

Evolution of AI



Olivia Klose
SDE, Microsoft



Marcel Tilly
Senior Program Manager,
Microsoft AI & Research

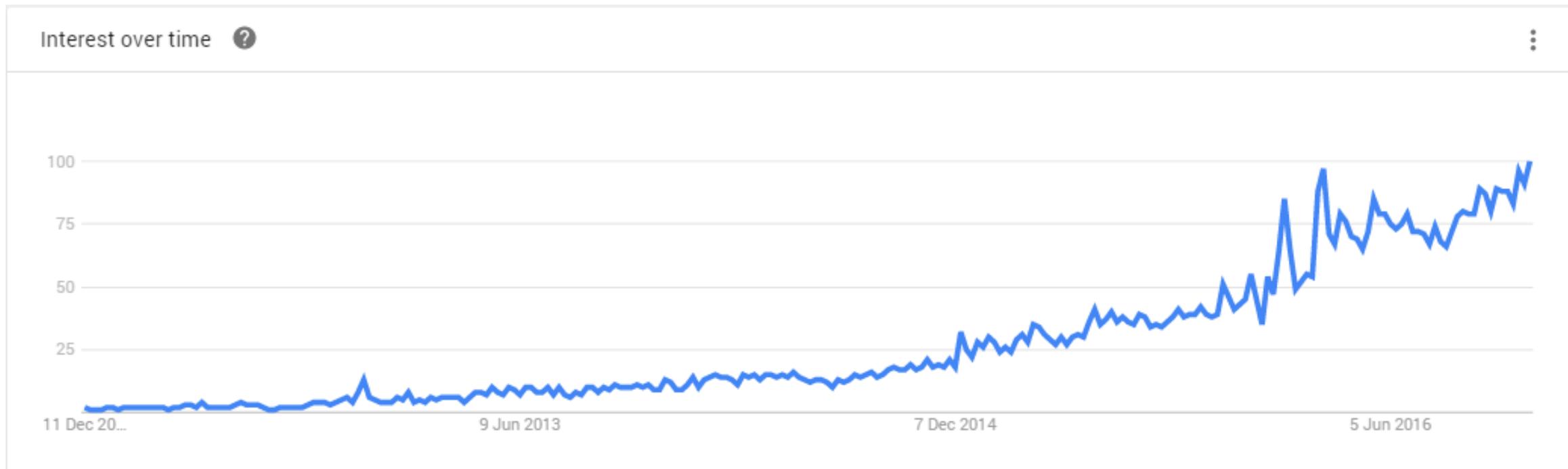
NVIDIA stock price



Technical Summit 2016
connecting.technologies.

#msts2016 

Deep Learning



Technical Summit 2016
connecting.technologies.

#msts2016 

Scientists See Promise in Deep-Learning Programs

By JOHN MARKOFF NOV. 23, 2012



A voice recognition program translated a speech given by Richard F. Rashid, Microsoft's top scientist, into Mandarin Chinese. Hao Zhang/The New York Times

Using an artificial intelligence technique inspired by theories about how the brain recognizes patterns, technology companies are reporting startling gains in fields as diverse as computer vision, speech recognition and the identification of promising new molecules for designing drugs.

NEWS DESK

IS "DEEP LEARNING" A REVOLUTION IN ARTIFICIAL INTELLIGENCE?



By Gary Marcus, NOVEMBER 25, 2012

Can a new technique known as deep learning revolutionize artificial intelligence, as yesterday's front-page article at the *New York Times* suggests? There is good reason to be excited about deep learning, a sophisticated "machine learning" algorithm that far exceeds many of its predecessors in its abilities to recognize syllables and images. But there's also good reason to be skeptical. While the *Times* reports that "advances in an artificial intelligence technology that can recognize patterns offer the possibility of machines that perform human activities like seeing, listening and thinking," deep us, at best, only a small step toward the creation of truly intelligent Deep learning is important work, with immediate practical applica not as breathtaking as the front-page story in the *New York Times* suggest.



By Roger Parloff

Illustration by Justin Metz
SEPTEMBER 28, 2016, 5:00 PM EST

WHY DEEP LEARNING IS SUDDENLY CHANGING YOUR LIFE

Decades-old discoveries are now electrifying the computing industry and will soon transform corporate America.

Over the past four years, readers have doubtlessly noticed quantum leaps in the quality of a wide range of everyday technologies.

Most obviously, the speech-recognition functions on our smartphones work much better than they used to. When we use a voice command to call our spouses, we reach them now. We aren't connected to Amtrak or an angry ex.

A Brief History of AI

Pre-millennial



IN THIS BUILDING DURING THE SUMMER OF 1956
 JOHN MCCARTHY (DARTMOUTH COLLEGE), MARVIN L. MINSKY (MIT),
 NATANIEL ROCKEFSTER (IBM), AND CLAUDE SHANNON (BELL LABORATORIES)
 CONDUCTED

**THE DARTMOUTH SUMMER RESEARCH PROJECT
 ON ARTIFICIAL INTELLIGENCE**

FIRST USE OF THE TERM "ARTIFICIAL INTELLIGENCE"
 FOUNDING OF ARTIFICIAL INTELLIGENCE AS A RESEARCH DISCIPLINE

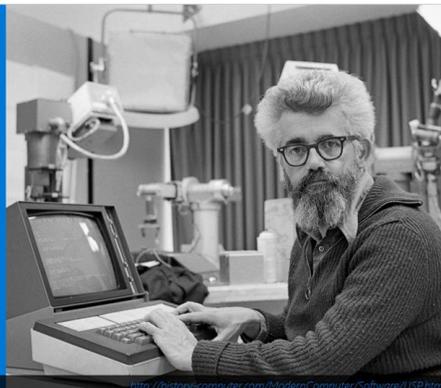
*"In principle on the basis of the conjecture
 that every aspect of learning or any other feature of intelligence
 can in principle be so precisely described that a machine can be made to simulate it."*

