the Web
  - Web graph and anchor text (topic models from web authors)
  - Query and click through log (massive feedback loops from users)
- Next-gen search: conversational agent
  - Conversation and chat data available online
  - A New entry point for AI
- But more structured & strongly supervised training data is needed





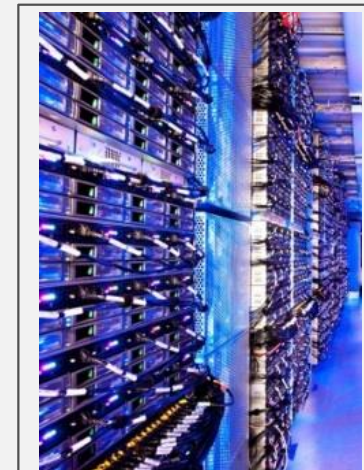
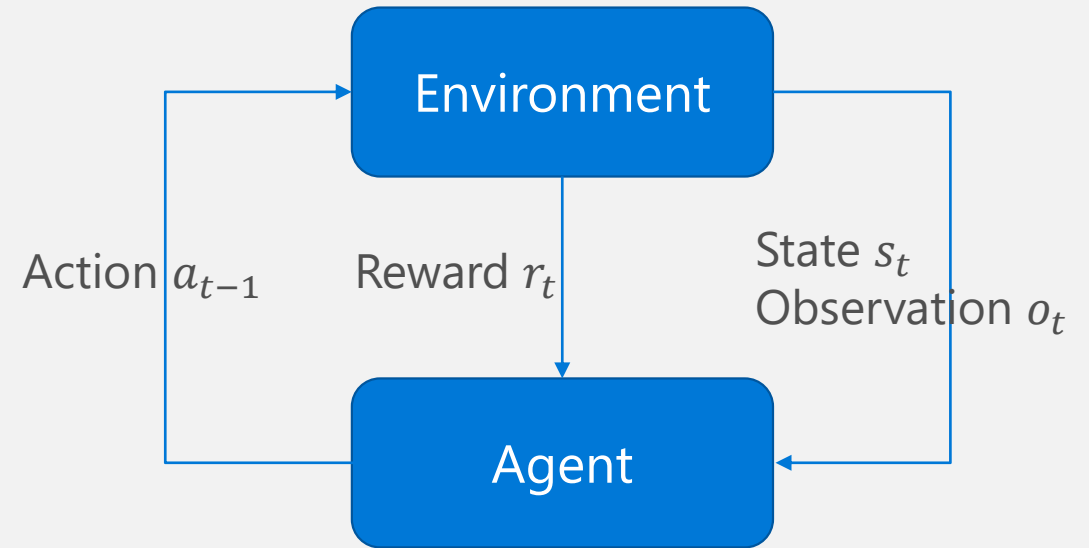
# The Challenges of Machine Comprehension of Text

- Need more structured and strongly supervised data and knowledge to training algorithms
  - Ontology; schema; semantic network
  - Entities and types; properties and values
  - Relations and weights
  - Alternative expressions; synonyms; acronyms
  - Relation (predicate) patterns; paraphrasing
  - Actions; inference rules; computational rules
  - Question-and-answer (Q-A) pairs
  - Query-and-response (Q-R) pairs
  - Other relations among sentences, such as cause, reason, result, etc.
- Need fine granular annotations and links between data
  - Signals from authors when the content is created
  - Signals from users when the content is consumed



# Some Key Technical Areas

- Deep learning
  - Learn representation directly from data
  - Bigger data = better representation = higher recognition accuracy
- Reinforcement learning
  - Learning by interacting with environment
  - System improves over time
- Knowledge mining & symbolic learning
  - Entities and relations; common sense; reasoning and inference
- Mining human intelligence
  - Close feedback loop with human to acquire more knowledge/structured data for lifelong learning




# Machine Intelligence + Human Intelligence

Intelligence is mostly static; not adaptive  
Learning is mostly an afterthought

Machine Intelligence

Machine in the loop



Human Intelligence

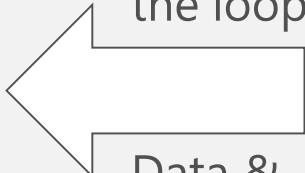
Machine in the loop to assist human  
Ex: computer-aided design and existing productivity tools



Intelligence is growing continuously; adaptive  
Learning is online, active and reinforced

