

# AI for Oil & Gas

Andrea Panizza



<https://github.com/AndreaPi/docker-training-2019-public>



# AI is around us



Self-driving cars



Health



Wearables



Voice assistant



Surveillance



Smart home



Translation

# Video synthesis

<https://www.youtube.com/watch?v=p1b5aiTrGzY>

---

# Which of these persons is real?

1

2

3

4



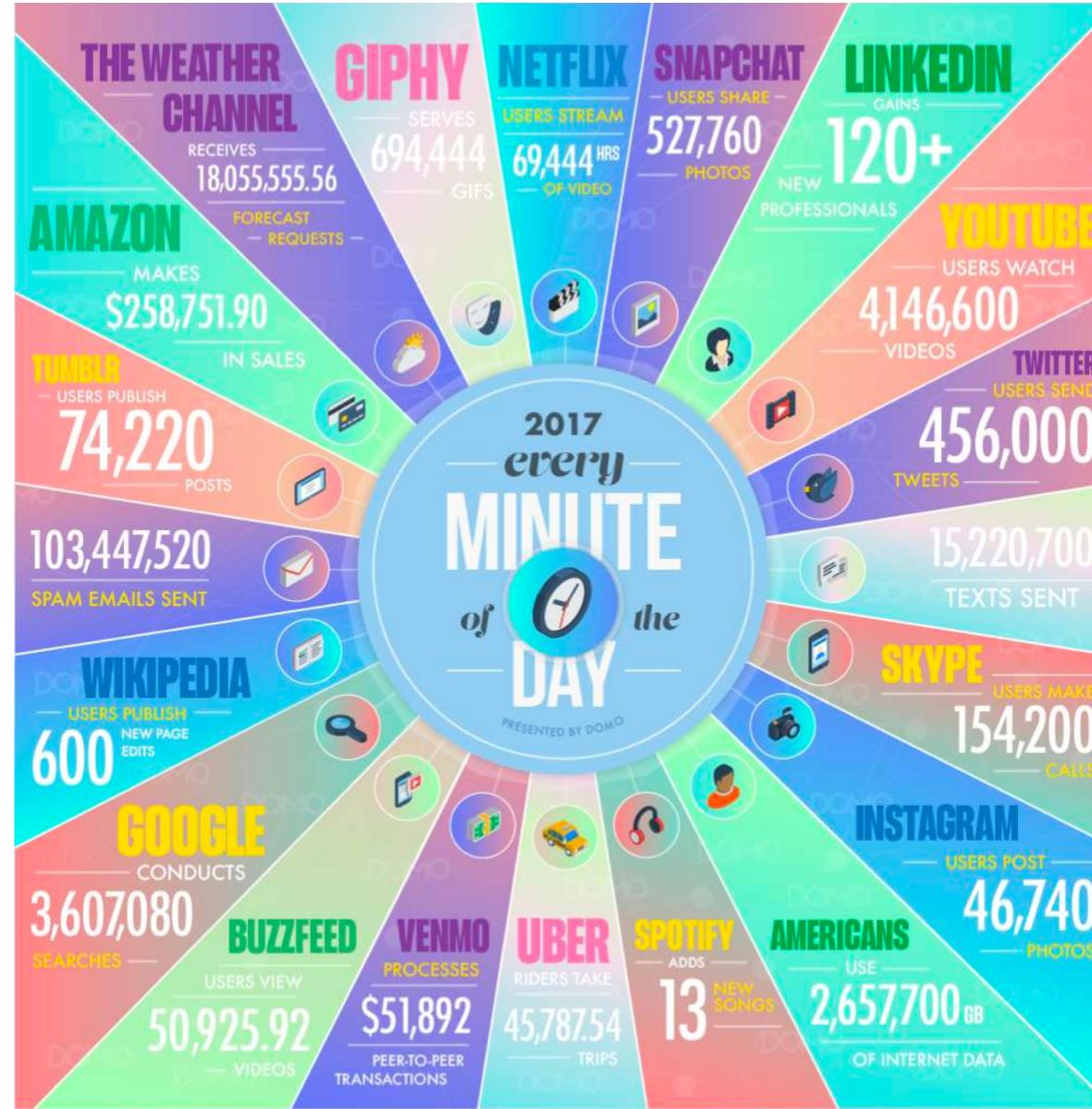
---

# Video synthesis

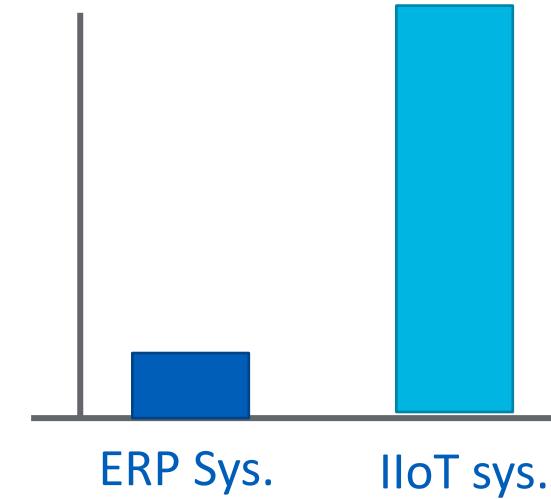
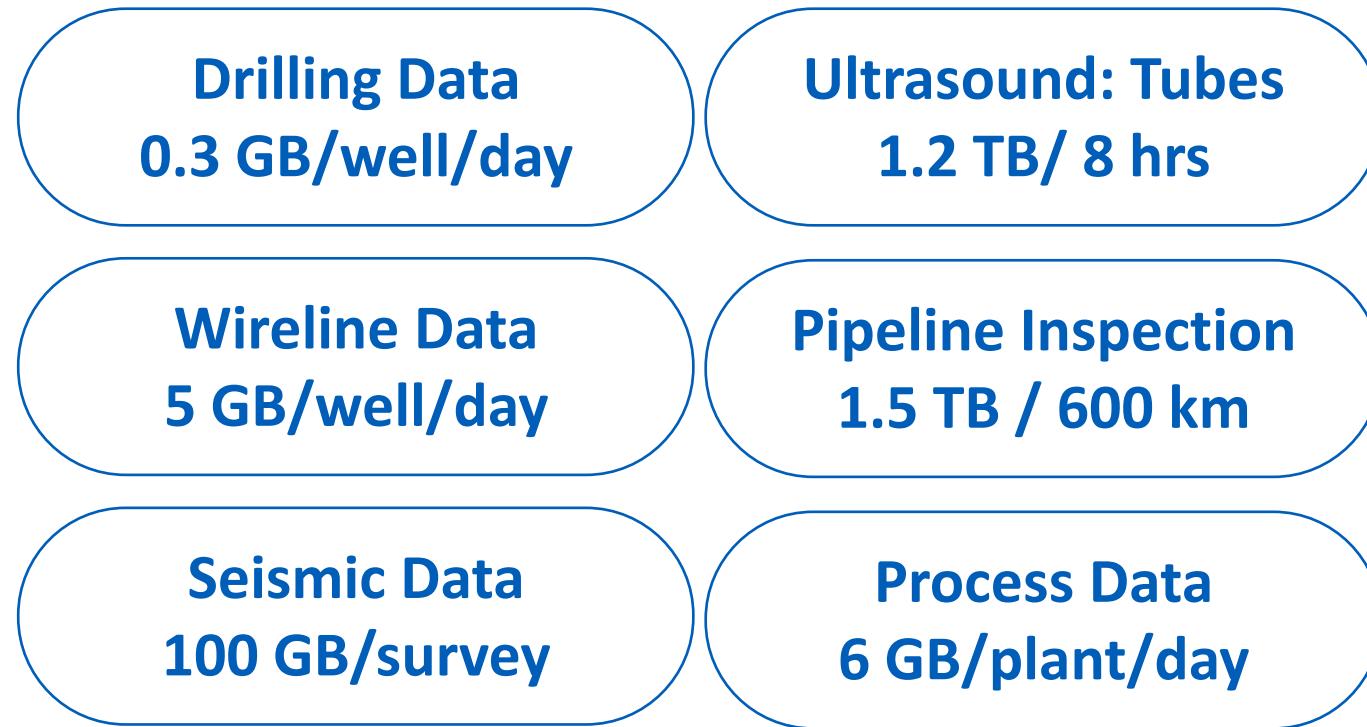
<https://youtu.be/c-NJtV9Jvp0>