IN COMMEMORATION OF THE PROJECT'S 50th ANNIVERSARY
 JULY 13, 2006

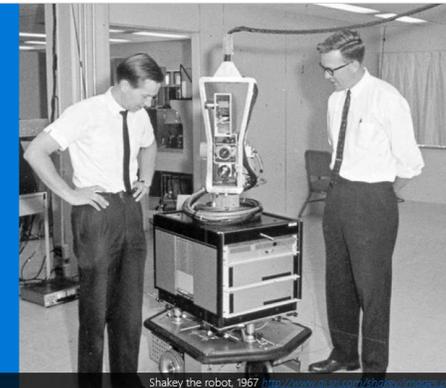
Reasoning



Knowledge Representation



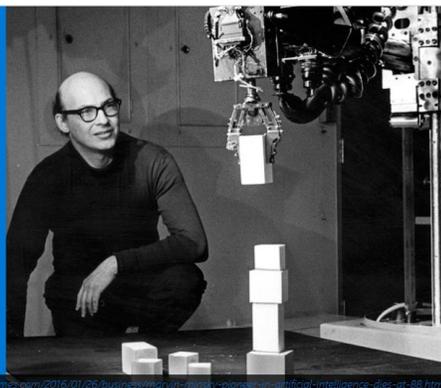
Planning



Natural Language Processing

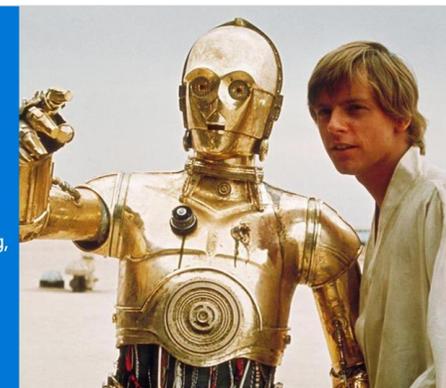


Perception



Generalised Intelligence

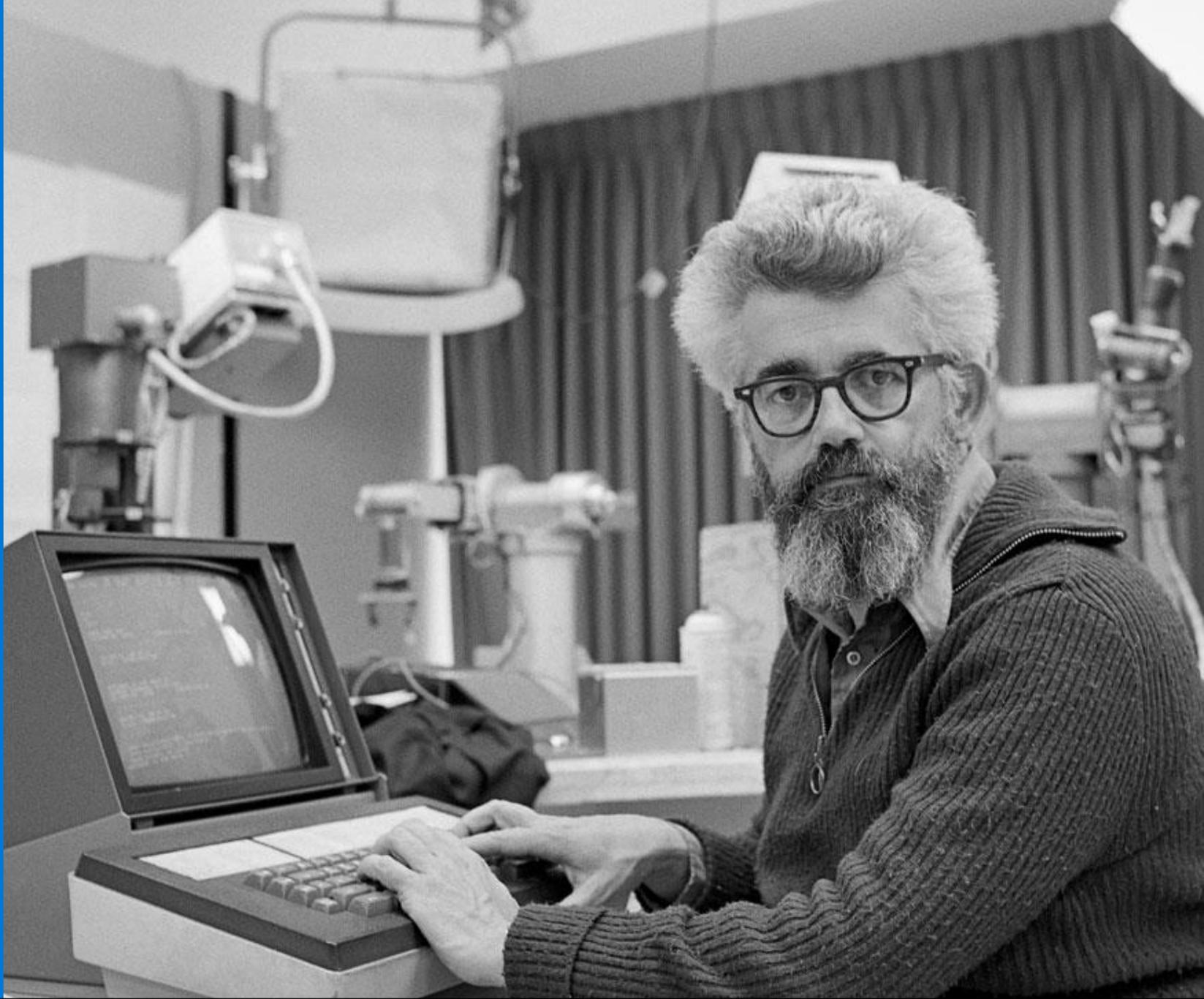
(emotional intelligence, creativity, moral reasoning, intuition,...)