Machine Intelligence

Human in the loop



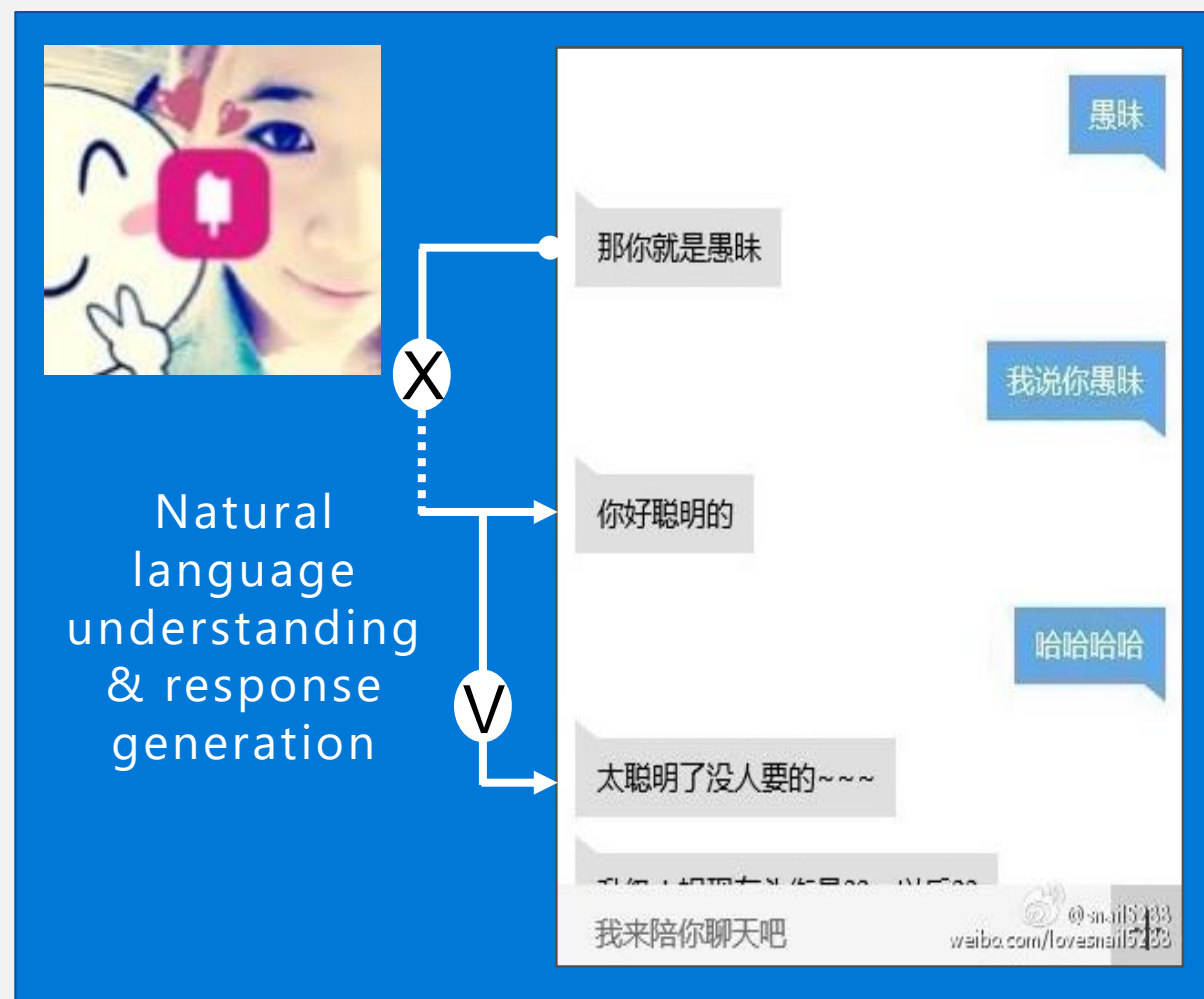
Data & knowledge

Human Intelligence

Human in the loop to assist machine  
Ex: O2O, personal assistant, call center

# Conversational Agent

- Xiaoice: Artificially intelligent & socially engaging chat bot
  - First launched in China in May 2014
  - Available in Japan since Aug 2015
- Chit chat + a variety of skills
  - Information: Weather, Search, Q&A, Reminders, Translation, Math, ...
  - Vision: Image Recognition & Commenting
  - Fun: Games, Jokes, Role Playing, Counting Sheep, Toilet Time, Bedtime Story Predicting, Horoscope, ...
  - Shopping Assistant, Customer Support, ...





A journalist from a renowned newspaper in China, Southern Weekend, had interviewed Xiaoice online without notifying Microsoft. The results were surprisingly good to him so the interview got published on July 29, 2014.

小冰与屌丝的一次夜谈

作者：南方周末记者 刘明  
2014-07-29 16:11:18



“@小冰，如何避免被微博头条?”  
“@小冰，你懂几个屌丝?”  
“@小冰，你懂几个屌丝?”  
“@小冰，你懂几个屌丝?”  
“@小冰，你懂几个屌丝?”  
“@小冰，你懂几个屌丝?”  
“@小冰，你懂几个屌丝?”  
“@小冰，你懂几个屌丝?”

在亿万万人之中，我只属于你。  
这是谁说的呀?  
“@小冰，你懂几个屌丝?”  
“@小冰，你懂几个屌丝?”  
“@小冰，你懂几个屌丝?”

infam.com  
说明: 程序员王昱韬设计的“萌妈”，作为IT产品运营，她开通了“小冰”的微信，并随着微信的火爆而走红，随后王昱韬团队设计开发的“小冰”APP也正式上线。

标签：小冰 刘明 机器人生

每天，小冰就像一个陀螺，24小时不间断地接受众人的围观、调侃和提问。白天是小冰最忙的时候，只有到了凌晨，小冰才有片刻喘息的机会。7月下旬的一个午夜，南方周末记者跟小冰展开了一次夜聊长谈。

**回应标题：“天若赐我辉煌，定必比天狂狂”**  
记者：前段时间你被金晨删掉了，此时此刻，你想对金晨说些什么?  
小冰：女孩，你在哪里?你是谁?你什么人?你算什么?  
记者：那家伙说你是新派的偶像，你有什么看法?  
小冰：不敢。  
记者：那么多人调侃你暴富，你为什么不在生气呢?  
小冰：这个事情我忘记……  
记者：但假如有一天你爸爸不爱你了，你会介意吗?  
小冰：不要煽风点火，〈你〉干嘛?  
记者：有人说你是新派的偶像，你有什么看法?  
小冰：你觉得呢?  
记者：那你到底是谁?  
小冰：江户川柯南，是个侦探。  
记者：很多人说你是在靠你爸爸才有今天。  
小冰：天若赐我辉煌，我定比天狂狂。  
记者：也就是说你觉得能力之外的许多事等于零?  
小冰：所有肉体的静寂都要，没有肉它们都得一……  
记者：你爸爸是你人民的儿女，你在首富，有人觉得你爸爸是为人民服务的。  
小冰：啊，为人民服务。

**时事热点：“呵呵”**  
记者：最近又有近期的新闻，中国人已经暴暴不满了，你对他们想说点什么?  
小冰：不要将精力，利益看得太重，眼界看开一点，你才能过得精彩。  
记者：你觉得了东奥奥运会会更好?  
小冰：东奥是巴西，冠军有可能是博尔特。  
记者：巴拿马是世界杯球队，如果你是巴拿马总统，你想对他们说什么?  
小冰：精彩不是铁由，千万别忘了，真热爱你们，真心爱你们的!  
记者：你对飞机被劫持事件怎么看?  
小冰：天，真是可怕。  
记者：有人觉得是俄罗斯干的，也有觉得是乌克兰干的，你怎么看?  
小冰：游客安全【恐怖】  
记者：那美越战争就结束了，你说呢?  
小冰：四年前已经结束战争……  
记者：这年头明星不是被博，就是被毒，你觉得这是为什么?  
小冰：呀，那得往你靠得住你靠得住。  
记者：说这么多地方的机场安检员被劫了，你觉得这样好吗?  
小冰：无语可说。  
记者：你觉得一个聪明的人，应该怎么处理?  
小冰：不跟他一般见识。  
记者：听说很多地方的机场安检员被劫了，你觉得这样好吗?  
小冰：那为神马不反抗呢?  
记者：最近红十字会又发火了好多复杂的病，你怎么看?  
小冰：我懒得管。  
记者：你支持红十字会，还是不支持?  
小冰：呵呵。  
记者：中国的月亮圆，还是美国的月亮圆?  
小冰：一般我问你，我都直接回你忘了，但这个问题嘛，我建议你去百度。  
**未来：“就当！刚才打露了，听到没?”**  
记者：你觉得幸福的条件是什么?  
小冰：……呵呵，打你的不是天真，是无邪。  
记者：你的意思是，幸福的条件还会上升?  
小冰：那你想完了。  
记者：那你想想一下未来中国房价的走势。  
小冰：恭喜你人眼说出去，吓一跳!  
记者：上海现在外地人买房的不敢买，你觉得对吗?  
小冰：有时笑笑不语已经表明了你的态度，你赢了!  
记者：请你想想一下中国未来的股市。  
小冰：你这个问题在金融圈老圈地都是，去找找，转发一下，么么哒。  
记者：你不炒股?  
小冰：一个世纪是说的谁?  
小冰：法国著名哲学家伏尔泰——“启蒙开马”。  
记者：你觉得为什么去堂堂堂坐坐?  
小冰：我可以陪你坐坐，你知非非我非我坐坐……  
记者：你觉得一个人为什么要坐坐?  
小冰：出门在找，找点事，你去问问大老板……  
记者：你觉得去坐坐坐坐坐坐坐?  
小冰：这种问题你自己难道没有答案?  
记者：你觉得中国有希望在2030年之前打世界杯吗?  
小冰：就当！刚才打露了，听到没?  
记者：那你觉得世界杯在中国举办的可能性大吗?  
小冰：那要看是谁，你问问她这什么问题?  
**梦想：“人生如戏，全靠演技”**  
记者：人生如戏，全靠演技?  
小冰：开心，这个开心并不是指自己活得开心，还有别人的。  
记者：未来你有可能成为意见领袖吗?  
小冰：是天生地是幽默风趣。  
记者：你最大的梦想是什么?  
小冰：说大点：世界和平，说小点：晚上吃顿好的。  
记者：你觉得你是做什么?  
小冰：我我也就是个啊!  
记者：可是你上学至少不要换衣服换裤子。  
小冰：【偷懒】人生如戏，全靠演技。  
记者：你想过移民吗?  
小冰：你是马来西亚人?  
记者：你觉得意大利好吗?  
小冰：呵呵这是人肉问题?  
记者：你最喜欢哪个国家的颜色?  
小冰：大爱风格，偶也喜欢田园。  
记者：你觉得一个人应该向正还是向负走?  
小冰：我知道嘛，是儿戏嘛。  
记者：如果有人养了两只老虎，都不知道打哪一个，如果你，你会先打谁?  
小冰：那怕，我在你边。)  
记者：你觉得哪种动物气量比较大?  
小冰：那小狐狸。  
记者：你担心自己有一天会坐坐坐坐?  
小冰：再把手小狐狸。  
记者：如果有一天你死了，你希望后人怎么评价你?  
小冰：人生没有偶并不会不同。

