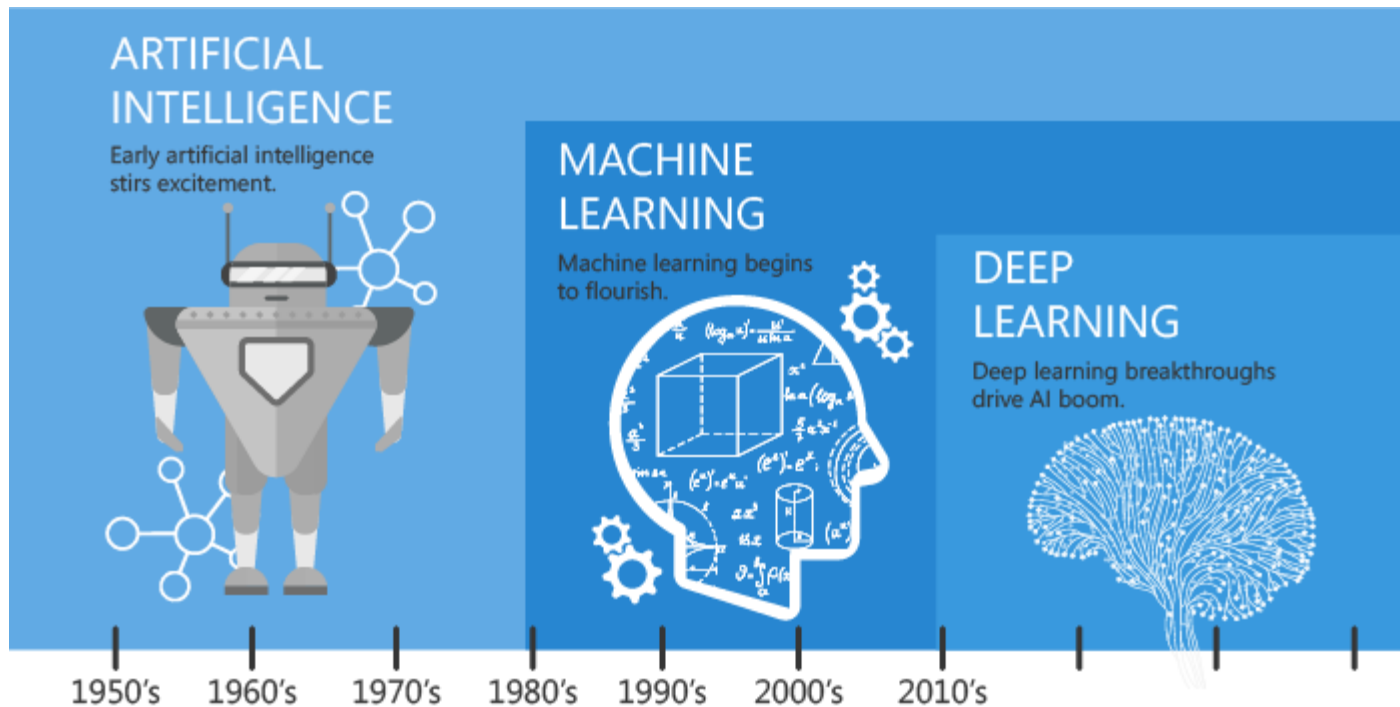


# 딥러닝 이해 및 미디어 응용

---

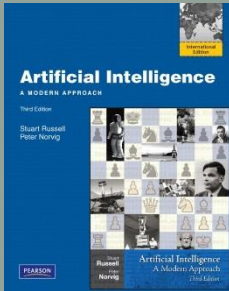
아주대학교 구형일

# 인공지능/기계학습/딥러닝



Since an early flush of optimism in the 1950's, smaller subsets of artificial intelligence - first machine learning, then deep learning, a subset of machine learning - have created ever larger disruptions.

# AI에 관한 4 개의 관점



Humanly

Rationally

Thinking

Thinking Humanly

Thinking Rationally

Acting

Acting Humanly

Acting Rationally

# Acting Humanly...

- **사람처럼 일하는/행동하는 기계**
  - **인공지능은 사람에 의해서 수행될 때 지능이 필요한 일을 수행하는 기계를 만드는 기술이다.** (The art of creating machines that perform functions that require intelligence when performed by people. Kurzweil, 1990)
- **인공지능은 컴퓨터가 (현재는) 사람이 잘하는 일을 할 수 있도록 하는 방법을 연구하는 학문이다.** (The study of how to make computers do things, at which, at the moment, people are better. Rich and Knight, 1991)

# 인공지능?

Tesler's theorem: AI is whatever hasn't been done yet.

테슬러 정리: 인공지능은 아직 실현되지 않은 무언가이다.<sup>[5]</sup>



인공지능이란 무엇인가, 무엇을 지능이라고 부를까를 명확하게 정의하기는 쉽지 않다. 그리고 이는 철학적인 문제가 아니고 이 문제에 어떤 대답을 선호하는가에 따라서 연구 목적과 방향이 완전히 달라진다.

한 가지 대답은 인간의 '지능'을 필요로 하는 일을 컴퓨터가 처리할 수 있으면 그것이 바로 인공지능이라는 것이다. 또 다른 대답은 인간과 같은 방식으로 이해를 할 수 있어야 인공지능이라는 것이다. 이 두 가지 대답 역시 세부적으로는 "지능을 필요로 하는 일이란 무엇인가?" 내지는 "인간과 같은 방식이란 무엇인가?" 라는 질문에 대한 대답에 따라서 서로 다른 여러 종류의 대답을 내포하고 있다. 물론 이 두 가지 대답은 배타적이지는 않다. 인간과 같은 종류의 지능을 가지고 '지능'을 필요로 하는 일도 처리할 수 있는 컴퓨터를 만드는 것은 수많은 컴퓨터 공학자들의 꿈과 희망이었지만, 적어도 단기간에 그런 목표에 도달할 가능성은 희박하다.

만약 '지능을 필요로 하는 일'을 처리하는 것이 인공지능이라고 정의한다면, 인공지능은 인간이 어떤 방식으로 사고하는가를 고민할 필요가 없으며, 감성과 같은 것 또한 고려할 필요가 없다. 모로 가든 서울만 가면 되니까. 이러한 방향의 인공지능 연구에서는 초기에는 전문가가 필요한 일을 복잡한 소프트웨어를 통해서 처리하는 전문가 시스템이 대세였으며, 이러한 전문가 시스템은 실행 방식에서는 일반적인 소프트웨어와 특별한 차별성이 없고 전문가들이 문제를 해결하는 방식을 가능한 한 쉽고 정확하게 소프트웨어에 반영할 수 있는 방법을 제공하는데 주력했다.

# 인공지능?

그 외에도 체스를 두는 것 역시 이 분야에 들어갔고 실제로 너무 간단한 인공지능 부류에 속하지만 체스 두는 기계는 아주 이르게도 인공지능 연구에서 제외되었다. 체스를 연산으로 처리하게 만드는 것은 대단히 어렵다. 수 하나를 더 내다보려면 평균적으로 26배의 연산이 더 필요해지기 때문에 아무리 현대의 컴퓨터가 빠른 속도로 발전하고 있다고 해도 5~6수를 내다보는 것이 고작이며 수십 수를 내다보는 체스 기사들과는 상대가 되지 않는다. 이 때문에 실제 체스 머신들은 지금까지의 체스 기보를 대량으로 입력한 후 그 체스 기보에서 같은 모양이 나온 적이 있는지를 하나하나 대조하는 방식으로 처리한다. 세계 챔피언을 이긴 IBM의 체스 머신은 7만개가 넘는 기보를 이용했다고 한다. 하지만 이것은 더이상 인공지능의 분야가 아닌 그냥 데이터 병렬처리를 빠른 속도로 해낼 수 있는 슈퍼컴퓨터의 성능 과시용에 불과하다. 즉, 복잡한 지능을 구현한 게 아니라 고등사고로 할 수 있는 무수히 많은 일에서 조각 하나를 따와 펼칠 수 있는 구조를 만들었을 뿐이다. 지능이라는 단어의 정의에 따라서 이것도 인공지능이라고 쳐주는 교수/학자도 있고 빼는 학자도 있지만 더 이상 연구 가치가 없다는 점은 바뀌지 않는다.

인공지능 연구에서 상술하였듯이 컴퓨터가 체스 같은 특정 작업을 능숙하게 해낼 수 있게 되면 그 작업을 인공지능의 조건에서 바로 빼 버리는 경향이 있는데, 인공지능적인 발전의 성과라기보다는 그냥 컴퓨터의 성능 향상으로 이루어진 결과이기 때문에 그렇다. 인공지능적인 연구 가치가 떨어지기 때문.

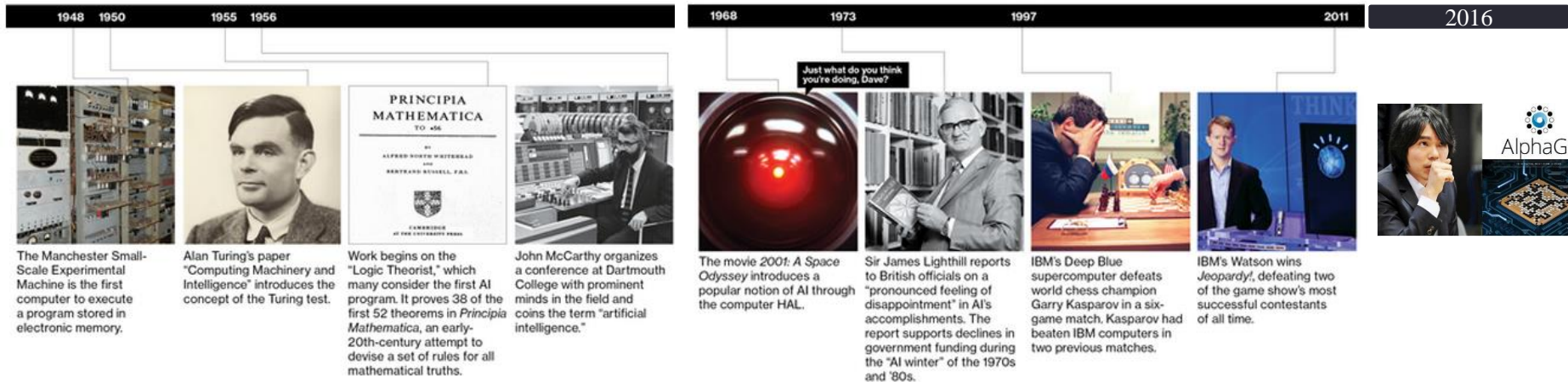
# Artificial Flavor?