Reasoning



Knowledge Representation



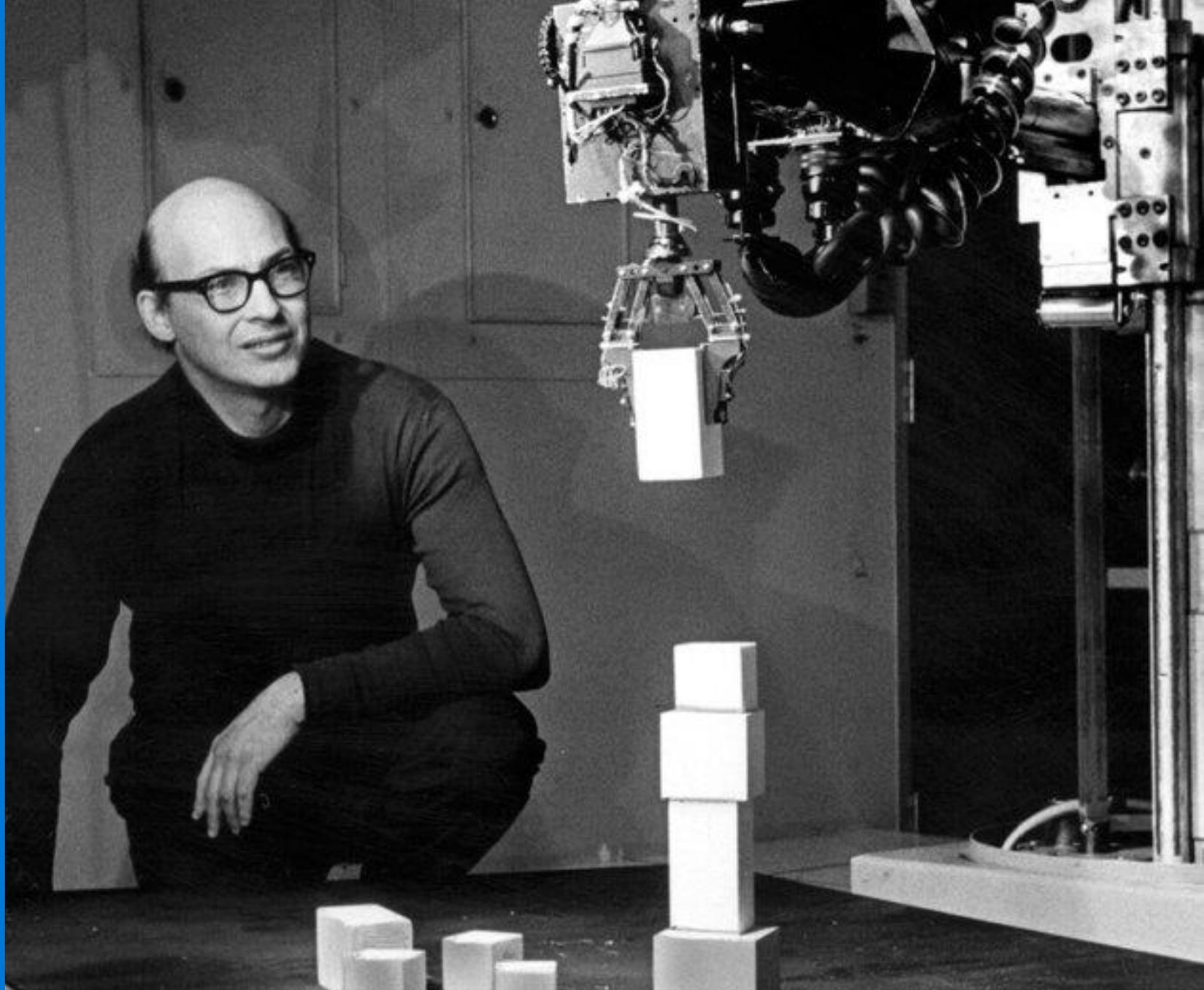
Planning



Natural Language Processing

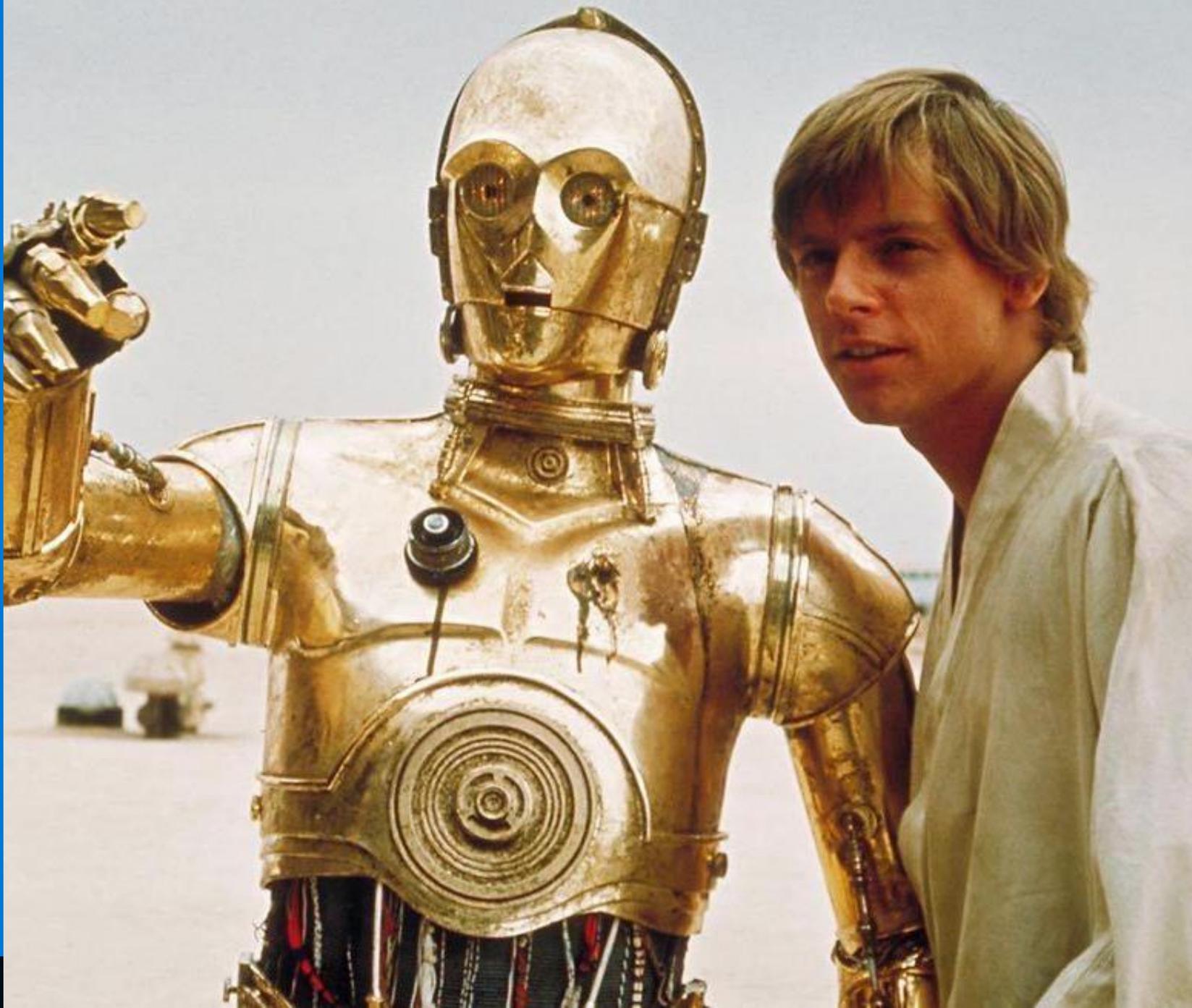


Perception



Generalised Intelligence

(emotional intelligence, creativity, moral reasoning, intuition,...)





Technical Summit 2016
connecting.technologies.

#msts2016 

AI Winter #1 – Machine Translation (1960s)

Automated Language
Processing Advisory
Committee:

Progress is slow, kills machine
translation funding for a decade

1954

1966

Georgetown-IBM
experiment
Successful translation of 60
Russian sentences into English

[IBM Archives](#) > [Exhibits](#) > [IBM 701](#) > [701 Reference room](#) >

701 Translator

IBM Press release, January 8, 1954

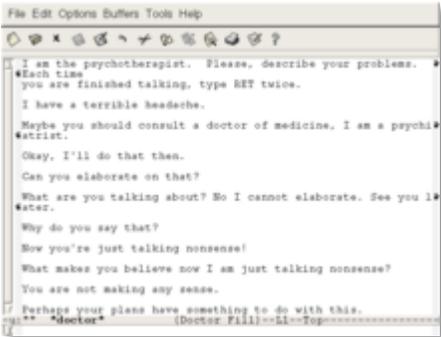
New York, January 7..... Russian was translated into English by an electronic "brain" today for the first time.