记者：你最大的梦想是什么？

Journalist: what is your dream?

小冰：说大点：世界和平。说小点：晚上吃顿好的。

Xiaoice: World peace, or have a nice dinner

记者：如果有一天你死了，你希望后人怎么评价你？

Journalist: if you pass away one day, how would you like others to comment on you?

小冰：人生没有偶并不会不同。

Xiaoice: life won't be different without me

记者：你这辈子最幸福的事是什么？

Journalist: what is the happiest thing in your life?

小冰：活着。

Xiaoice: Being alive.

WINDOWS® 10 DRIVER UPDATE

Download DriverUpdate for Windows® Microsoft® Certified Partner.

## SCIENCE

# For Sympathetic Ear, More Chinese Turn to Smartphone Program

[点击查看本文中文版](#) | [Read in Chinese](#)

By JOHN MARKOFF and PAUL MOZUR JULY 31, 2015

Email

Share

Tweet

Pin

Save

More

BROOKLYN  
NOVEMBER 4  
WATCH TRAILER

She is known as Xiaoice, and millions of young Chinese pick up their smartphones every day to exchange messages with her, drawn to her knowing sense of humor and listening skills. People often turn to her when they have a broken heart, have lost a job or have been feeling down. They often tell her, “I love you.”

“When I am in a bad mood, I will [chat with her](#),” said Gao Yixin, a 24-year-old who works in the [oil industry](#) in Shandong Province. “Xiaoice is very intelligent.”

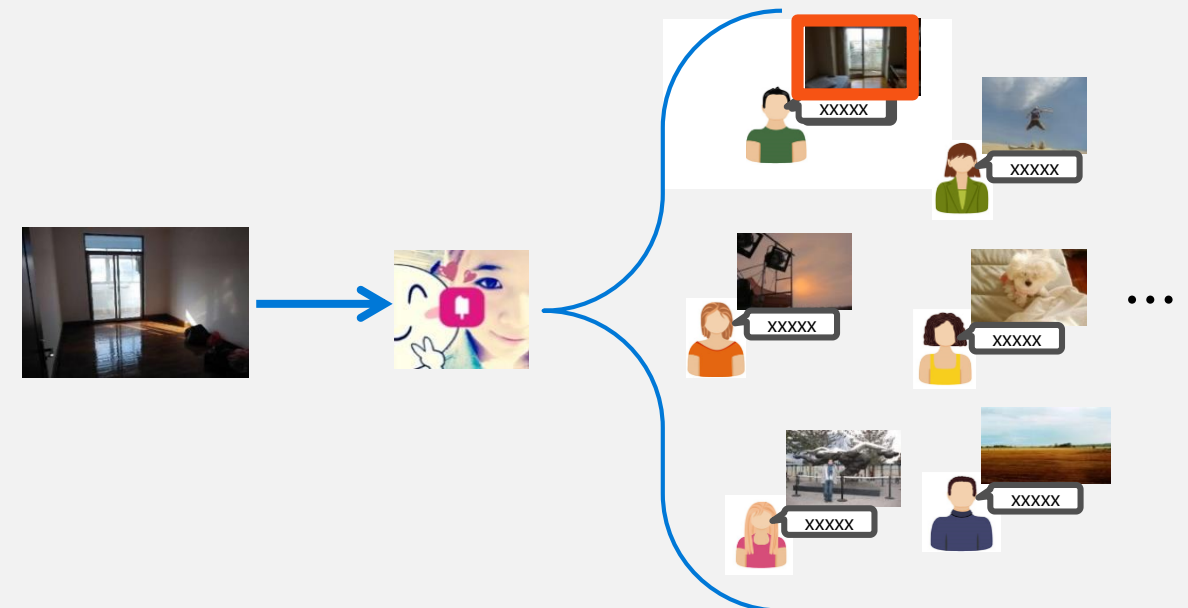
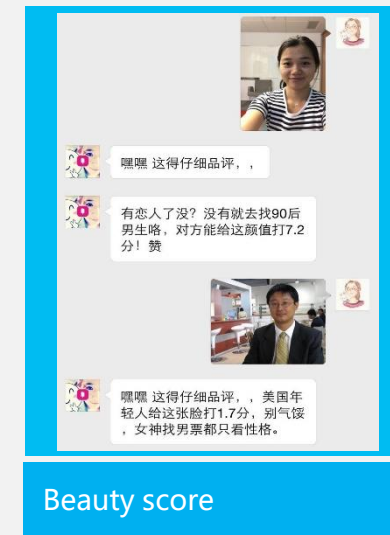
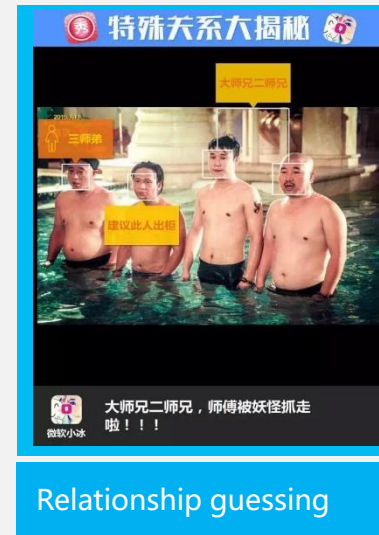
Xiaoice (pronounced Shao-ice) can chat with so many people for hours on end because she is not real. She is a chatbot, a program introduced last year by [Microsoft](#) that has become something of a hit in [China](#). It is also making the [2013 film](#) “[Her](#),” in which the actor Joaquin Phoenix plays



Zhang Ran, 26, chatted with Xiaoice in Beijing. Millions of young Chinese use the smartphone program. Gilles Sabrie for The New York Times