# AI 역사

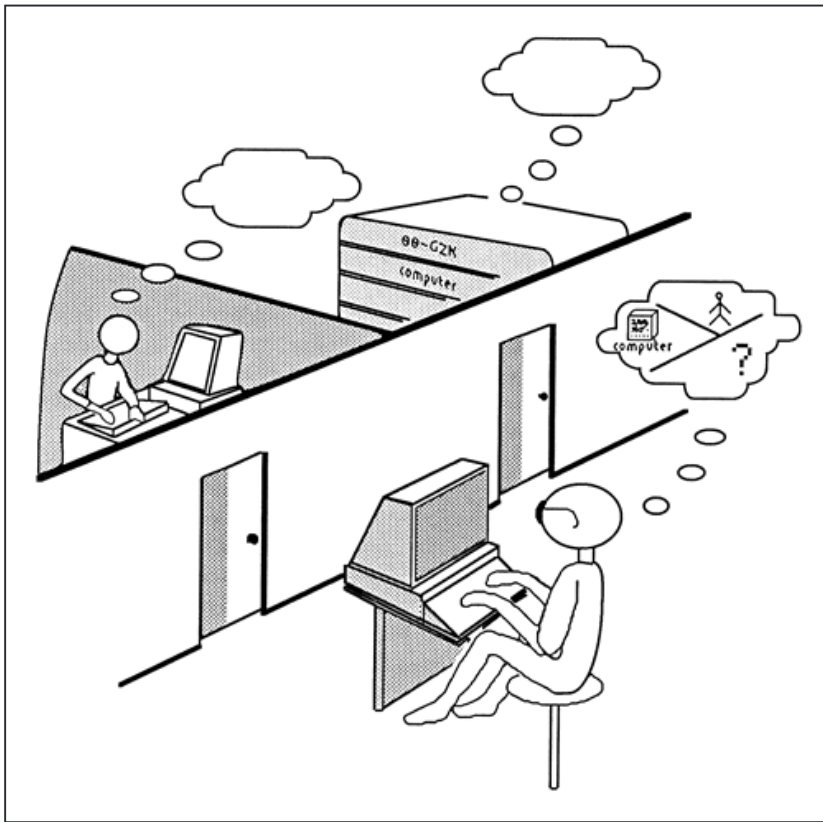
MIT  
Technology  
Review

## AI's Evolution





# 튜링테스트



B.J. Copeland 2000

- 튜링이 1950년 에 제안
- 질문자(interrogator)가, 질문에 대한 대답을 바탕으로 사람과 기계(로봇)을 구별할 수 없다면 기계가 지능을 가지고 있다고 하자.

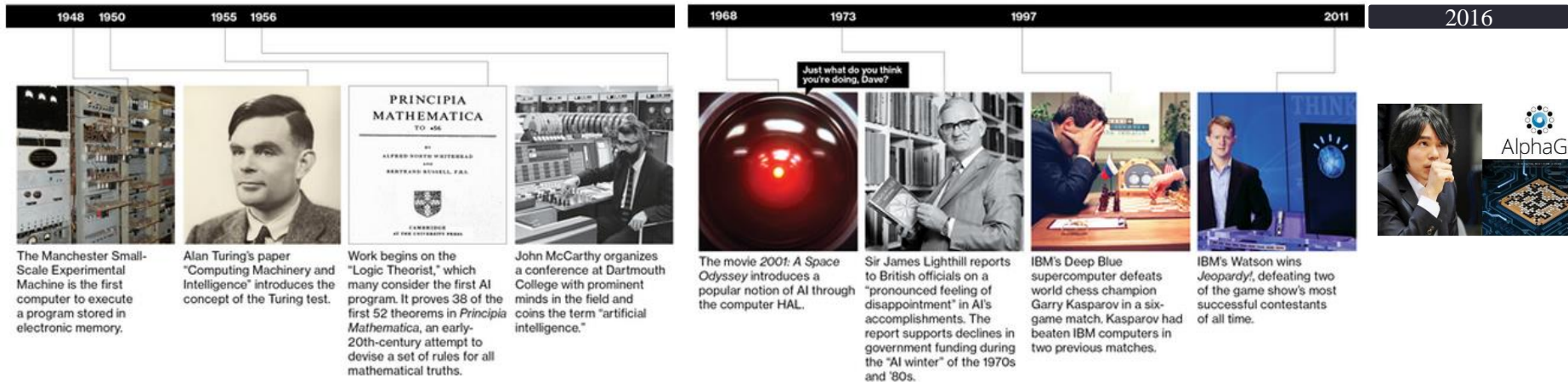


Alan Turing

# AI 역사

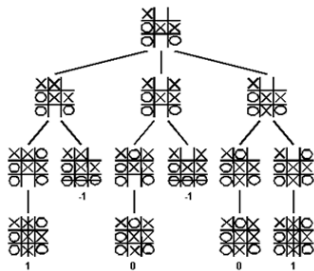
MIT  
Technology  
Review

## AI's Evolution



# 딥블루

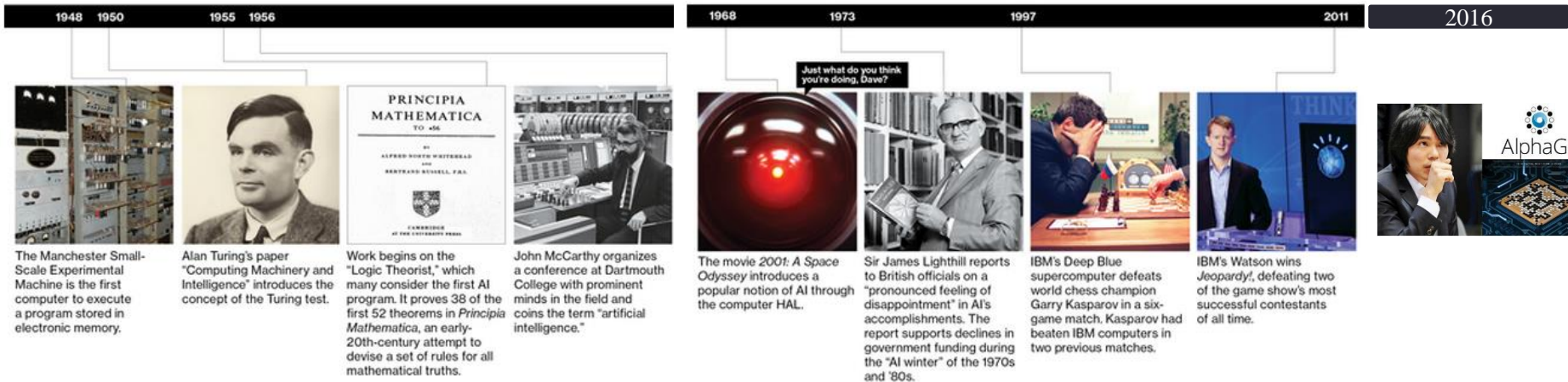
- 딥블루 vs 게리 카스파로프, 1997
  - Deep Blue vs Kasparov
    - $3\frac{1}{2}$  vs  $2\frac{1}{2}$
  - Brute-force search power
    - 6~8 수를 내다봄



# AI 역사

MIT  
Technology  
Review

## AI's Evolution



# IBM Watson 슈퍼컴퓨터

- 질문예시)

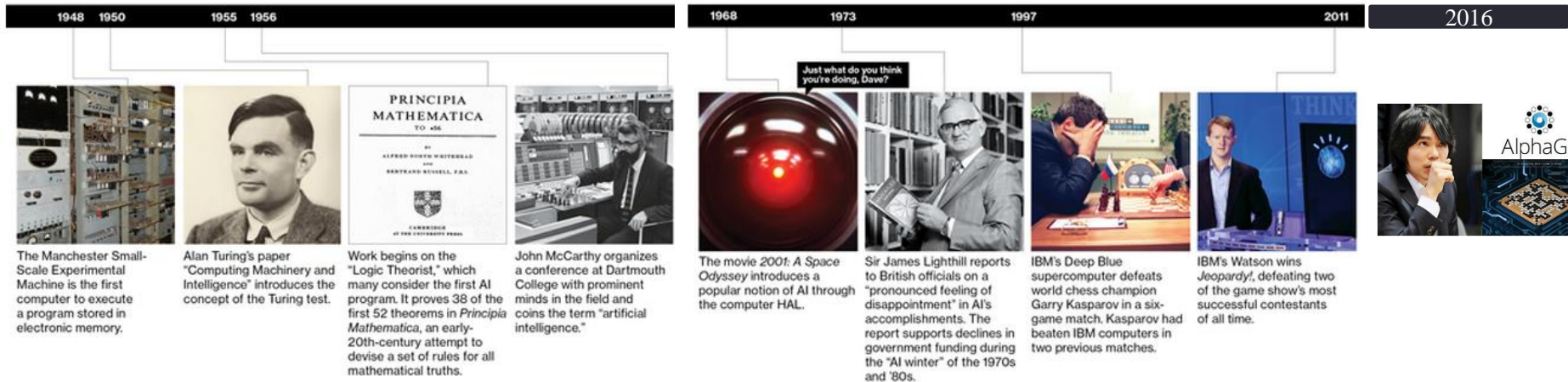
- Kathleen Kenyon's excavation of this city mentioned in Joshua showed the wall had been repaired 17 times
  - WHAT is "Jericho"
- This child star got his first on-screen kiss in "MY GIRL"
  - WHO is "Macaulay Culkin"



# AI 역사

MIT  
Technology  
Review

## AI's Evolution





# 알파고



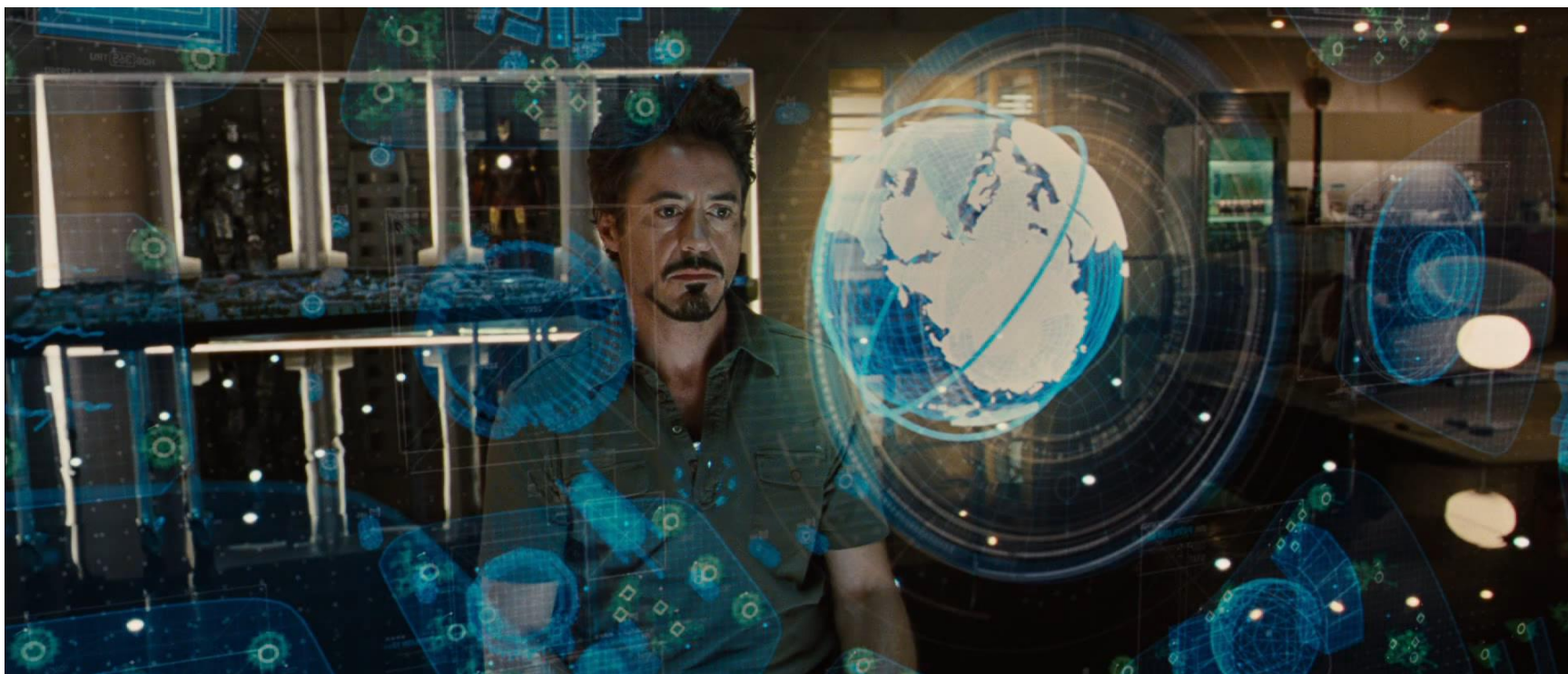
# 머신러닝

---

인공신경망을 중심으로...