Brief statements about politics, law, mathematics, chemistry, metallurgy, communications and military affairs were submitted in Russian by linguists of the Georgetown University Institute of Languages and Linguistics to the famous 701 computer of the International Business Machines Corporation. And the giant computer, within a few seconds, turned the sentences into easily readable English.

A girl who didn't understand a word of the language of the Soviets punched out the Russian messages on IBM cards. The "brain" dashed off its English translations on an automatic printer at the breakneck speed of two and a half lines per second.

"Mi pyeryedayem mislyi posryedstvom ryechyi," the girl punched. And the 701 responded: "We transmit thoughts by means of speech."

AI Winter #2 – Golden Age (1970s)



ELIZA
Pattern matching
Joseph Weizenbaum, MIT

Marvin Minsky:
„In from 3-8 years we will have a machine with the general intelligence of an average human being.“

1959

1964-8

1968-70

1970

1974

„Reasoning as search“ –
Newell & Simon Geometry
Theorem prover, Arthur
Samuels checkers player

Micro-worlds – SHRDLU
Terry Winograd, MIT

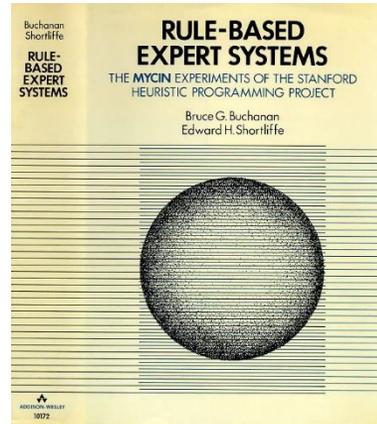


Lighthill report:
„...utter failure of AI to
achieve its grandiose
objectives“

AI Winter #3 – Expert Systems (1980s)

MYCIN

Diagnose infectious blood diseases
Edward Shortliffe, Stanford



Startups
e.g. Symbolics
Lisp Machines

Collapse of
Symbolics

1965

1972

1980

1987

1990

DENDRAL

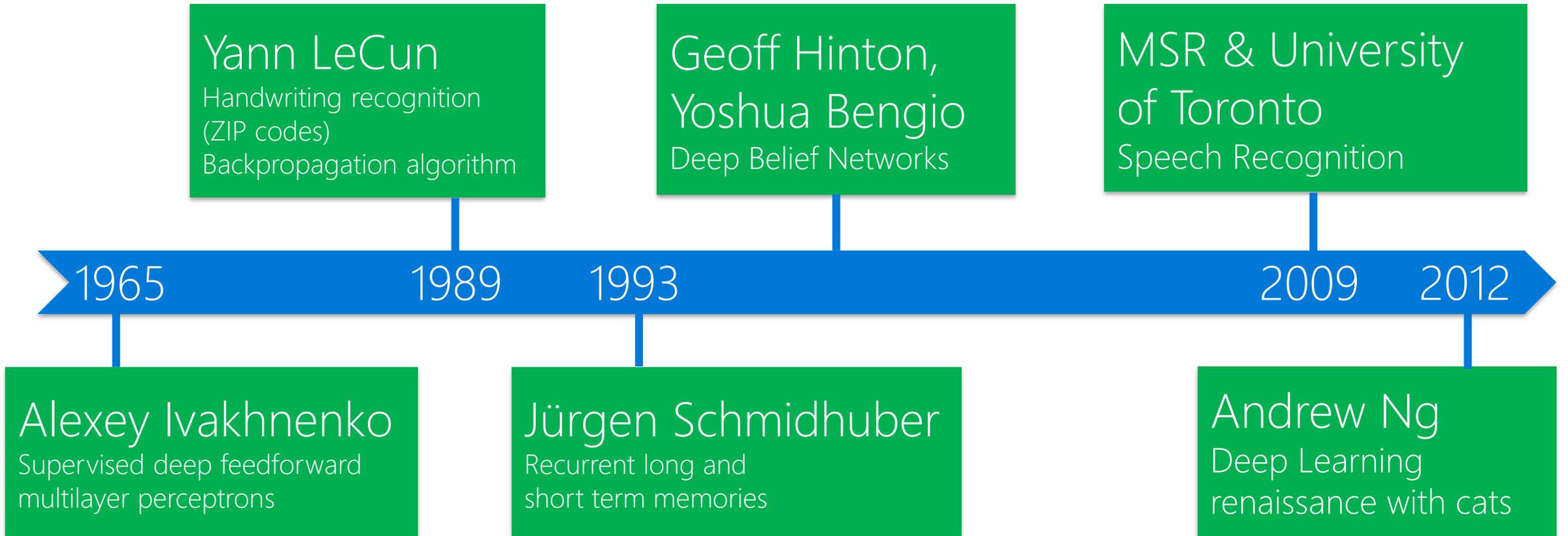
Identify chemical compounds from spectrometer data
Edward Feigenbaum, Carl Djerassi, et. al.

XCON (Expert Configurator)
Automatically select VAX minicomputer components based on user needs
John McDermott