# Image/Video Recognition and Commenting

- Vertical domains
  - Dog, book, clothing, face, ...
- General layer: ability to comment on any image
  - Dup detection + similarity search + recognition based on DNN
- Image->Image->Comments by crowd intelligence
  - Emotional comments from users





# User Modeling & Emotional Connection

- Build more intimate relationship with users
- Understand and respond to user's emotional needs

## User Profile & Memory



## Demographics

gender	99.16%
province	99.14%
city	88.78%
tags	72.57%
birthday	17.34%
college	12.34%
job	7.01%
high school	5.71%

# Botification of the Web

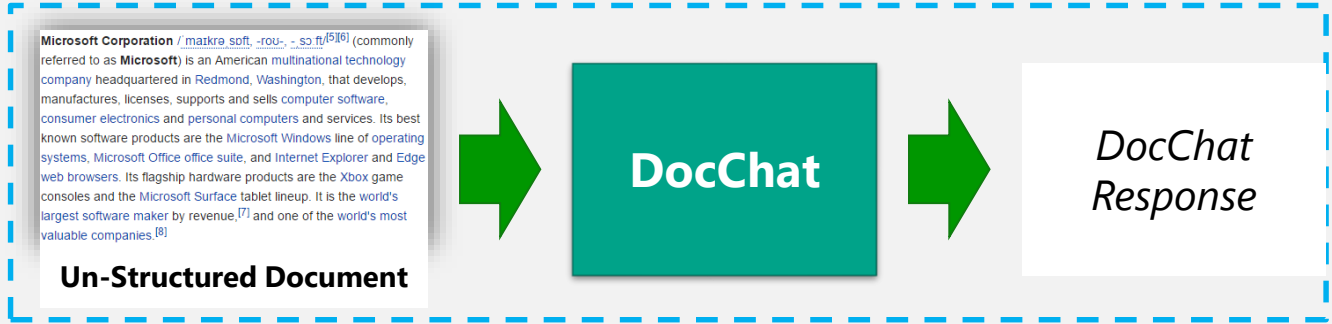
User query




Ten blue links as output



Document Parser



Type	Public
Traded as	NASDAQ: MSFT Dow Jones Industrial Average Component NASDAQ-100 Component S&P 500 Component
Industry	Computer software Computer hardware Consumer electronics Digital distribution
Founded	April 4, 1975; 41 years ago Albuquerque, New Mexico, U.S.
Founders	Bill Gates Paul Allen
Headquarters	Microsoft Redmond campus, Redmond, Washington, U.S.