---

# BIG data



# BIG data in Oil & Gas

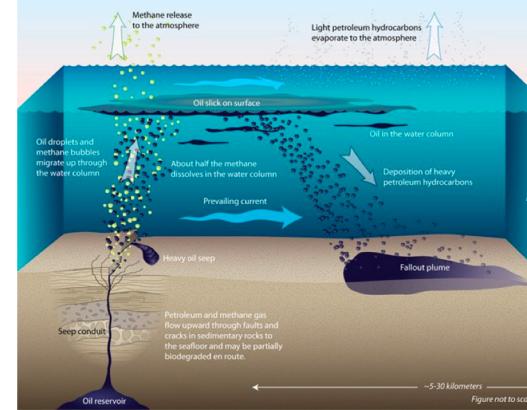


- ~10-100x more volume
- ~100-1000x more velocity

# AI in Oil & Gas

## ExxonMobil

- working with MIT to design AI robots for ocean exploration
- The robots fly over the ocean to detect natural seeps which occur when oil escapes from rock found in the ocean floor



## Sinopec

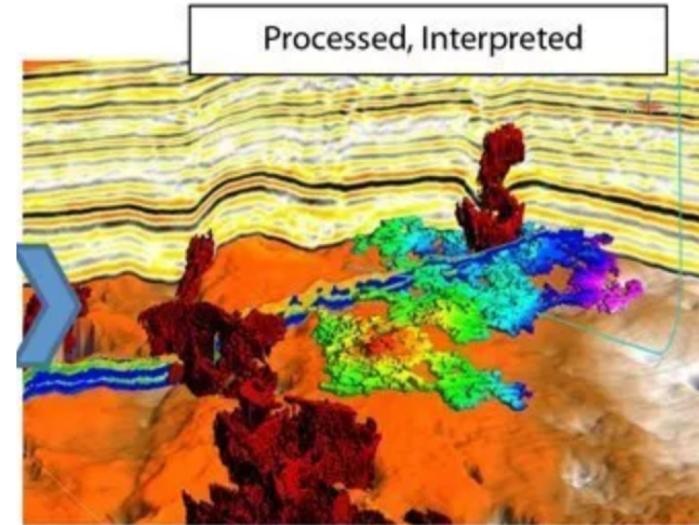
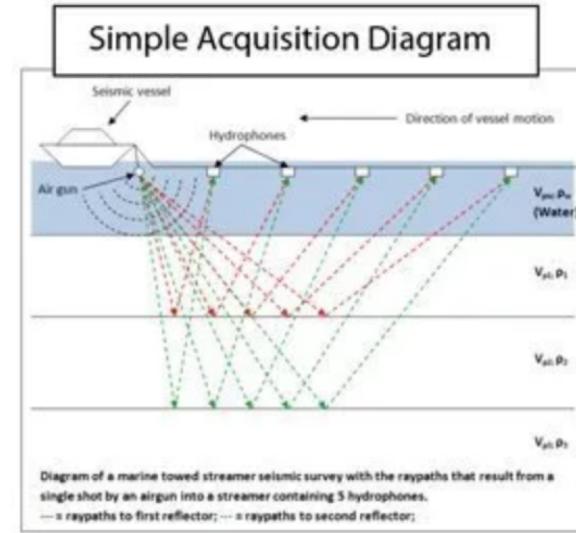
- long-term plan to build 10 AI-driven plants, 20% reduction in operation costs
- Collaboration with Huawei (an AI giant) in April 2017 to design a “smart manufacturing platform” for centralized data management and integration of data to manage factory operations
- AI will interpret data and look for opportunities to improve factory operations