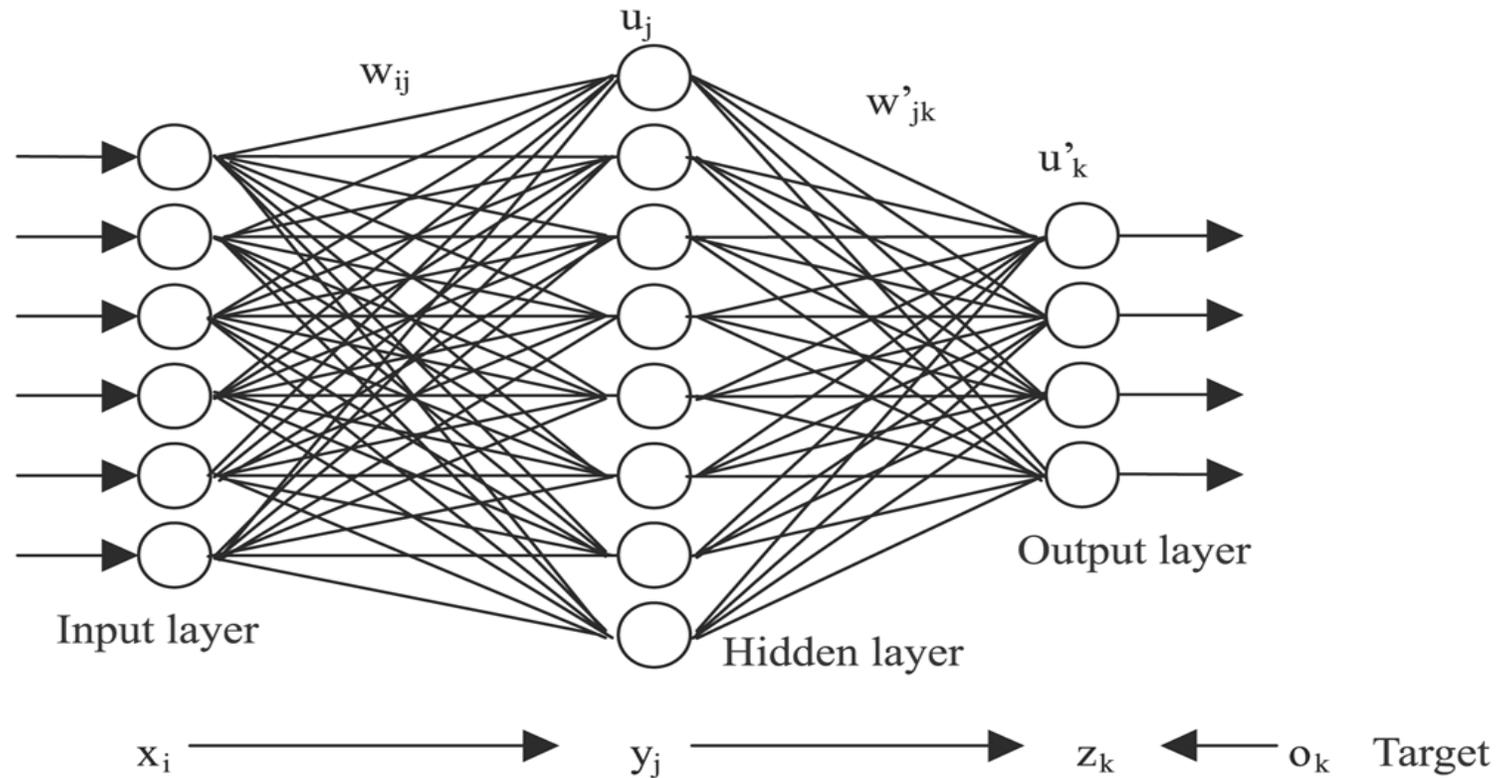
The Integrated Reasoning Shell
IBM

Deep Learning

Beginnings of Deep Learning



What is Deep Learning?



Why so successful?

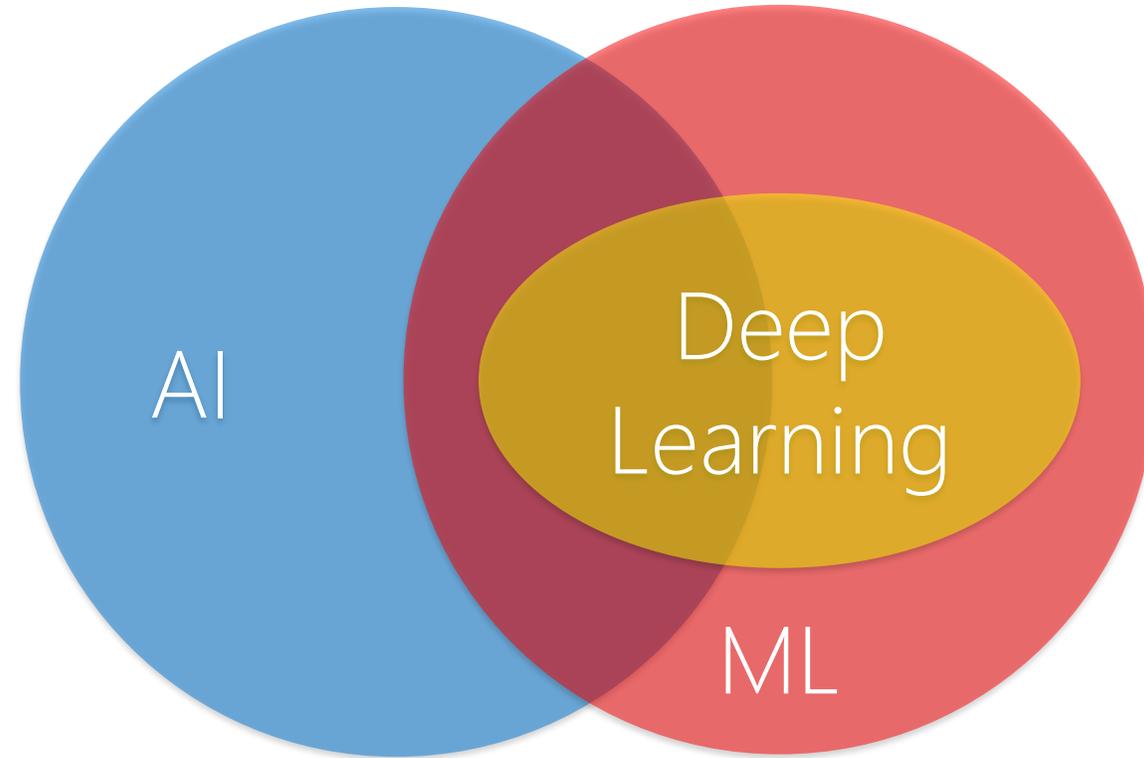
Scale in
Compute

Scale in
Data

Better
Algorithms

More
Investment

Our View on AI, ML & Deep Learning



Deep Dive in Deep Learning

Let's see what this all means

Image Classification

Classification



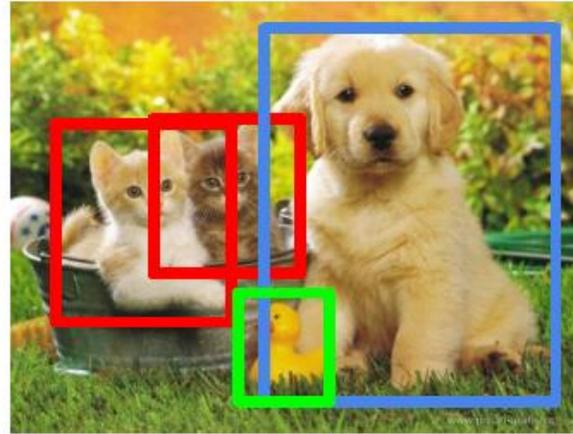
CAT

**Classification
+ Localization**



CAT

Object Detection



CAT, DOG, DUCK

**Instance
Segmentation**



CAT, DOG, DUCK

Single object

Multiple objects

The Microsoft Cognitive Toolkit

A free, easy-to-use, open-source, commercial-grade toolkit that trains deep learning algorithms to learn like the human brain.