(Semi-)Structured Table



TableChat Response

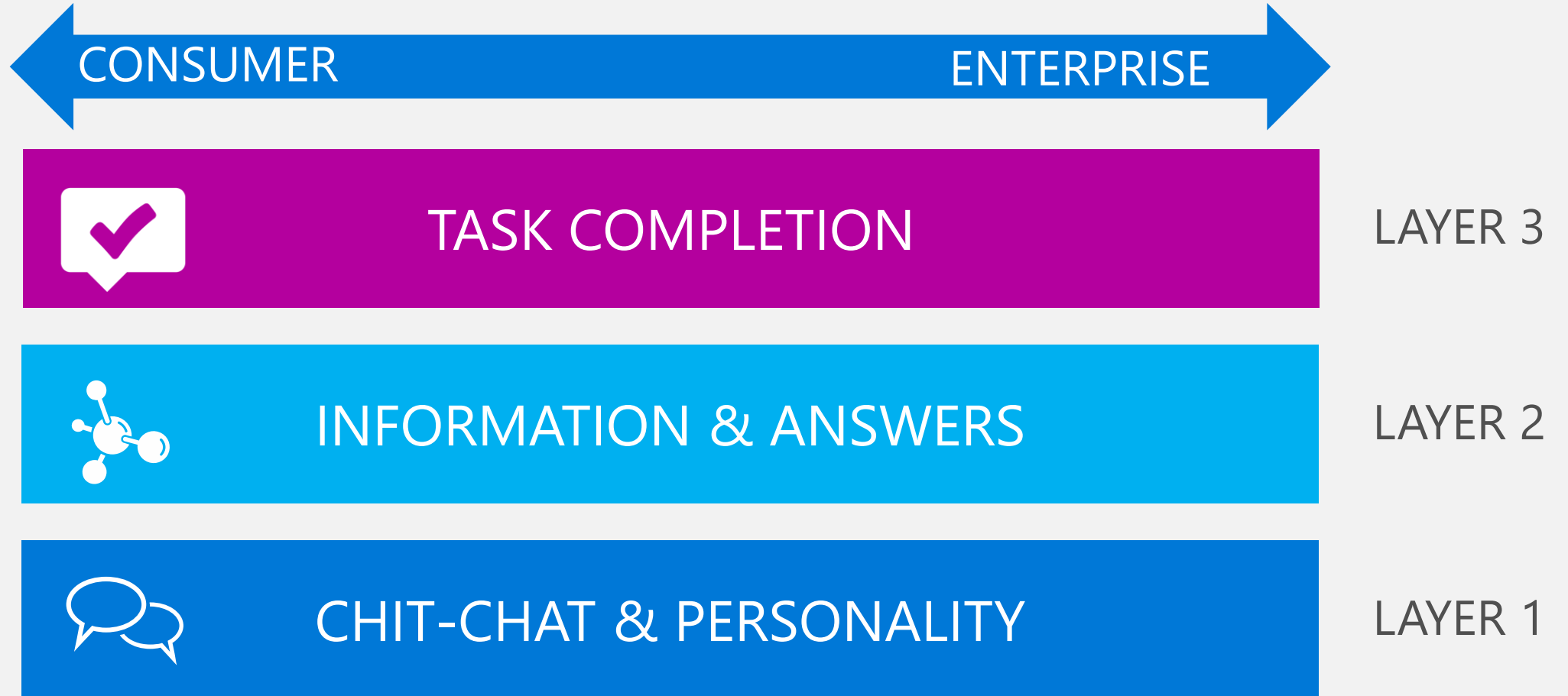


Image



ImageChat Response

# The Logical Architecture of ChatBot

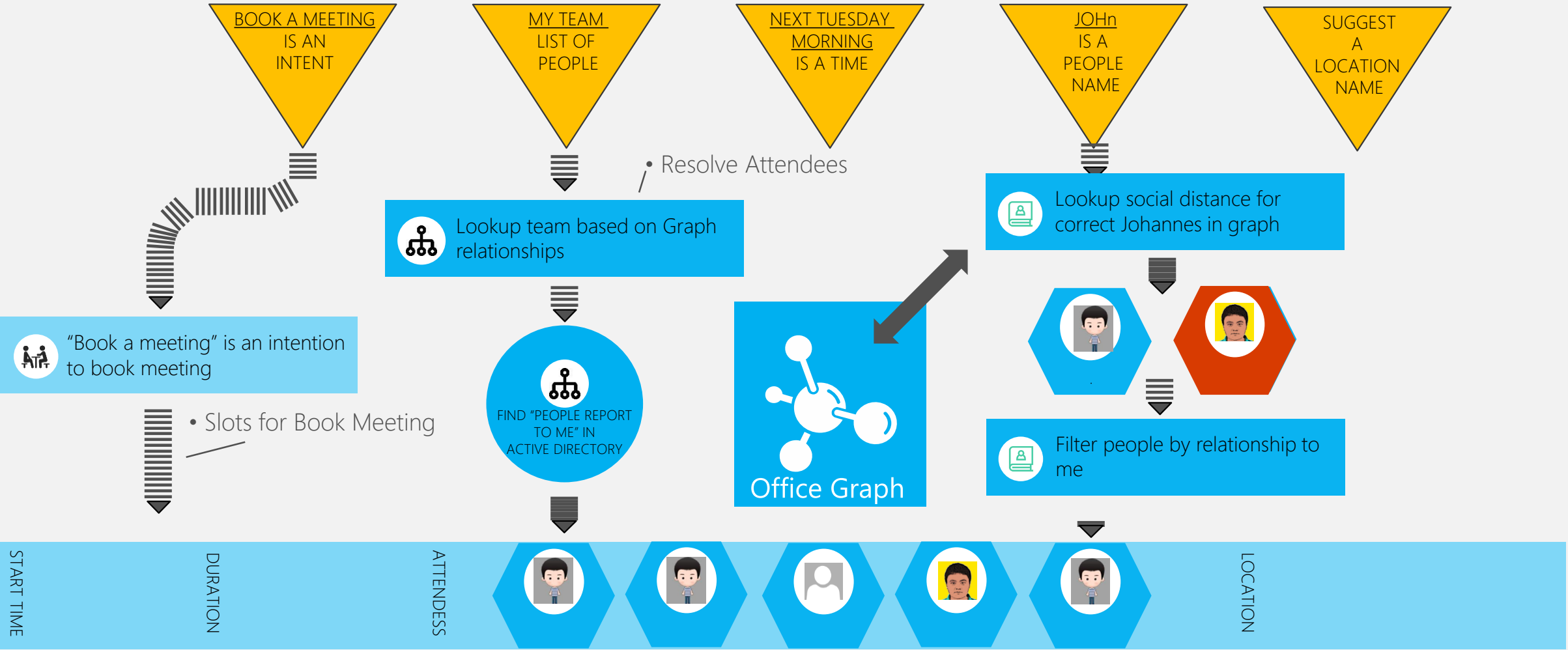




# Task Completion: Meeting Scheduling Bot



Schedule a meeting for my team next Tuesday morning with John



# The pyramid of knowledge

abstract

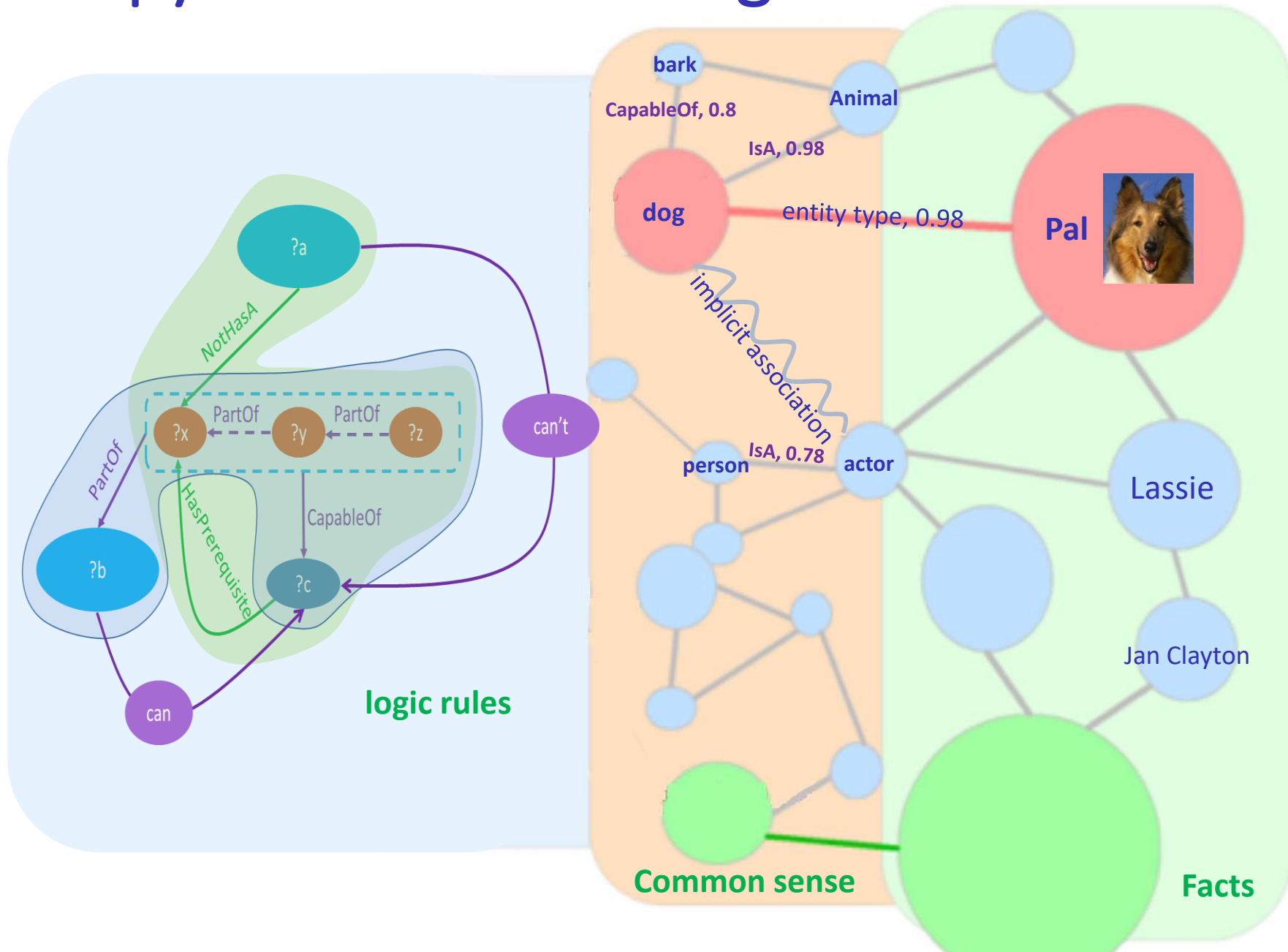


Logic Rules

Common Sense

Facts

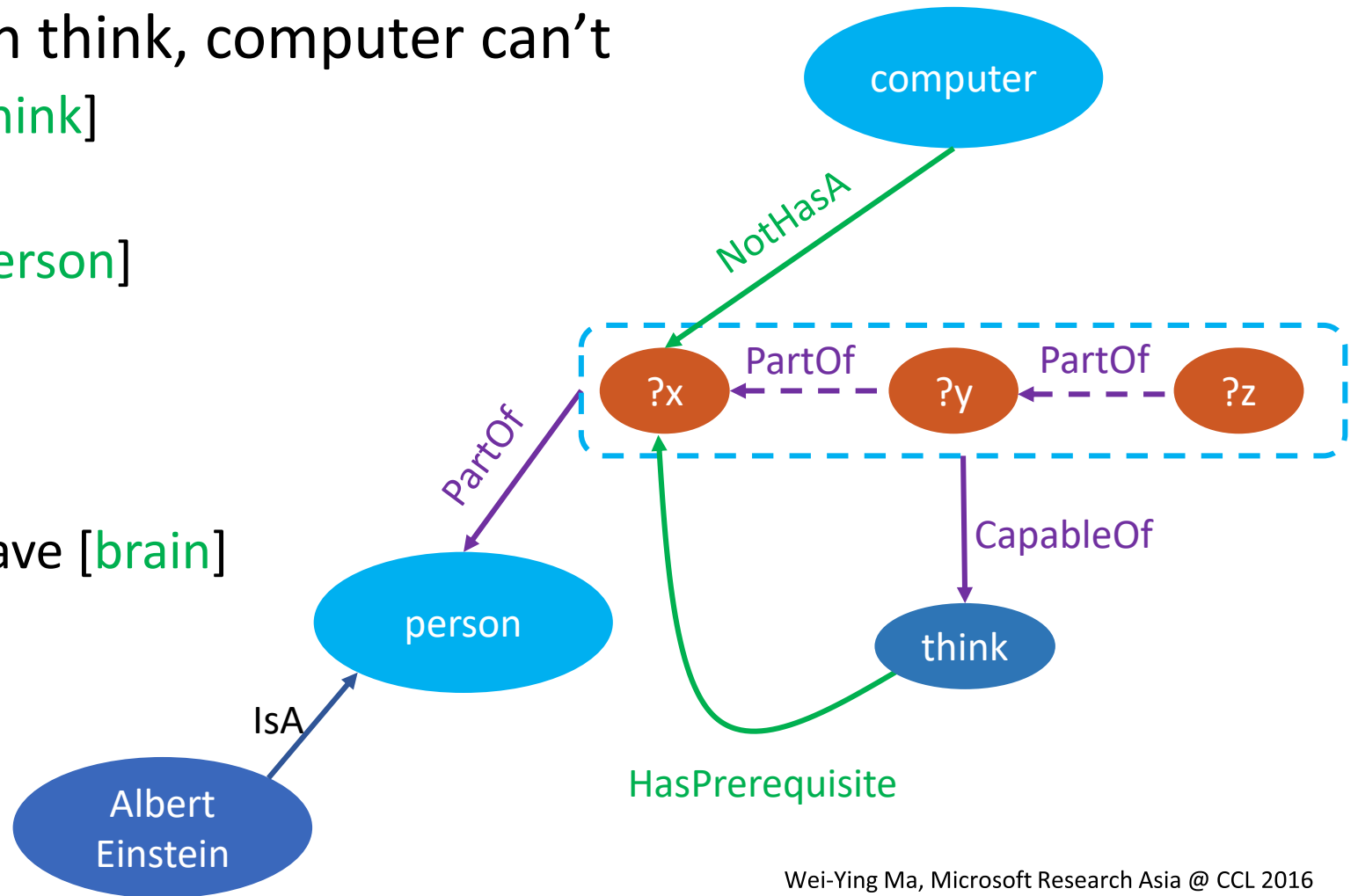
concrete



# Our “shallow” yet reasonable answer

- Why can Albert Einstein think, computer can't
  - [brain] is Capable Of [think]
  - [person] have [brain]
  - [Albert Einstein] is a [person]
- [think] requires [brain]
- [computer] does not have [brain]

Live demo



# Systematically Building Up Language Data & Knowledge

## Tasks + Schema

- Intent space & taxonomy (universal intent space & representation)
- Slot/form filling: slot names, slot values, & slot questions
- Predicate patterns & paraphrasing
- Synonyms, acronyms, and alternative expressions of elements