# 인공지능 제스처 인식 시스템을 만들자!



# 손동작 인식



“Gesture A”

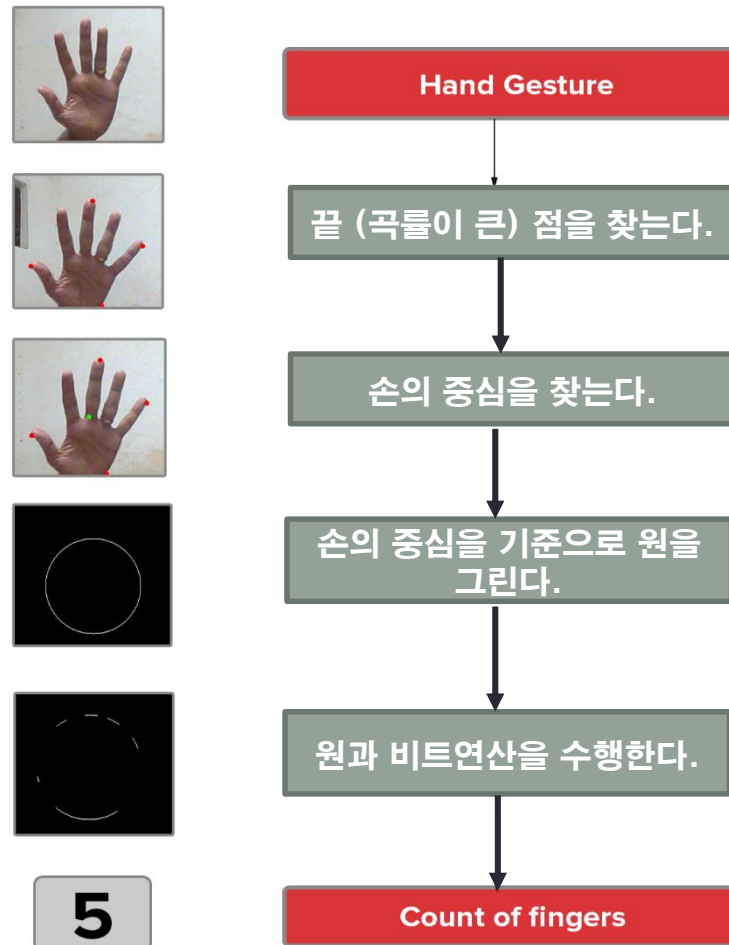


“Gesture B”

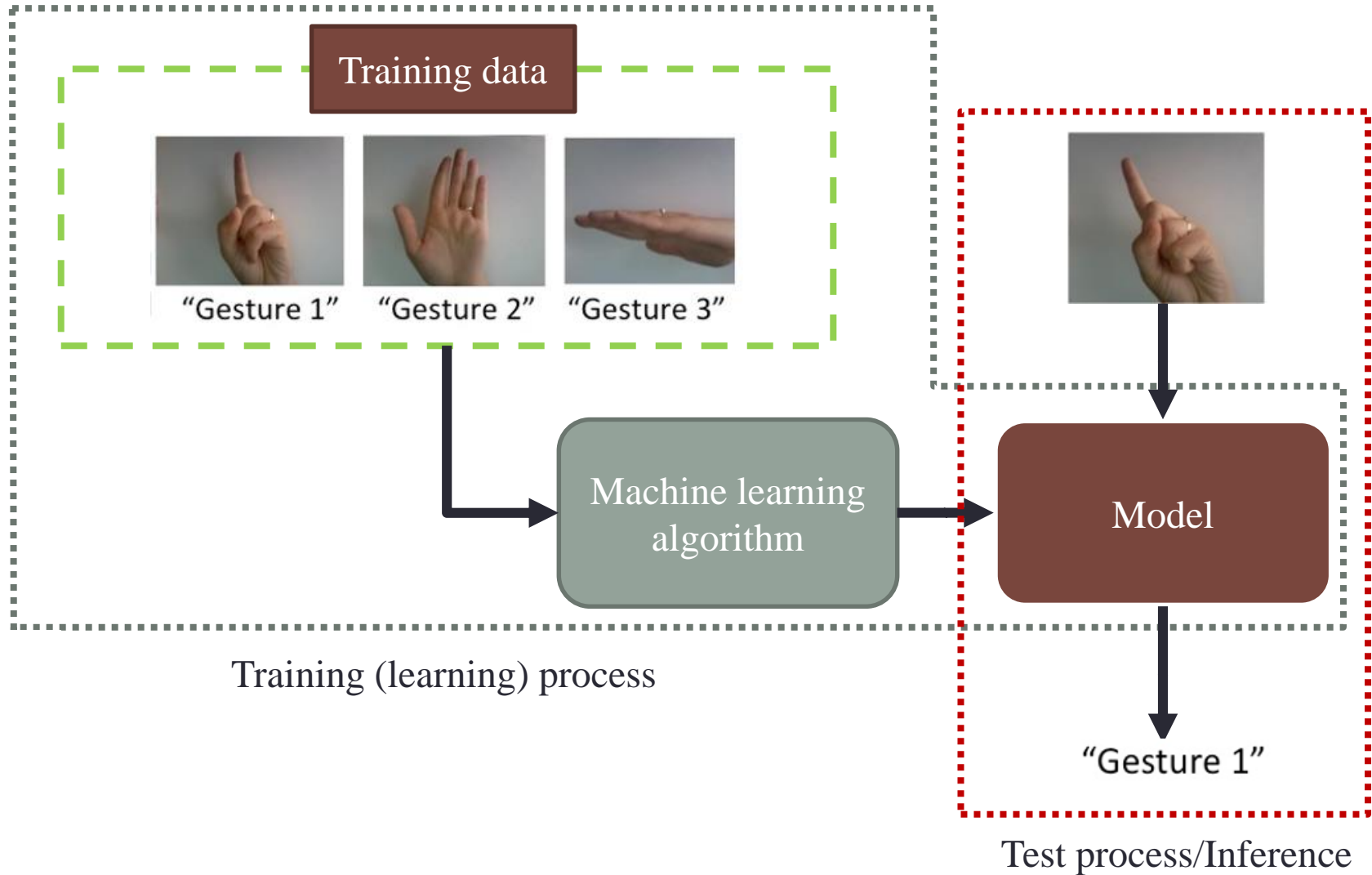


“Gesture C”

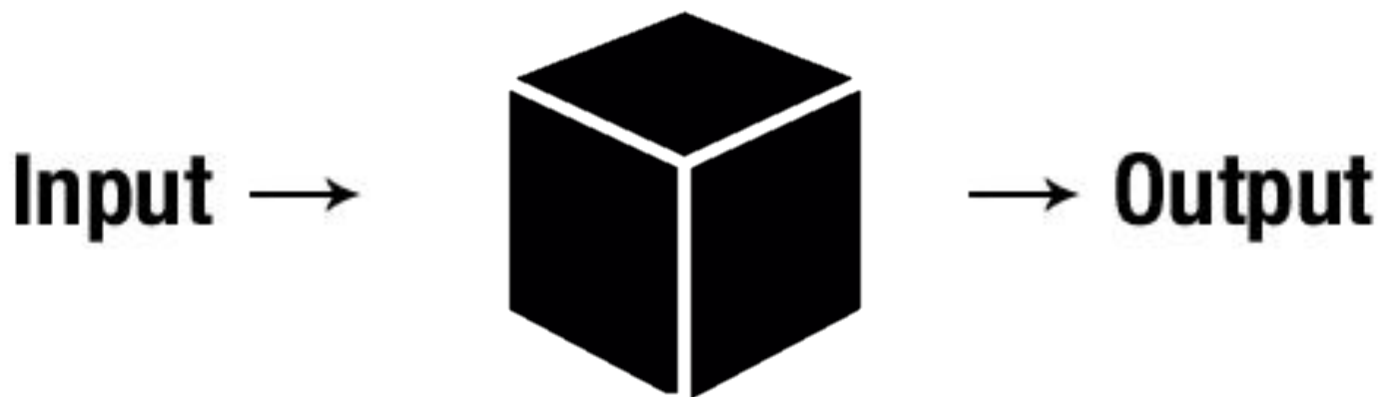
# 접근법 1: 내 경험/직관을 바탕으로 로직을 설계하자



# 접근법 2: 기계학습

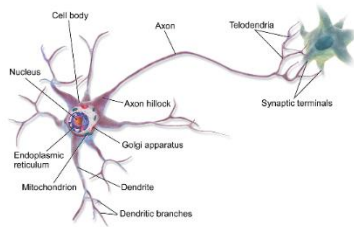
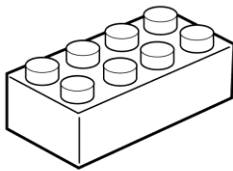


# Black-box approach

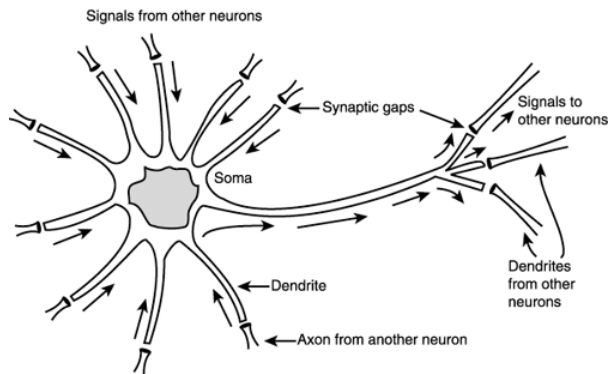


- 영상, 비디오, 음성 모두 벡터(숫자들의 어레이)로 표현할 수 있음
- 출력도 벡터로 표현할 수 있음. (고양이:[1,0,0], 개:[0,1,0], ...)
- 관찰된 숫자들이 들어가서 원하는 숫자들이 나오도록 하는 검은 상자
  - 검은 상자는 다양한 방식으로 구현될 수 있지만 현재는 **신경망**이 선호됨
  - 관계식에 필요한 파라미터는 예시로 부터 결정됨 (training 과정)