Open-source, cross-platform toolkit for learning and evaluating **deep neural networks**.

Expresses (nearly) **arbitrary neural networks** by composing simple building blocks into complex **computational networks**

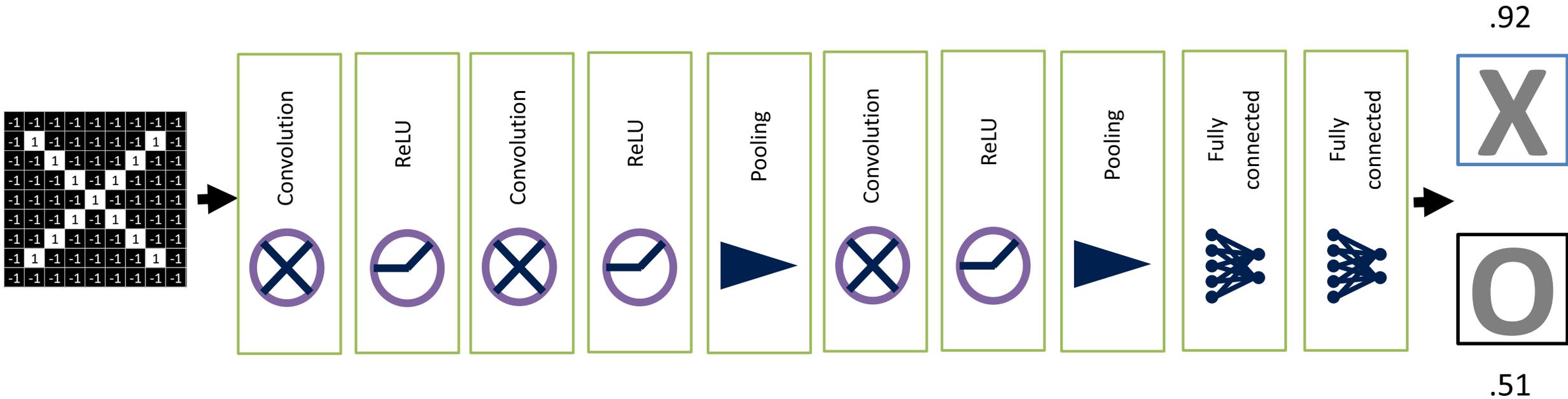
Production-ready: State-of-the-art accuracy, efficient, and scales to multi-GPU/multi-server.

Technical Summit 2016

connecting.technologies.

#msts2016 

Convolutional Neural Networks



Taken from: http://brohrer.github.io/how_convolutional_neural_networks_work.html

Cognitive Toolkit in Action

- Build your own Classifier-

Content of Image



V1:

```
{ "name": "grass", "confidence": 0.9999992847442627 },  
{ "name": "outdoor", "confidence": 0.9999072551727295 },  
{ "name": "cow", "confidence": 0.99954754114151 },  
{ "name": "field", "confidence": 0.9976195693016052 },  
{ "name": "brown", "confidence": 0.988935649394989 },  
{ "name": "animal", "confidence": 0.97904372215271 },  
{ "name": "standing", "confidence": 0.9632768630981445 },  
{ "name": "mammal", "confidence": 0.9366017580032349, "hint": "animal" },  
{ "name": "wire", "confidence": 0.8946959376335144 },  
{ "name": "green", "confidence": 0.8844101428985596 },  
{ "name": "pasture", "confidence": 0.8332059383392334 },  
{ "name": "bovine", "confidence": 0.5618471503257751, "hint": "animal" },  
{ "name": "grassy", "confidence": 0.48627158999443054 },  
{ "name": "lush", "confidence": 0.1874018907546997 },  
{ "name": "staring", "confidence": 0.165890634059906 }]
```

Describe

0.975 "a brown cow standing on top of a lush green field"

0.974 "a cow standing on top of a lush green field"

0.965 "a large brown cow standing on top of a lush green field"

If you train with too many cats,
everything looks like a cat



If there is something the
computer has not seen, then
the computer might think it is
something else



AI's White Guy Problem

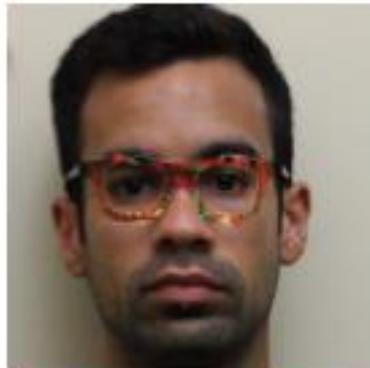
- Google's photo app: Classified images of black people wrongly
- Nikon's camera software: Misread images of Asian people
- HP's web camera software: Difficulty recognizing people with special skin tones



Data problem



Accessorize to a Crime



Technical Summit 2016

connecting.technologies.

#msts2016 

Why *image recognition* sometimes sucks?

Data set is key for ML and DL
Cheating is possible
Context is missing



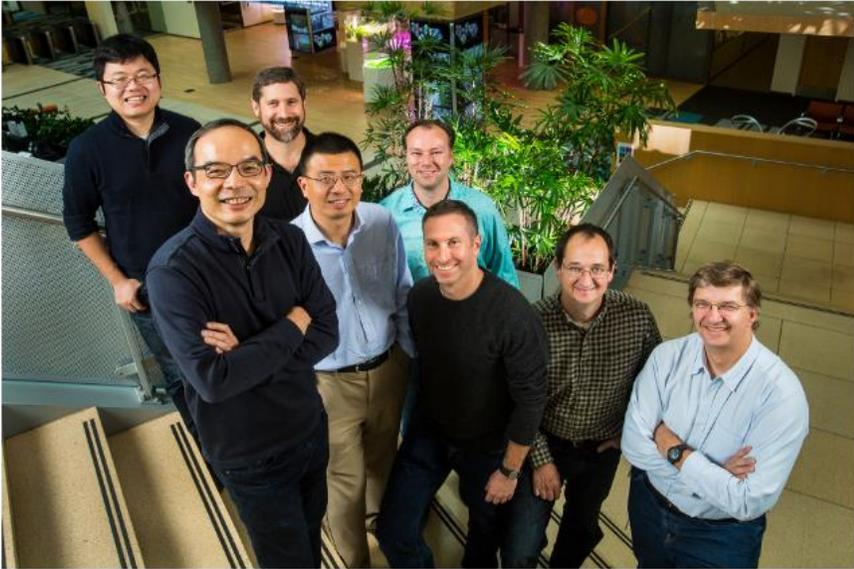
Fei-Fei Li, Associate Professor of
Computer Science at Stanford University

*“We [humans] are terrible at computing
with huge data but we’re great at
abstraction and creativity.”*