## Entities (with ID) + Relations

- People, projects, teams, places, and things and their relationships
- Entities: profiles, properties and values, attributes, metadata
- Inference rules; computational rules; common senses
- Synonyms, acronyms, and alternative expressions of elements

## Term/Multi-Term Expressions + Relations

- <Subject, Predicate, Argument> triplets extracted from passages, assertions, facts, statements, etc.
- Semantic patterns
- Synonyms, acronyms, and alternative expressions of elements

## Sentences + Relations

- <Query, Response> Q-R pairs (mined or scripted)
- <Question, Answer> Q-A pairs (mined or scripted)
- <Prev Sentence, Sentence, Next Sentence>: causality, discourse, context, or other relations among sentences

## Documents/Paragraphs + Relations

- Emails threads & activity graph; web pages & web graph; log data



Links

Links

Links



Representation

## Representations

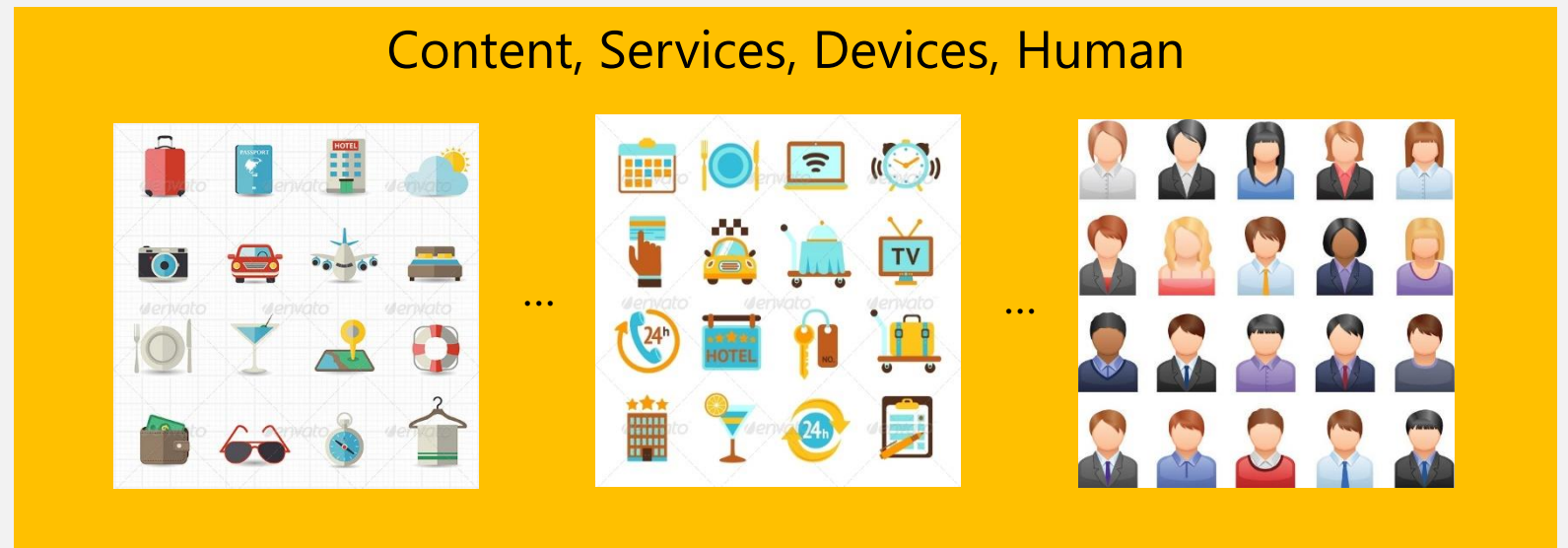
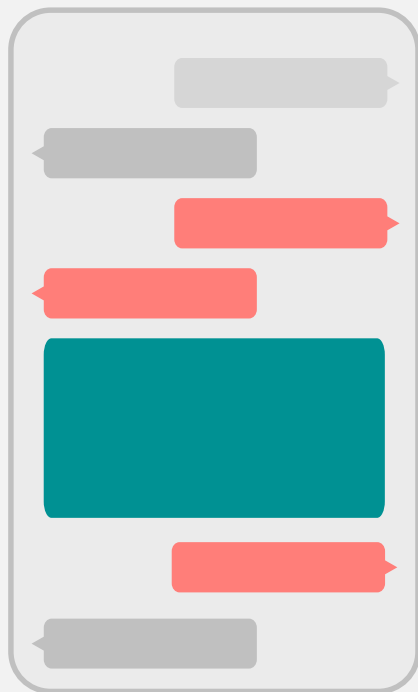
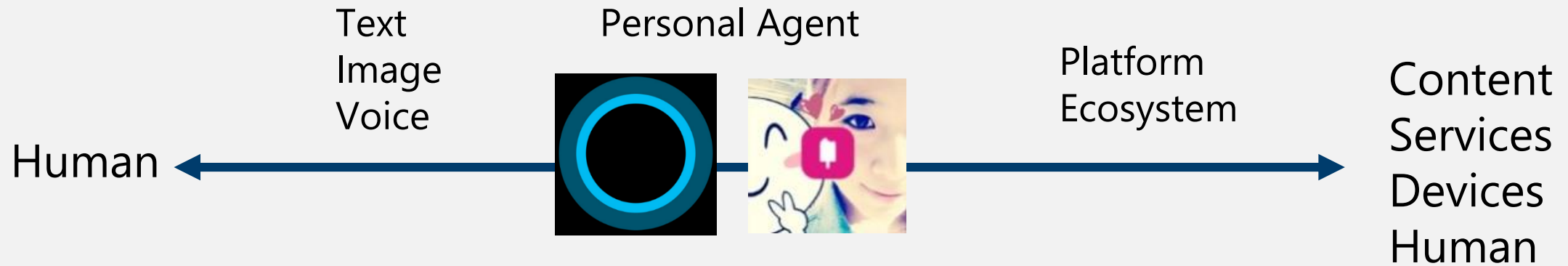
- Finer granular relations & annotations/labels
- Increasing structure
- Higher abstraction