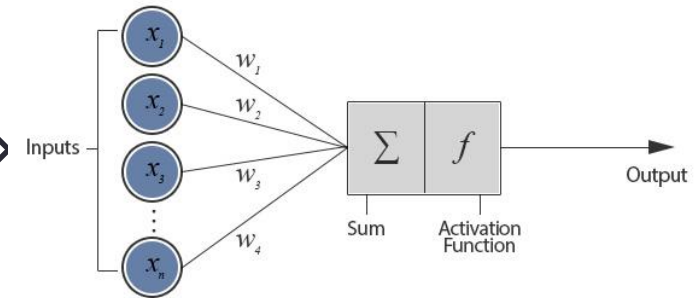
# 뉴런: 신경망의 기본 단위



# 인공 뉴런(Artificial Neuron)

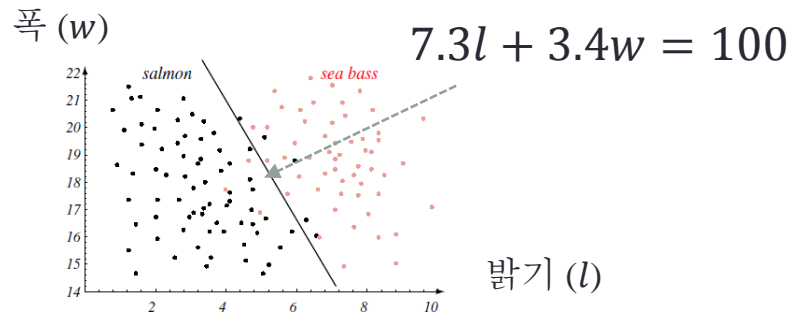


실제 뉴런

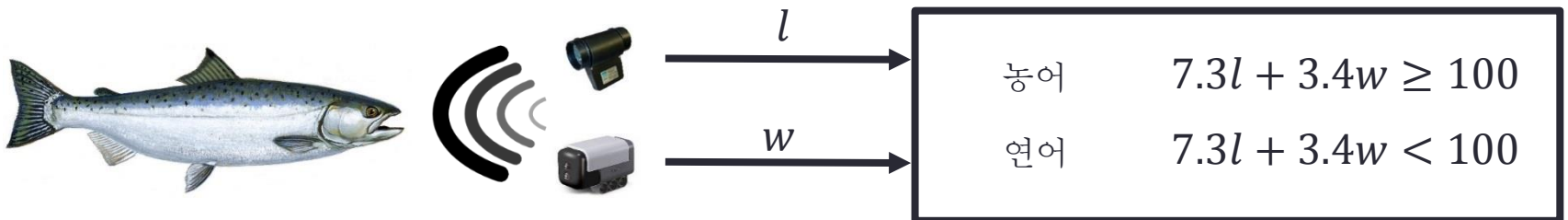


뉴런의 수학적 모델

# 예시: 연어와 농어의 구별

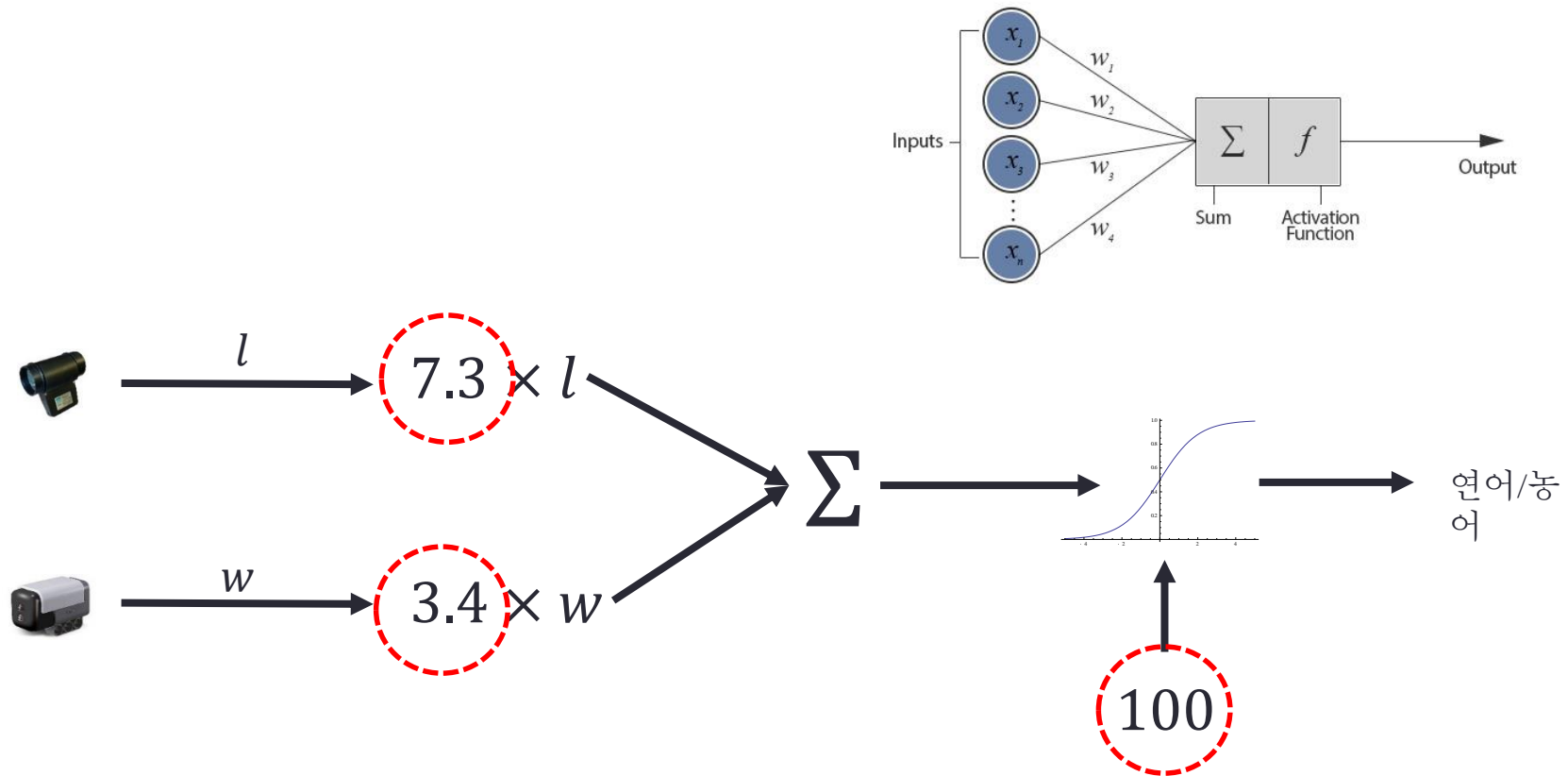


**FIGURE 1.4.** The two features of lightness and width for sea bass and salmon. The dark line could serve as a decision boundary of our classifier. Overall classification error on the data shown is lower than if we use only one feature as in Fig. 1.3, but there will still be some errors. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.

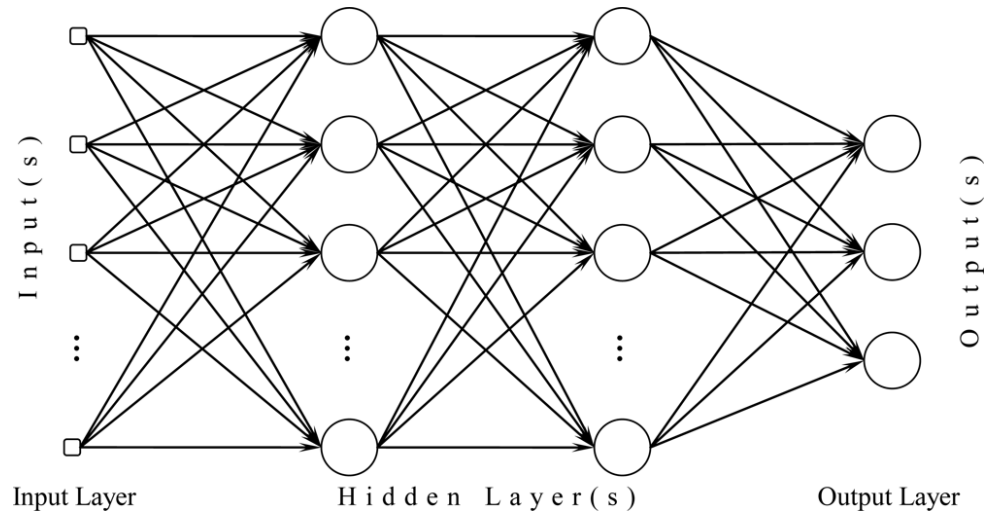
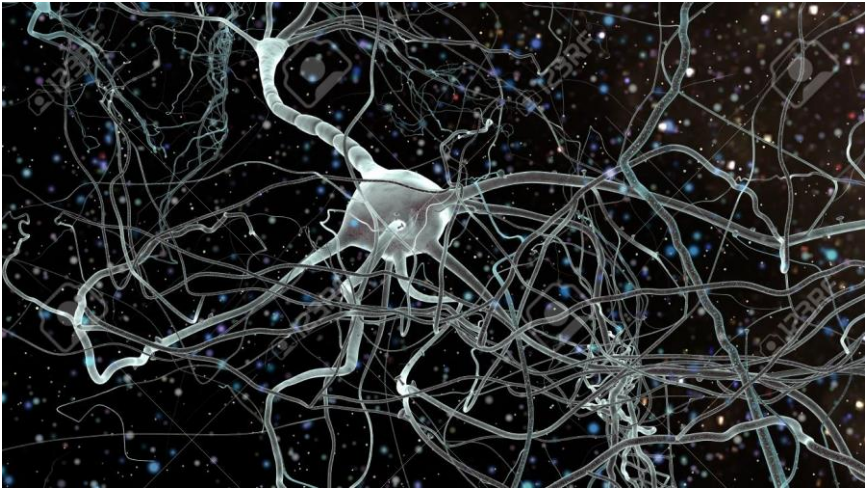