# AI in Oil & Gas

## Total & Google Cloud Platform

- Traditionally subsurface seismic images at Total were analyzed manually by geologists
- Total & GCP are collaborating to develop AI systems that can help geologists interpret such images using **computer vision, NLP & conversational AI**
- Goal: build a tool which analyzes image files, correlates findings with info from technical docs & summarizes info into a queryable AI assistant that can answer questions on the subsurface data in natural language.
- Total is also using machine learning for more “classical” applications such as predictive maintenance of turbines, pumps, and compressors



# AI in Oil & Gas

## Siemens

- Uses AI to control GT combustion (GT-ACO, Gas Turbine Autonomous Control Optimizer)
- NOx emission reduced by 20% with respect to settings from an expert control engineer
- Used in pilot operations in the United States and South Korea



# AI in Oil & Gas

## Many others

- Repsol uses ML to simplify and increase the accuracy of geophysical analyses
- Saudi Aramco uses clustering to reduce the control effort for its wells by using a common model for wells in the same cluster
- Shell uses ML to:
- manage unmanned service stations and check that no one smokes near a pump
- control drones which survey inhospitable locations
- perform predictive maintenance
- predict energy market trends
- It even has an AI residency program, like Google, Facebook, etc.!



[Shell RechargePlus: Managed Smart Charging for Electric Vehicles](#)



[A bionic inspector rolls in](#)

Sensabot is a robot that can help maintain oil and gas field equipment in hostile environments.



[Eye in the sky](#)

See how we are using high-tech drones to go places humans would rather avoid.

# AI in Oil & Gas

## And of course, Baker & Hughes!

- **Document validator**
- Significantly reduce the need to outsource classification to an external supplier
- Eliminate the need for outsourcing, if CV is coupled to NLP (increase in effort/risk)
- Accelerate classification process significantly

**ORDER ACCEPTANCE FORM/ACCEPTATION DE COMMANDE**

Attn Thermodyn S.A.S  
C/o este soita  
Av. de l'Europe, CCL 2  
90127 France  
Fax +33 03 84 51 53 93  
Email [thermodyn.sas@thermodyn.fr](mailto:thermodyn.sas@thermodyn.fr)

Date 09/04/2019

AI 8963 00  
TAHUD Elec

Nous acceptons la réception de la commande ci-dessous à nos conditions générales de vente et nous la concéderons à l'acheteur suivant : + **commentaires annotés SKF sur PO**

Wen et reçus du succès de la réception de la commande suivante et nous confirmons notre acceptation de la même.

Référence : Ammendement dated 08-APR-2018 14:36  
Commande/Purchase Order # A392-G7085 du 21-OCT-2018 15:49  
Acheteur : Martel, Mr. Anthony

Veuillez faire parvenir le présent document signé par les moyens suivants au destinataire :  
Courrier recommandé  
Fax : +33 03 84 51 53 93  
E-mail : [thermodyn.sas@thermodyn.fr](mailto:thermodyn.sas@thermodyn.fr)

Vous êtes prié de nous renvoyer le présent document signé par le moyen suivant au destinataire :  
Courrier recommandé  
Fax : +33 03 84 51 53 93  
E-mail : [thermodyn.sas@thermodyn.fr](mailto:thermodyn.sas@thermodyn.fr)

Si je vous en prie, veuillez me renvoyer le présent document signé par le moyen suivant :  
Courrier recommandé  
Fax : +33 03 84 51 53 93  
E-mail : [thermodyn.sas@thermodyn.fr](mailto:thermodyn.sas@thermodyn.fr)

SKF Magnetic Mechatronics S.A.S  
S.A.S. c.c. 001 de 5.961 K-4 E  
2, rue des Champs,  
27 950 Saint-Marcel, FRANCE  
Tel. 02.32.64.33.00 - Fax 02.32.21.25.99  
N°SIT: 304.174.176 - N°TVA: 27112

**ORDER ACCEPTANCE FORM/ACCEPTATION DE COMMANDE**

Attn: Thermodyn S.A.S  
c/o cette partie  
Ville de l'Europe, CCL 2  
90127 France  
Fax: +33 03 84 51 53 93  
Email: [thermodyn.sas@thermodyn.fr](mailto:thermodyn.sas@thermodyn.fr)

Date 28/11/18

Nous acceptons la réception de la commande référencée ci-dessous. Nous confirmons en avoir pris connaissance et en acceptons les termes.  
We acknowledge receipt of the following Purchase Order and we confirm our acceptance of its contents.