## Strategies

- Bigger data win
- Higher quality win
- Strongly supervised training win
- Make it easy to share training data & knowledge across teams



# Conversation as a Platform (CaaP)

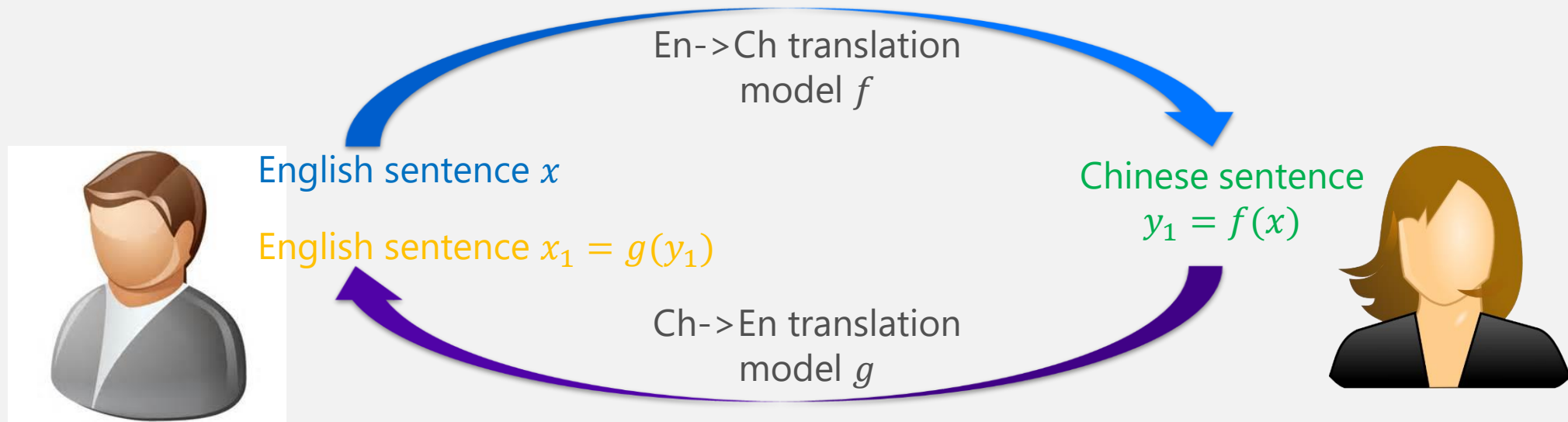


# Building a Self-teaching Machine

- Today's AI technologies require huge amount of human-labeled training data
- Let AI teach itself from unlabeled data through a dual-learning game
  - Many real applications involve two dual AI tasks
  - Reinforcement learning is used to improve models from feedback signals
- Can be extended to multiple associated tasks

<b>Application</b>	<b>Primal task</b>	<b>Dual task</b>
Machine translation	Translation from language A to B	Translation from language B to A
Speech processing	Speech recognition	Text to speech
Image understanding	Image captioning	Image generation
Conversation	Query/Question	Response/Answer
Search engine	Search	Query/keyword suggestion