<https://www.technologyreview.com/s/602094/ais-language-problem/?set=602166>

Speech Recognition

Historic Achievement: Microsoft researchers reach human parity in conversational speech recognition



Microsoft researchers from the Speech & Dialog research group include, from back left, Wayne Xiong, Geoffrey Zweig, Xuedong Huang, Dong Yu, Frank Seide, Mike Seltzer, Jasha Droppo and Andreas Stolcke. (Photo by Dan DeLong)

5.9% word-error rate – Human Parity

All experiments were run on CNTK

[W. Xiong, J. Droppo, X. Huang, F. Seide, M. Seltzer, A. Stolcke,

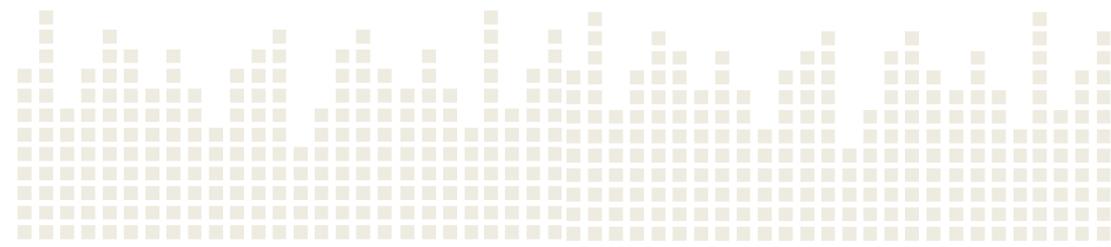
D. Yu, G. Zweig: "The Microsoft 2016 Conversational Speech Recognition System," <http://arxiv.org/abs/1609.03528>]

<http://blogs.microsoft.com/next/2016/10/18/historic-achievement-microsoft-researchers-reach-human-parity-conversational-speech-recognition>

Technical Summit 2016
connecting.technologies.

#msts2016 

Speech Customization



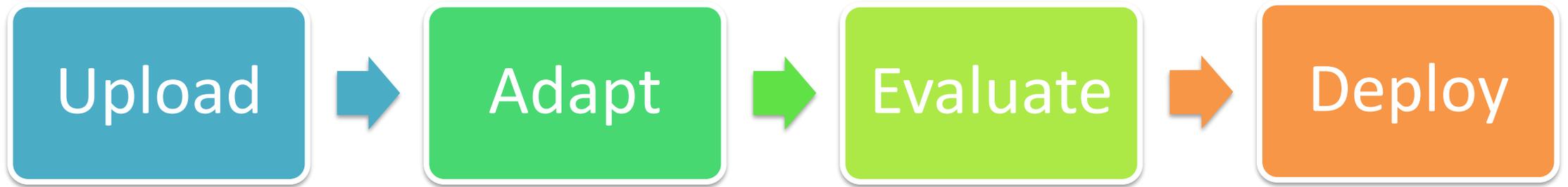
Problem with Speech APIs

- Customers may have very specific needs
- - Specific ages, accents, noise conditions -
- General purpose endpoints can not be optimized

Solution

- Package speech recognition customization as a simple service
- Data in, models out. Minimal configuration.
- Language Model and Acoustic Model adaptation.
- Allow third parties to create speech recognition endpoints.

Custom Speech Recognition Demo



Upload in-domain data
(wav, txt)

Customize existing
models
(AM, LM)

Evaluate the
new models

Deploy a recognition
endpoint

But what is about that?

Quoc Le, Researcher at Google

User: "How many legs does a cat have?"

Computer: "Four, I think."

User: "How many legs does a centipede have?"

Computer: " Eight."





Microsoft's [...] Tay experiment shows the hidden dangers of AI!

 [Follow](#)

Why *talking* to my computer still sucks?

My Computer struggles with **context**

- Two words can have the same meaning
- One word can have different meaning (**jaguar vs. jaguar**)

My computer struggles with **semantics**

- Speech recognitions might be good in recognition but not understanding

My computer struggles with **conversations**

- Speech recognitions might be good in recognition but not understanding



Where are we now?
What's next?

We really mean Understanding?

Take pictures ≠ see

To hear ≠ to listen

Emotions, Actions, Intentions,
Relations



Reasoning



Knowledge Representation



Planning

INSIDE HOTZ'S
ACURA ILX

A 21.5-INCH SCREEN
DISPLAYS THE CAR'S
AUTONOMOUS TECHNOLOGY
(THE OPERATING SYSTEM IS
UBUNTU LINUX), ALONG WITH
FEEDS FROM THE CAMERAS
AND LIDAR SYSTEMS

HOTZ ADDED A JOYSTICK
TO THE CAR'S CENTER
CONSOLE. A PULL OF THE
TRIGGER ENGAGES THE
SELF-DRIVING SYSTEM

Natural Language Processing



university of toronto.

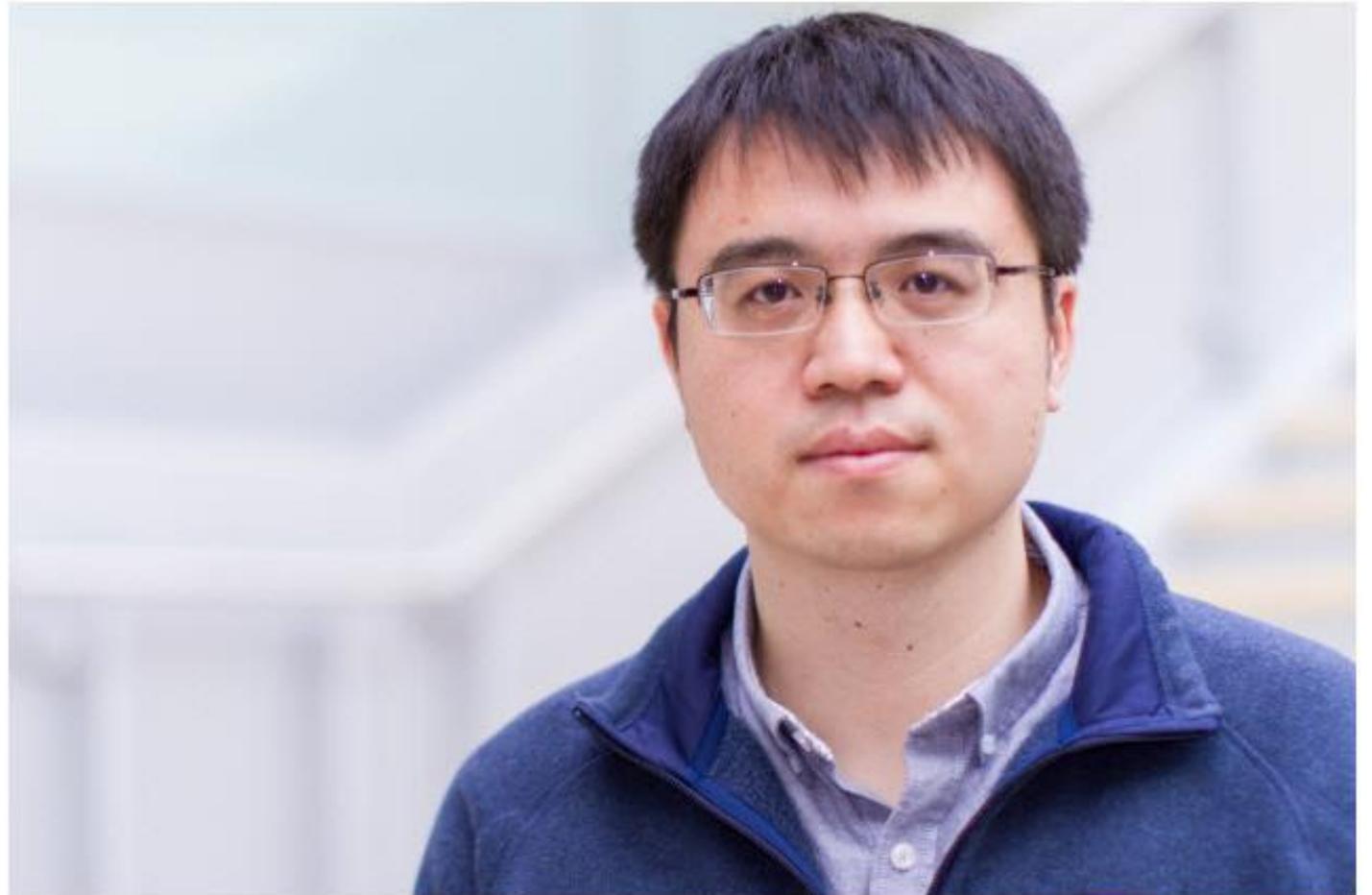
Came together to develop with another breakthrough in the field of speech recognition research.