Commande/Purchase Order # A39308637 du 26-NOV-2018 17:09  
Acheteur : YAHKSEEF, JAWAD

S 917341

Veuillez prendre du temps pour lire le présent document signé par les moyens suivants au choix :  
o) par fax au n°  
+39 055/405851993  
b) par e-mail à l'adresse : [oilganges.pocconformation@ge.com](mailto:oilganges.pocconformation@ge.com)  
You are kindly requested to send the present document duly signed by the following means:  
a) by fax to this number:  
+39 055/405851993  
b) in pdf format to the following mail address:  
[oilganges.pocconformation@ge.com](mailto:oilganges.pocconformation@ge.com)

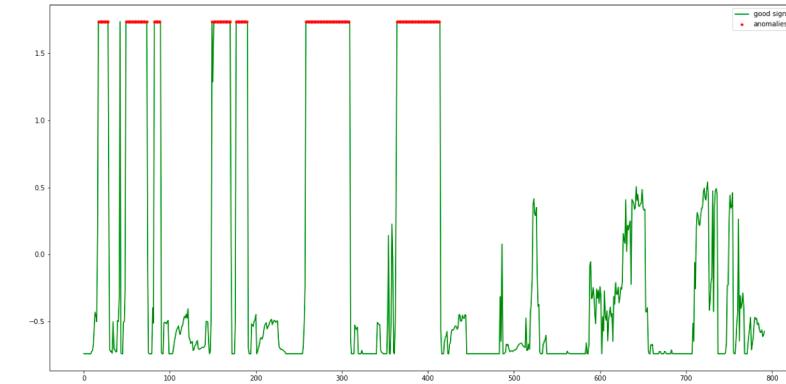
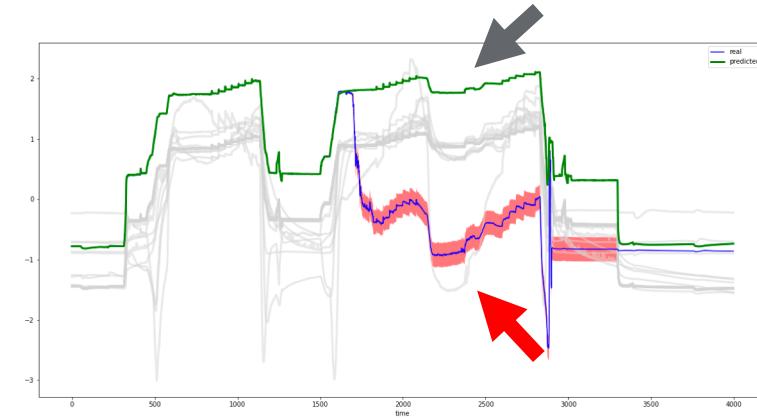
Solutions distinguées  
Confidential signature  
Stamp and signature

Waukesha Bearing Liu  
28 NOV 2018

# AI in Oil & Gas

## Baker & Hughes

- **Sensor Anomaly Detection**
- Identifies complex anomalies (scaled, saturation, drifting, multivariate anomaly)
- Uses an LSTM to predict the future behavior, and compares it with the actual measured signal
- Extremely accurate
- Many other applications I can't currently show