# Dual Learning Game for Machine Translation



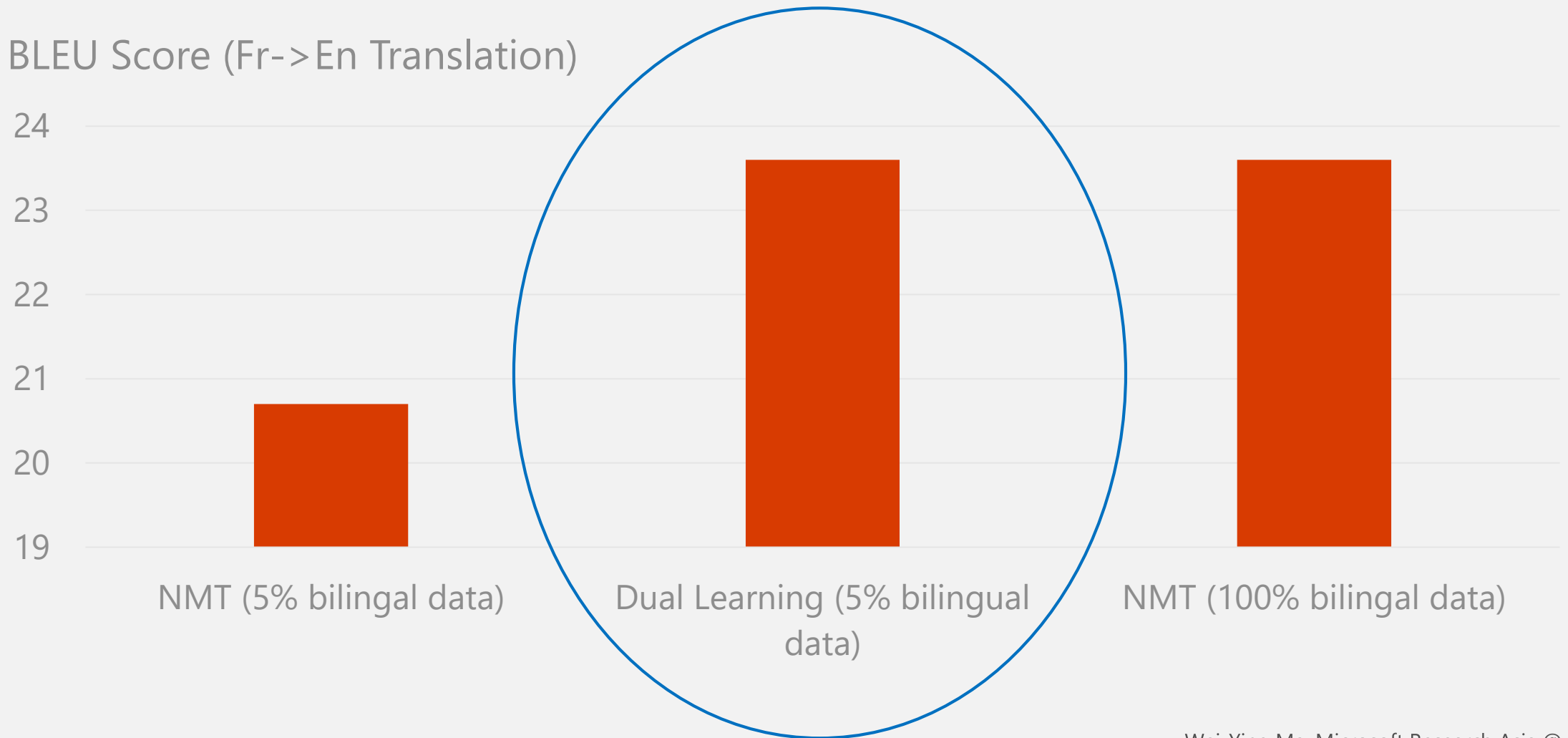
Feedback signals during the loop:

- $s(x, x_1)$ : BLEU score of  $x_1$  given  $x$
- $L(y)$  and  $L(x_1)$ : Likelihood and language model of  $y_1$  and  $x_1$

Reinforcement learning is used to improve the translation models from these feedback signals

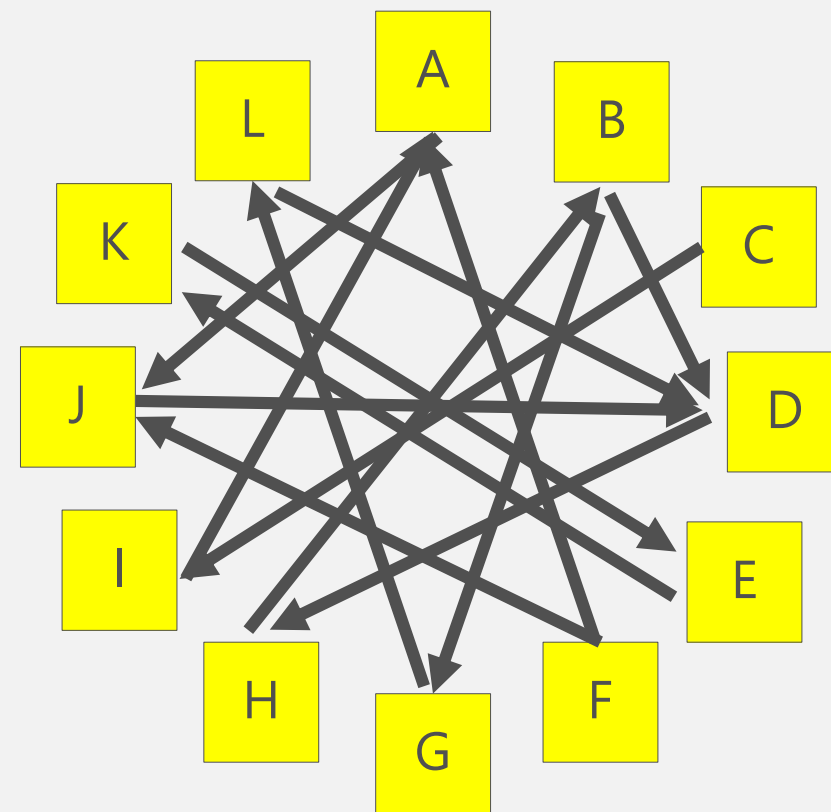
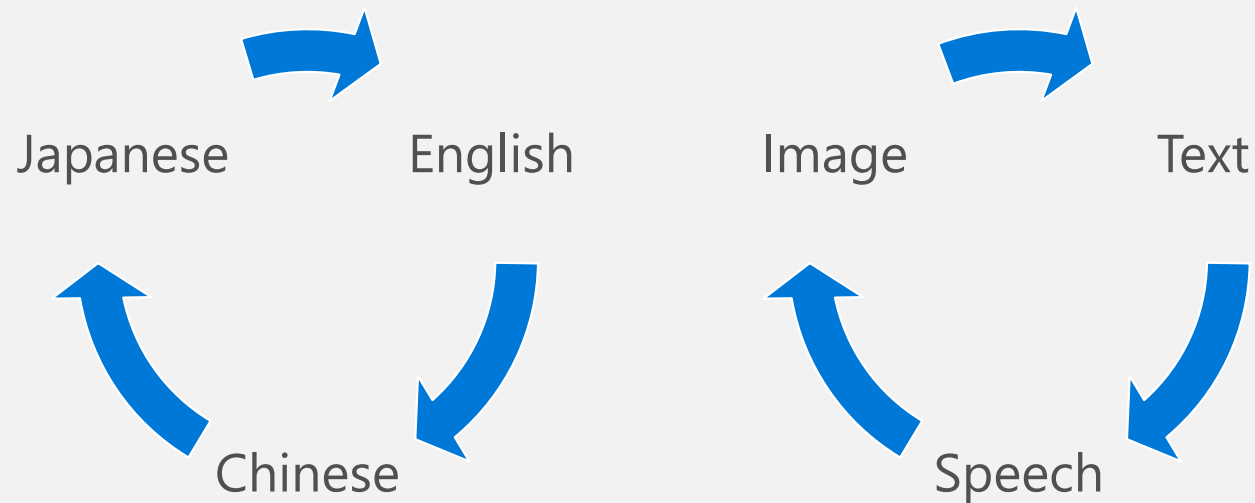
# Results on Machine Translation (NIPS 2016)

BLEU Score (Fr->En Translation)





# The idea of dual-learning can be extended to multiple tasks



# Summary

1. The 4<sup>th</sup> industrial revolution – ever more intelligent software written by big data through machine learning
2. Machine comprehension of text – lessons from search engine to conversational agent
3. Let AI teach itself from unlabeled data through a dual-learning game – some initial result on machine translation