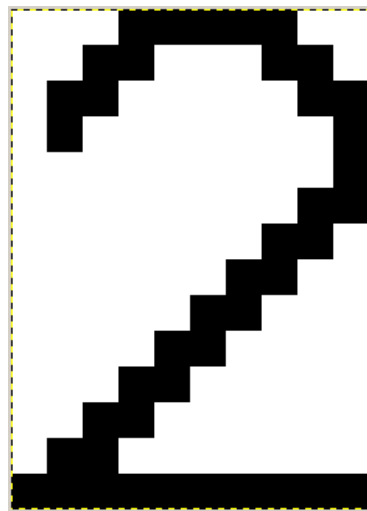
# 예시: 연어와 농어의 구별



# 다층 구조 네트워크 (인공신경망)

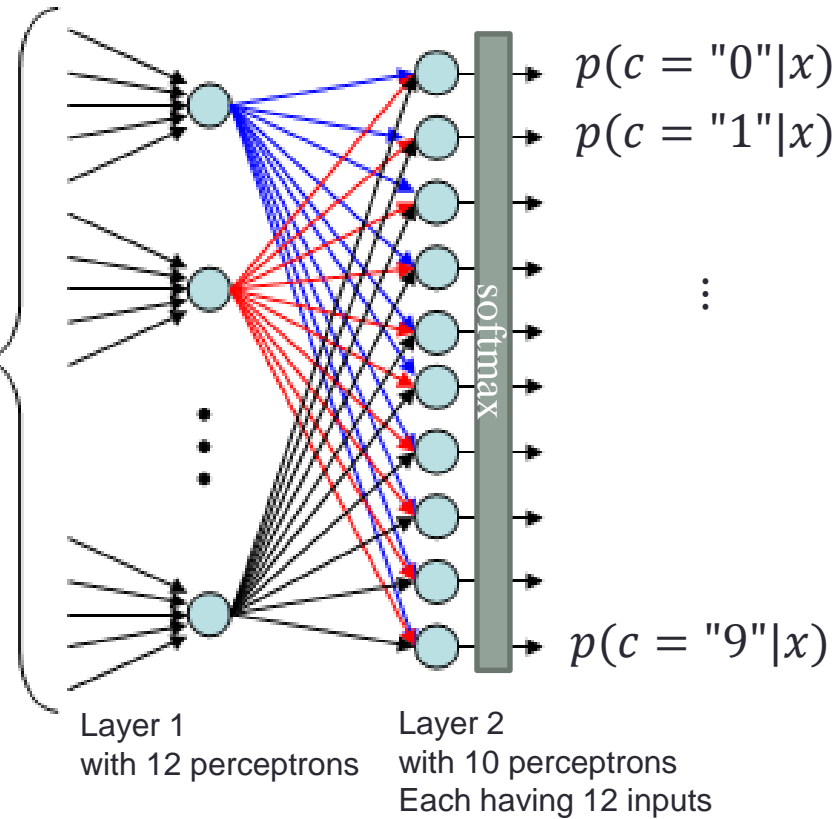


# 예시: 숫자 인식



$$x \in \{0,1\}^{10 \times 14}$$

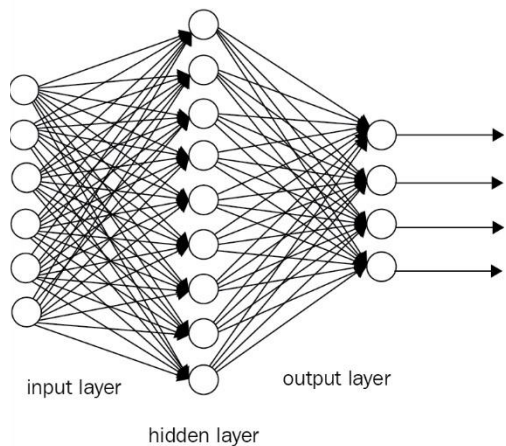
140 inputs



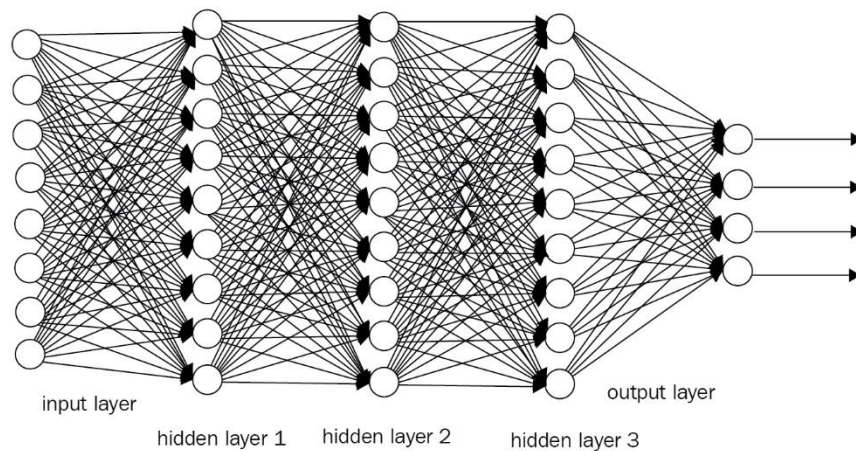
딥러닝

---

# 신경망 vs 깊은 신경망



신경망



깊은 신경망  
(Deep Neural Networks)



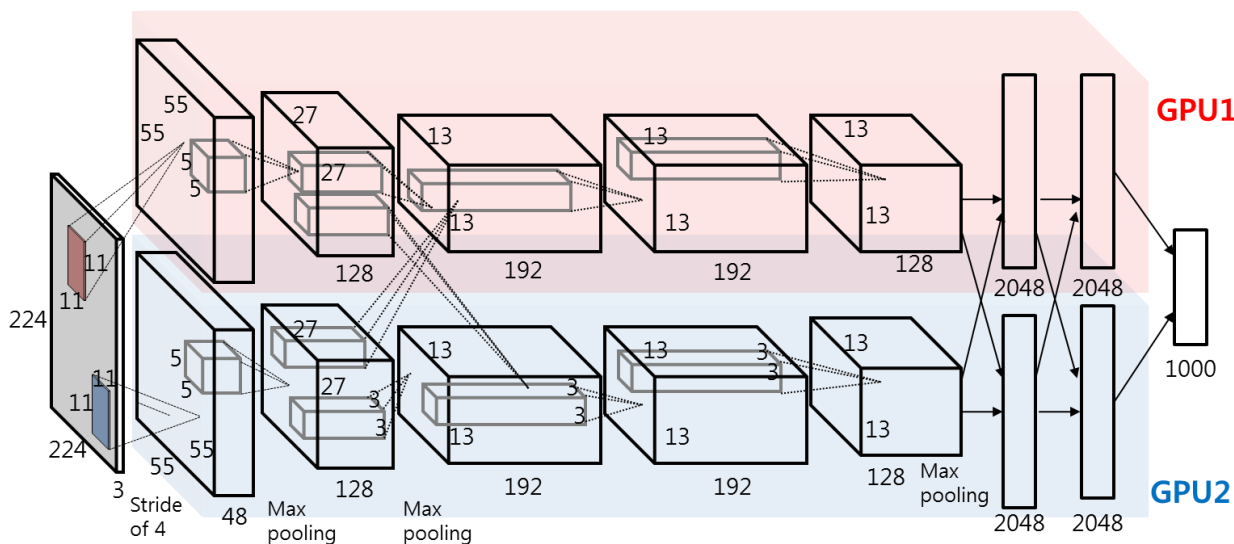
# Large-scale recognition

## 1,000,000 images and 1,000 categories



# AlexNet

- AlexNet won the 2012 ImageNet competition
  - 5 convolutional layers, 2 fully connected layers
  - The input is a color 224x224 image
  - 2 GPU architectures





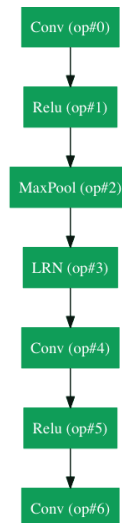
# AlexNet results (2012)

- [AlexNet TensorFlow codes and some results](#)

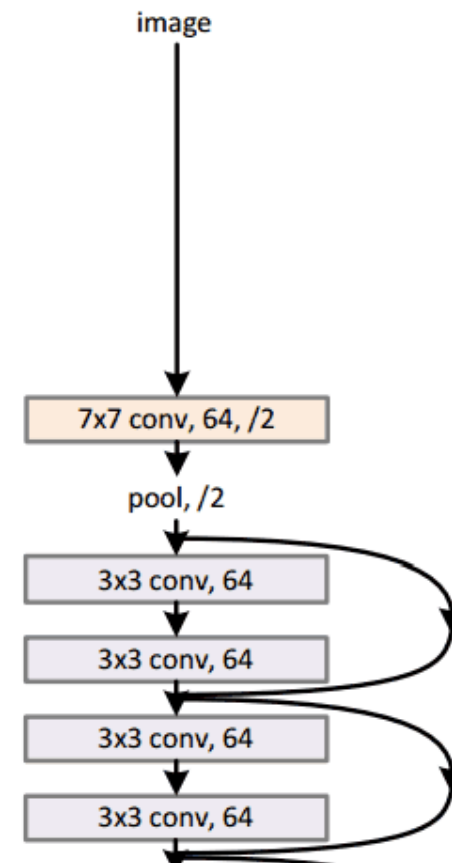




# GoogLeNet (2013), ResNet-34 (2014)



34-layer residual

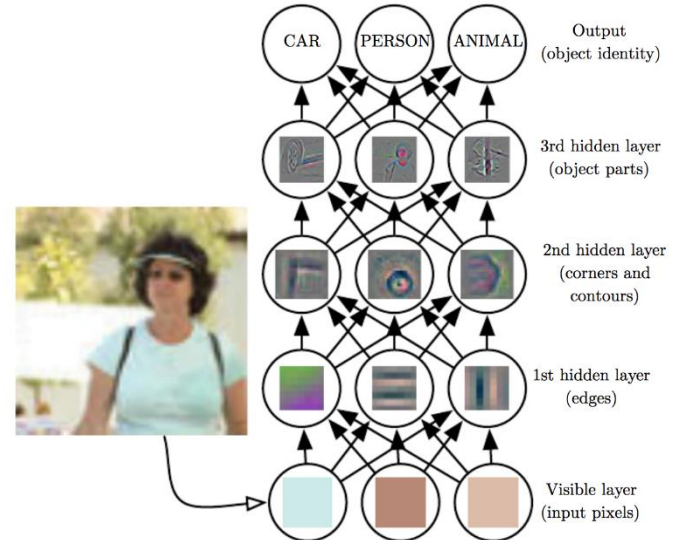
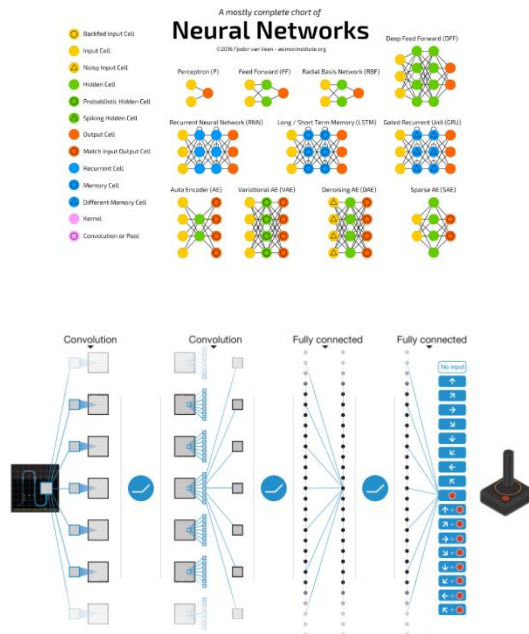


<http://ethereon.github.io/netscope/quickstart.html>

# 딥러닝의 특징

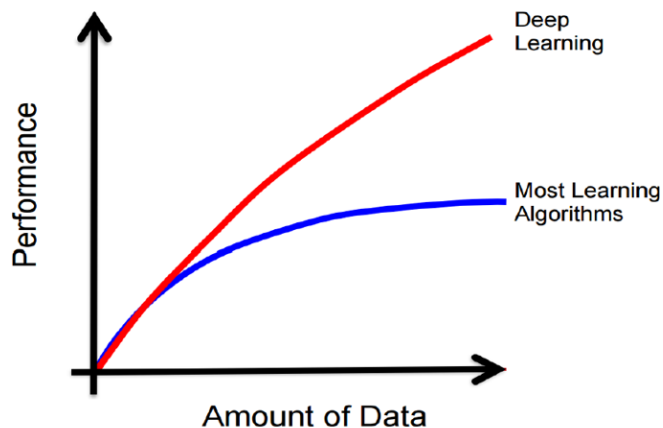
- 높은 유연성
  - 다양한 문제에 적용 가능
  - 다양한 구조 가능