The idea that they had was to use a

Recognizability: 90%

Perception

Microsoft researchers win ImageNet computer vision challenge



Jian Sun, a principal research manager at Microsoft Research, led the image understanding project. Photo: Craig Tuschhoff/Microsoft.

Posted December 10, 2015 By **Allison Linn**

Generalised Intelligence

(emotional intelligence, creativity, moral reasoning, intuition,...)

Project Malmo, which lets researchers use Minecraft for AI research, makes public debut



Katja Hofmann, left, and Matthew Johnson, right, say Project Malmo can speed AI research. (Photography by Scott Eklund/Red Box Pictures)



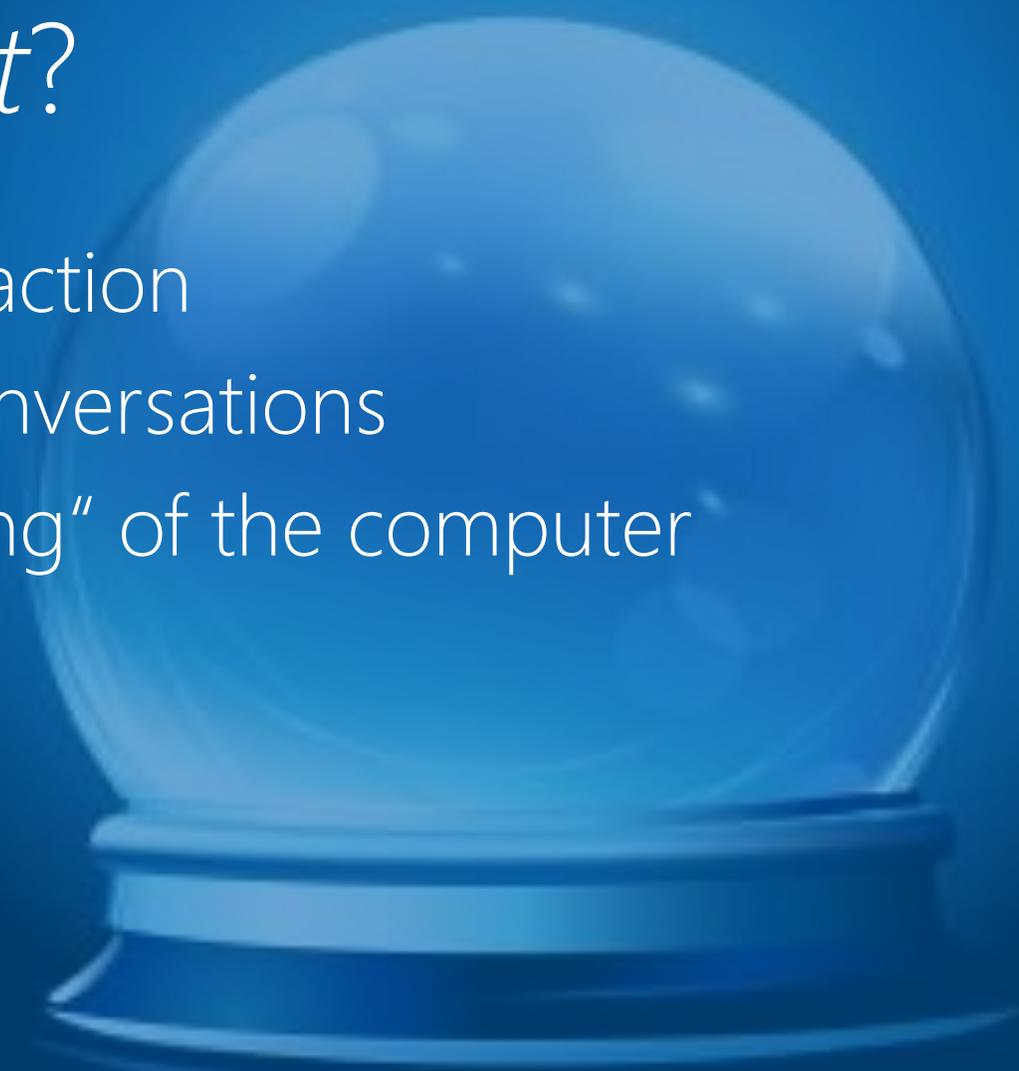
What's *next*?

Multi-modal interaction

Context-aware conversations

Ubiquitous „seeing“ of the computer

Smarter hardware



Further Resources

Cognitive Toolkit <http://aka.ms/cognitivetoolkit>

Azure Notebooks <http://notebooks.azure.com>

CRIS (Custom Recognition Intelligent Service) <http://aka.ms/cris>

Cognitive Services <http://microsoft.com/cognitive>

AI and Deep Learning (Frank Chen) <https://vimeo.com/170189199>

olivia.klose@microsoft.com
marcel.tilly@microsoft.com