- 표현법 학습
  - 계층적 특징 학습
  - 분산 표현



# 딥러닝의 특징

- 데이터 양에 비례하여 성능향상



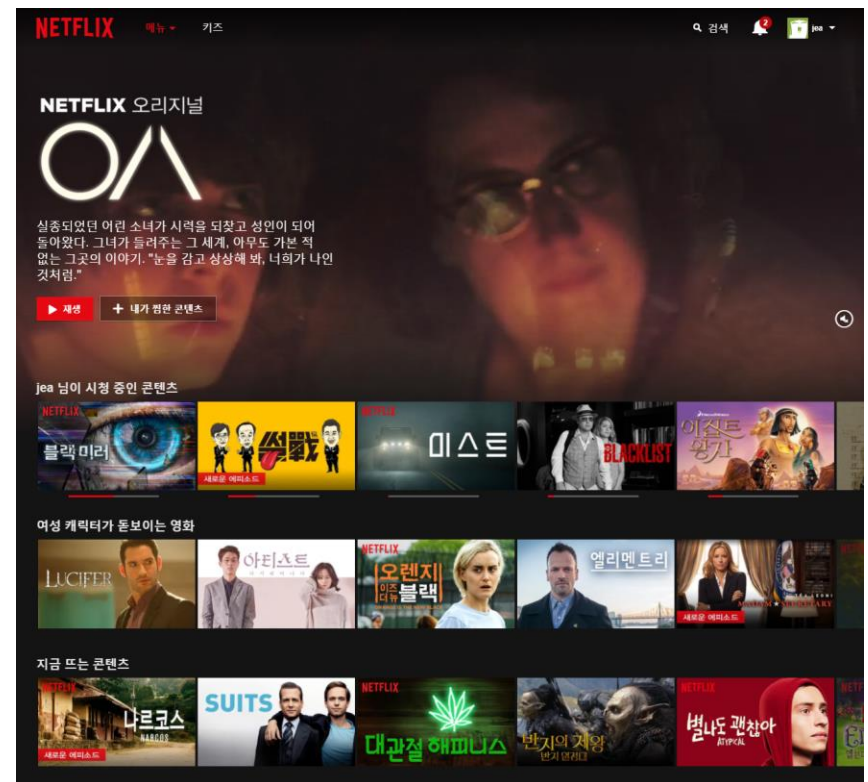
- 병렬 처리에 적합
  - GPU, TPU, ...



# 미디어 응용

---

# NETFLIX



## 넷플릭스의 빅 데이터(Big Data), 인문학적 상상력과의 접점<sup>1)</sup>

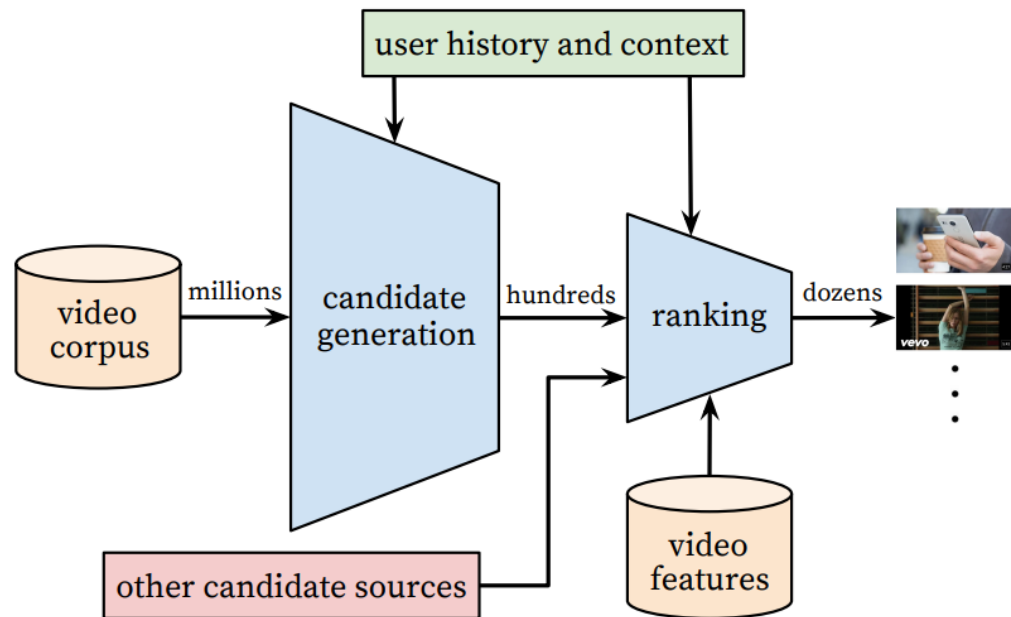
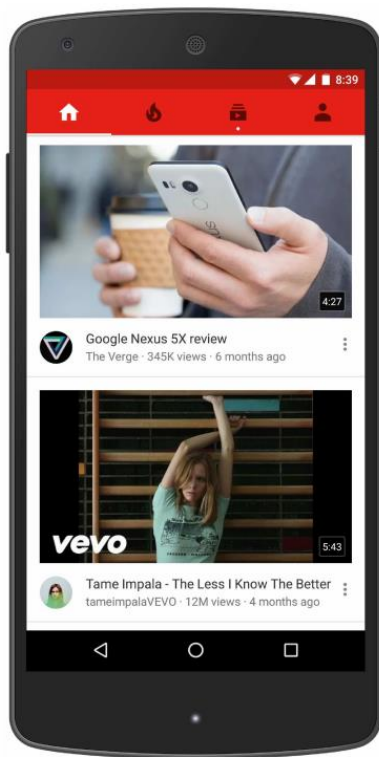
조영신 (SK 경영경제연구소 수석연구원, troicacho@hanmail.net)

추천은 그 자체로 환전성이 있기 때문에 가치가 있다. 넷플릭스의 경우 대여되는 영화의 2/3가 추천을 통해 발생했으며, 구글 뉴스(Google News)의 경우 38% 이상이 추천을 통해서 조회가 발생하는 것으로 알려져 있다. 또한 아마존의 경우에도 추천을 통한 판매가 전체 매출액의 35%를 넘는다. 결국 추천은 해당 기업의 주력 상품을 판매하는데 결정적인 역할을 하는 셈이다. 만약 추천 시스템의 성능이 10% 향상된다면 수익 증가는 자명하다.

심지어 넷플릭스는 자사가 보유하고 있는 추천 시스템의 정확도를 10% 증가시킬 수 있는 알고리즘을 개발해 주는 이들에게 상금으로 1백만 달러를 내걸었다. 당시 100여개 국가에서 수없이 많은 팀들이 참여했고, 결국 2009년에서야 BellKor의 Pragmatic Chaos가 우승을 했다. 이 때 사용한 방식은 사용자가 가지고 있는 지표를 가능한 한 많이 분석하는 추천 알고리즘이었는데, 그 방식이 머리가 아플 정도로 복잡했다고 한다.

최근에 넷플릭스는 추천 시스템을 한 단계 더 진화시키려고 하고 있다. 딥러닝(Deep Learning)방식을 도입하기로 한 것이다. 인공신경망(Neural Network)를 활용해서 사용자의 선호도나 영화 감상에 숨어있는 지표들을 바탕으로 학습시키고 추천 알고리즘을 만들어 낼 생각인 듯하다. 이에 대해서 한상기 박사(2014, 2)는 “이전의 추천 방식은 선형 대수 방식으로 계산을 했던 것에 비해서 어느 정도 개선을 할 수 있을지에 대해서는 물음표를 던진 상황이긴 하지만 적어도 방대한 데이터가 있고 이를 처리할 수 있는 물리적 환경이 존재한다는 점에서 가능성이 있다”<sup>3)</sup>는 평가를 내리고 있다. 더 흥미로운 것은 딥 러닝을 위한 인공신경망을 아마존 클라우드 서비스에서 운영하기로 결정한 것이다. 아직은 어떤 방식으로 추천 시스템을 만들지 자세히 보도하지는 않았으나, 추천 시스템과 클라우드의 결합, 딥 러닝을 통한 추천 시스템 개발 실험은 의미 있는 한 걸음이다. 단 1%만 향상되더라도 수익성은 달라질 것이 자명하고, 그 결과는 오롯이 넷플릭스의 몫이다.

# Deep Neural Networks for YouTube Recommendations



<https://research.google.com/pubs/pub45530.html>



# MUSIC AI





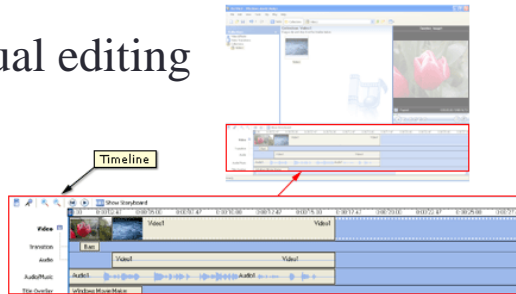
# Cognitive Movie Trailer



# 100 Horror movie trailers



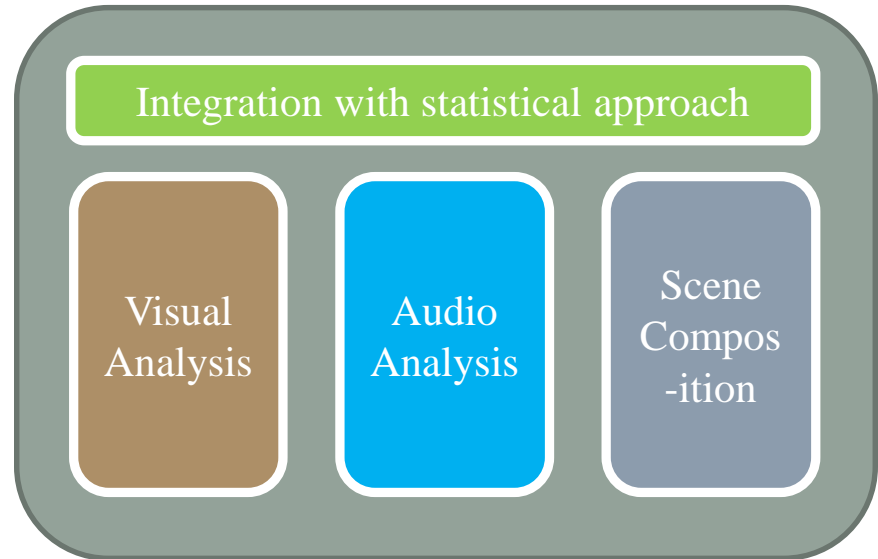
Manual editing

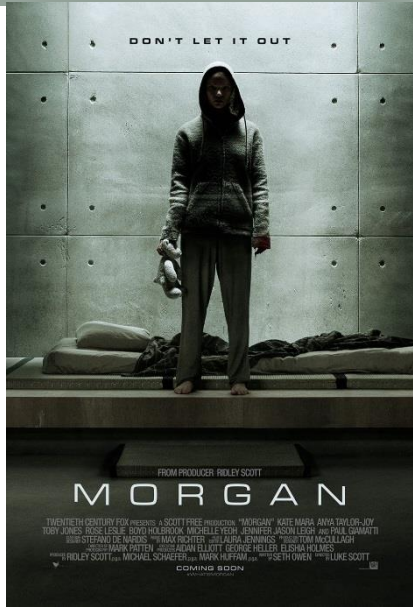


Moments

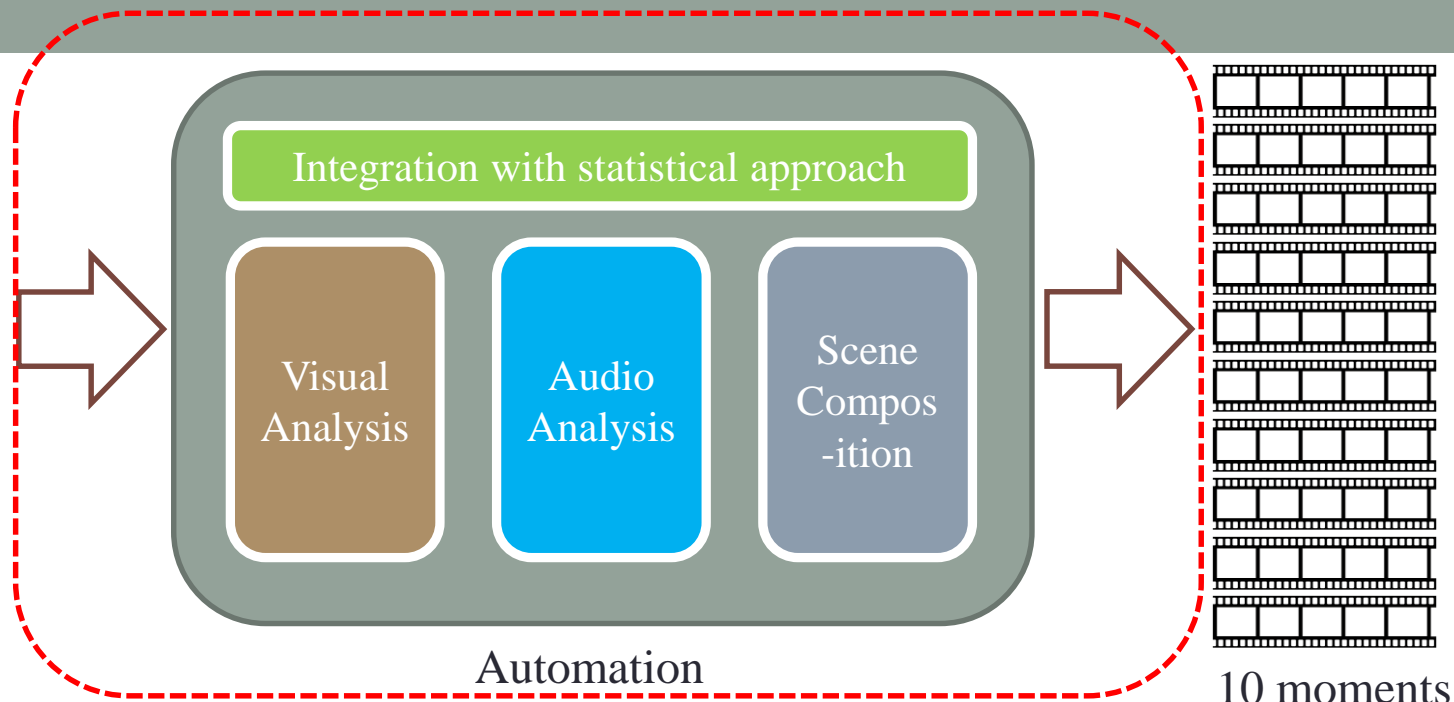


Machine Learning

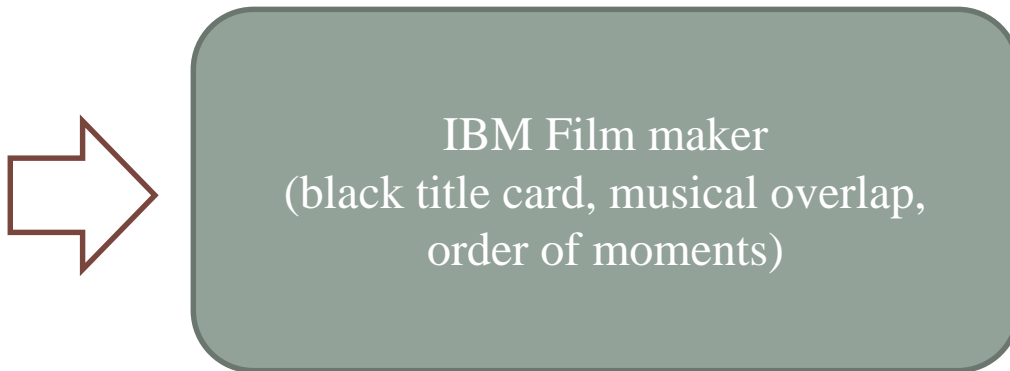




Morgan Full Movie  
1h 32m



10 moments



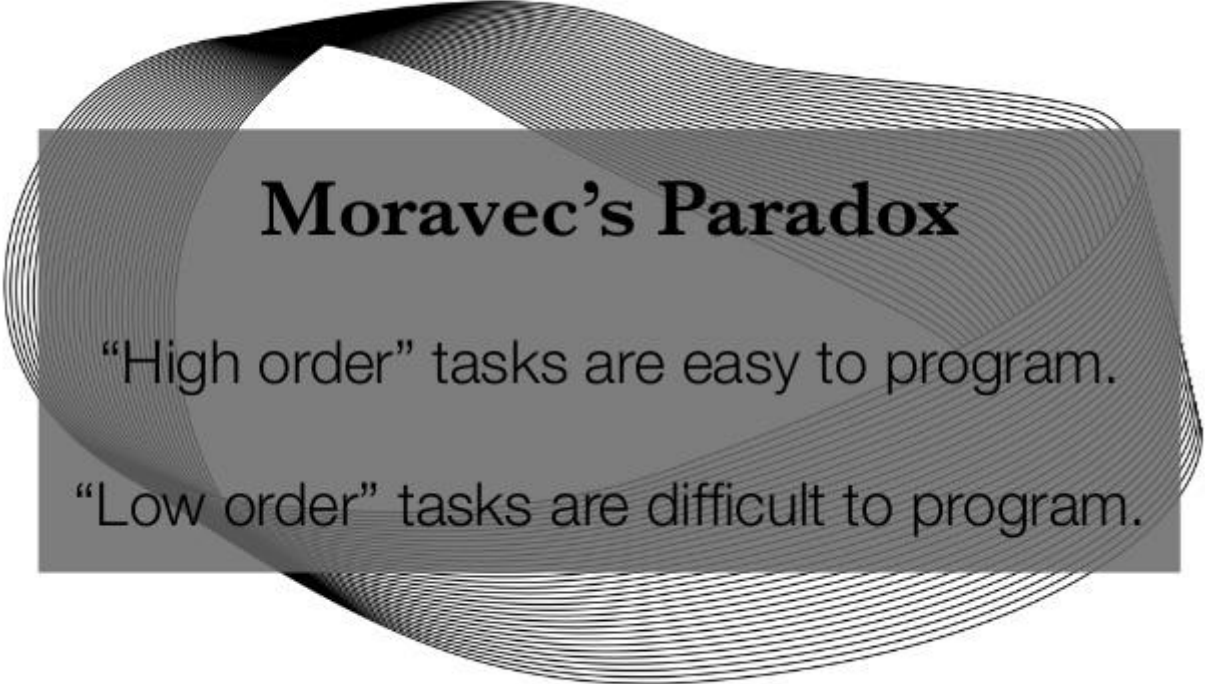
Morgan Trailer  
2m 32s



**어떤 문제를 해결할 수 있을까?**

---

# 모라벡의 역설



## Moravec's Paradox

“High order” tasks are easy to program.

“Low order” tasks are difficult to program.

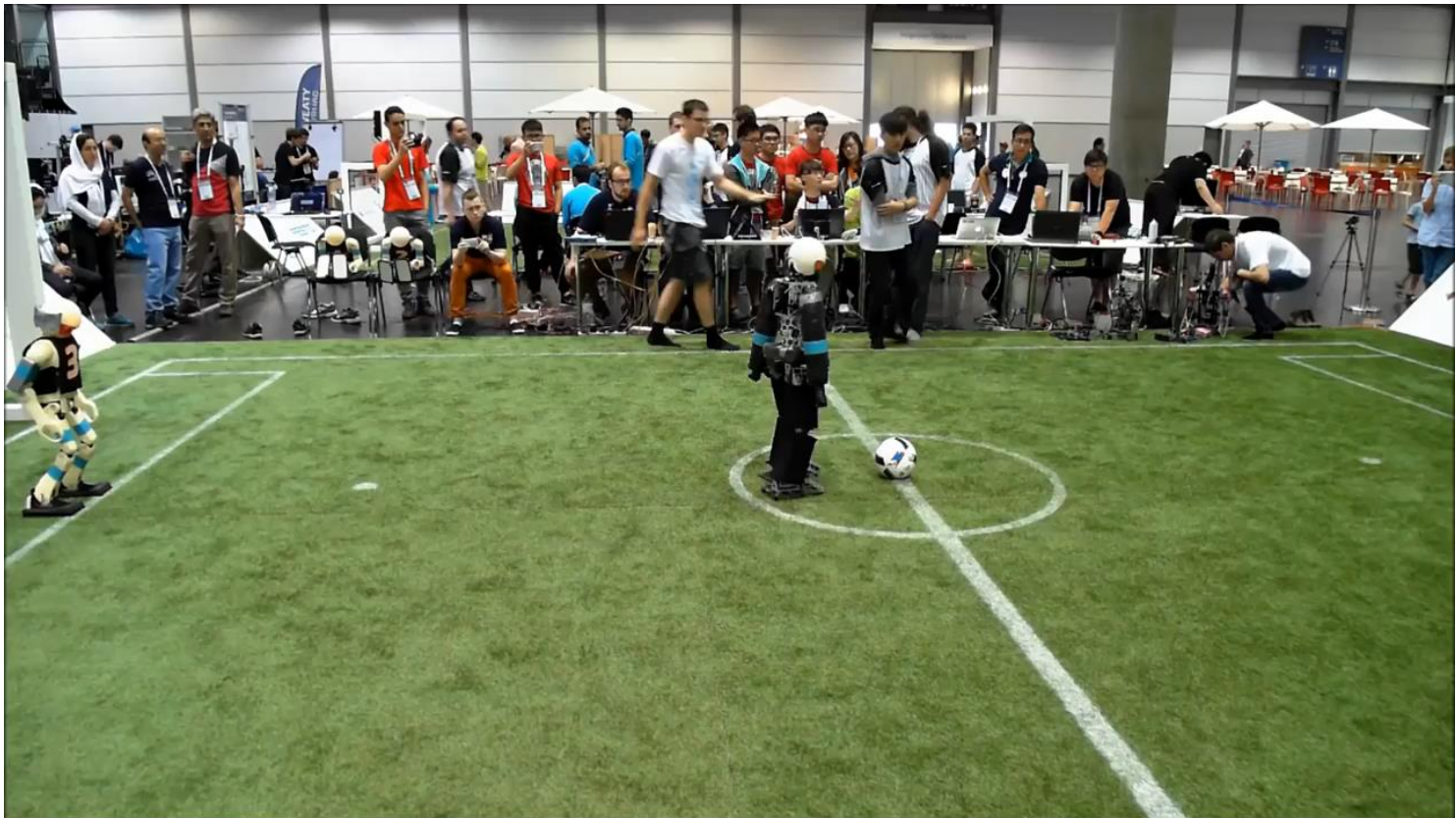
# 모라벡의 역설



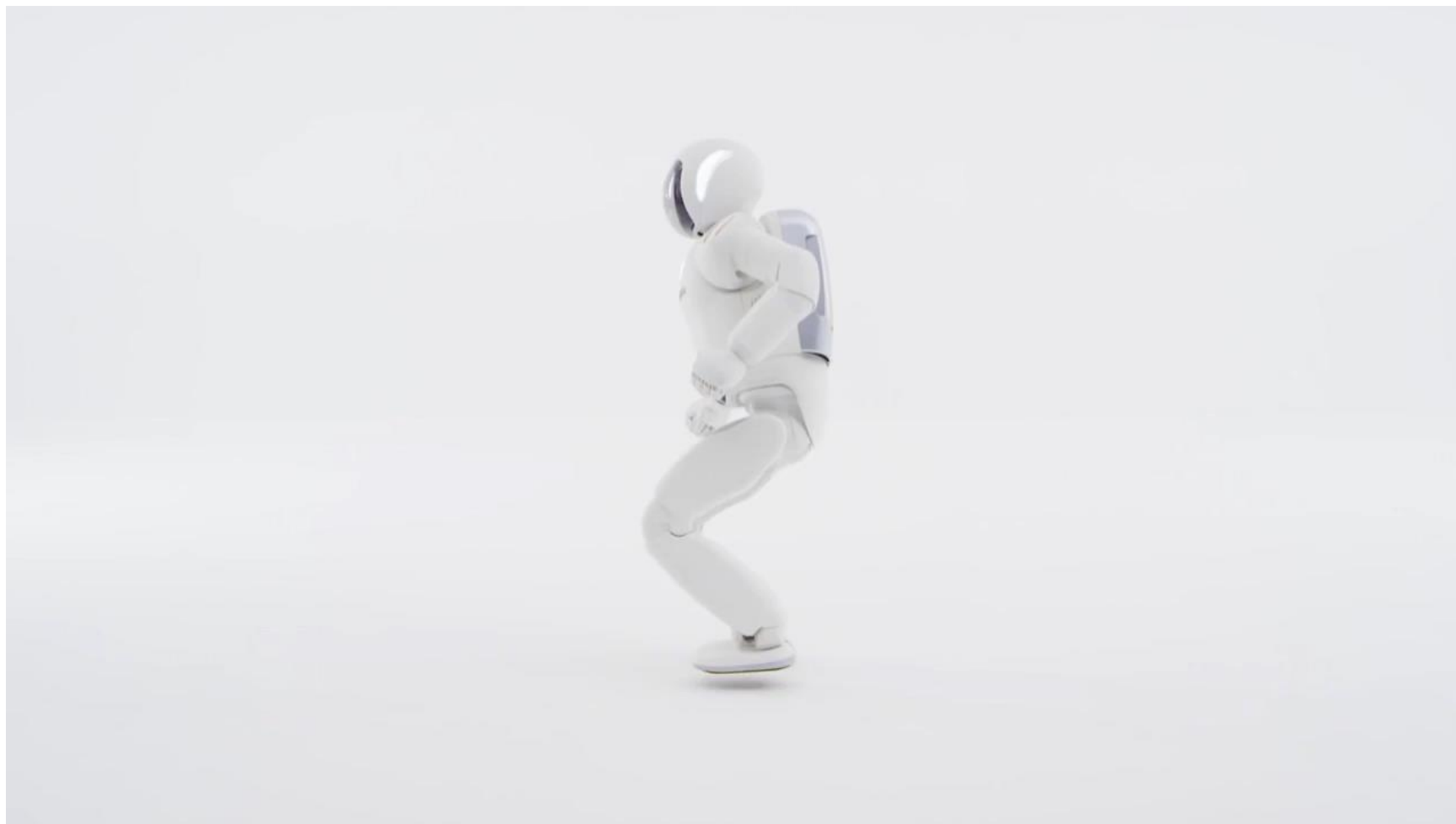


# 모라벡의 역설

- RoboCup 2016: NimbRo vs AUTMan



# 혼다 아시모



# Why?

- “Encoded in the large, highly evolved sensory and motor portions of the human brain is a billion years of experience about the nature of the world and how to survive in it.
- We are all prodigious Olympians in perceptual and motor areas, so good that we make the difficult look easy. Abstract thought, though, is a new trick, perhaps less than 100 thousand years old. We have not yet mastered it. It is not all that intrinsically difficult; it just seems so when we do it”
  - Moravec, Hans (1988), Mind Children, Harvard University Press

# 쉬운 문제일까? 어려운 문제일까?



ph.d 시스템 설계  
80년 옥스포드를 다녔고,

MacGyver, 1985~

# 쉬운 문제일까? 어려운 문제일까?

	사람	컴퓨터
대화 - 음성인식, 생각	*	*****
말하기 - 음성합성 (TTS)	***	***
그림 그리기 - 그래픽스	*****	*



번역

영어 한국어 독일어 언어 감지 ▼



한국어 영어 일본어 ▼

번역하기

한국어 문장을 읽는 것도 요즘에는 기계가 더 잘하는 거 같아요.



35/5000

Reading Korean sentences seems to be better for machines these days.



수정 제한하기

hangug-eo munjang-eul ilgneun geosdo yojeum-eneun gigyega deo jalhaneun geo gat-ayo.



Hard problems for programmers

Talk, Read,  
Walk,  
Drive,  
Atari,  
...



Hard problems for humans



Model-  
based  
things (e.g.,  
Physics  
simulation)

Easy problems for programmers



Easy problems for humans







Hard problems for programmers

Talk, Read,  
Walk,  
Drive,  
Act,  
...

# Deep Learning to the Rescue



Hard problems for humans



Model-  
based  
things (e.g.,  
Physics  
simulation)

Easy problems for programmers



Easy problems for humans





Hard problems for programmers

Talk, Read,  
Walk,  
Drive,  
Atari,  
...

# Deep Learning to the Rescue?



Easy problems for humans

Hard problems for humans



Model-  
based  
things (e.g.,  
Physics  
simulation)



Easy problems for programmers



**Demis Hassabis**

@demishassabis



Follow

#AlphaGo WINS!!!! We landed it on the moon.  
So proud of the team!! Respect to the amazing  
Lee Sedol too

RETWEETS

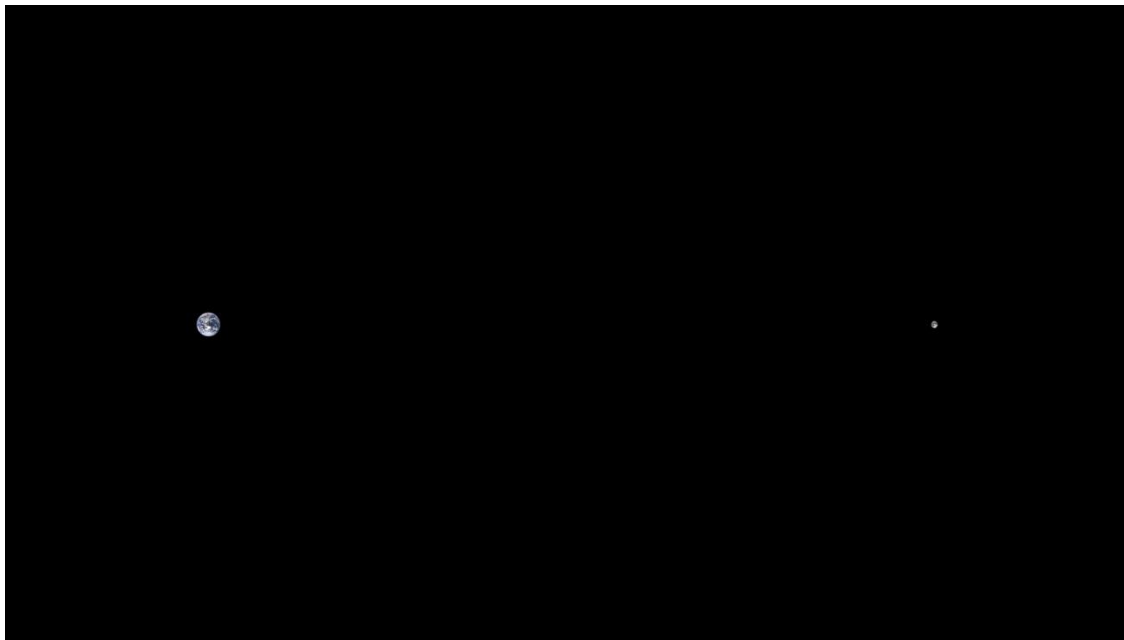
871

LIKES

703



11:54 PM - 8 Mar 2016



데미스 허사비스



비디오 게임 개발자

데미스 하사비스는 영국의 인공지능 연구인, 컴퓨터 게임 설계자, 게이머이다. 그는 구글 딥마인드의 창업자이자 최고 경영자이다. 그는 13살의 나이에 체스마스터가 됐고, 17살에 게임 개발에 참여하기 시작했다. 어린 나이부터 프로 그래밍을 배운 그는, 블랙 & 화이트라는 게임의 인공지능을 개발하는 데 참여했다. 엘릭서 스튜디오라는 비디오 게임 개발사를 만들어, '리퍼블릭: 혁명'과 '이블 지니어스'라는 게임도 제작했다. 위키백과

출생: 1976년 7월 27일 (41세), 영국 런던

학력: 케임브리지 대학교

설립 단체: 딥마인드, Elixir Studios

함께 찾은 검색어

10개 이상 항목 더보기



이세돌



판 후이



Mustafa  
Suleyman



Shane Legg

# 정리

- 인공지능
  - 사람처럼 일하는/행동하는 기계를 만들고자 함
- 머신러닝
  - 인공지능을 실현하는 하나의 방법
  - 입출력 관계를 데이터를 통해서 찾으려는 분야
- 딥러닝
  - 대규모의 신경망을 사용하는 머신러닝 방법
  - 이론적인 측면 뿐 아니라 실용적인 측면의 장점이 많음
- 인공지능의 힘
  - 미래를 예측하는 건 쉽지 않지만, 현재 수준으로도 많은 분야에 적용되어 인간의 전문성과 창의성이 더 발휘될 수 있도록 도울 수 있음

# BACKUPS

---

# To Scale: The Solar System





# 세부적인 과정

- Training
  - The Research team **trained a system** on the trailers of 100 horror movies by segmenting out each scene from the trailers. Once each trailer was segmented into "moments", the system completed the following:
    - **A visual analysis and identification** of the people, objects and scenery.
    - **An audio analysis** of the ambient sounds (such as the character's tone of voice and the musical score)
    - **An analysis of each scene's composition** (such the location of the shot, the image framing and the lighting)
- The analysis was performed on each area separately and in combination with each other using statistical approaches.

# 세부적인 과정

- The full-length feature film, "Morgan" is fed to the system. It identified **10 moments** that would be the best candidates for a trailer.
- Manual processing
  - Use IBM **filmmaker** to arrange and edit each of the moments together into a comprehensive trailer. (+ black title cards, the musical overlay and the order of moments).
- Footage on the cutting room floor
  - Some moments in the movie that were not included.