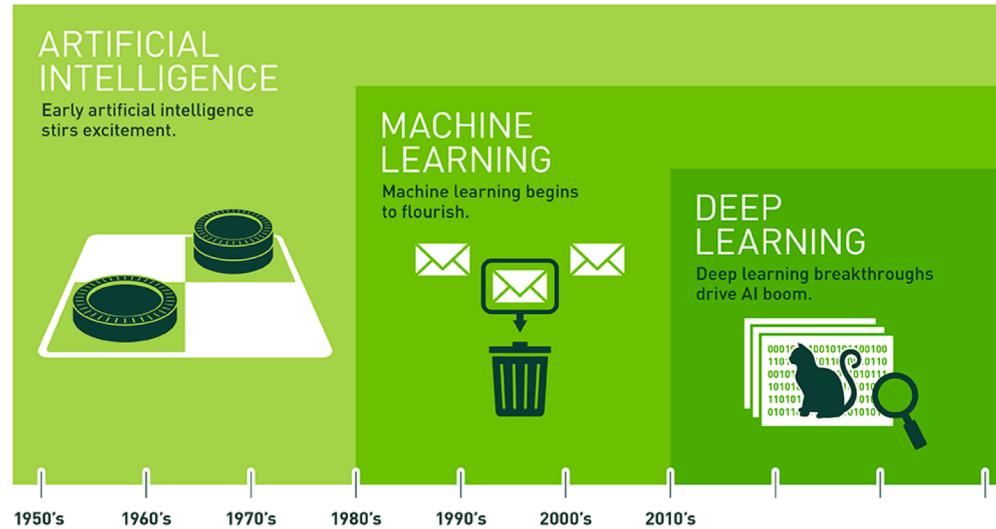


## But what is AI?

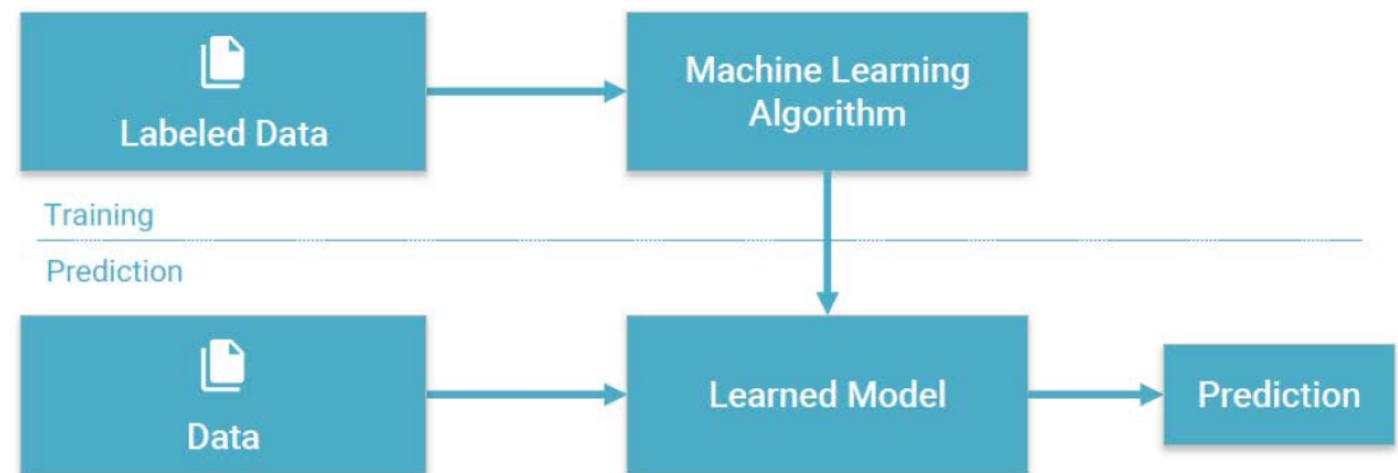
- At least 8 different definitions, according to the reference textbook
  - We define AI as the study of computer agents that interact with their environment, adapt to change, pursue goals and take actions to achieve the best expected outcome
  - AI is not new: its conventional birth date is the Dartmouth conference (1956)
  - Big enthusiasm in 1950-1970, but excessive optimism and failure to meet expectations led to AI winter in 1974
  - In the '80, success of expert systems (logic inference, symbolic AI) followed by another AI winter («expert systems work well in tightly defined subject areas [...] but are unable to learn concepts that children learn by the time they are 3 years old», Marvin Minsky)
-

# Machine Learning

- We skip through the rest of AI history
- The type of AI which is most successful today is **Machine Learning**, and in particular its subfield, **Deep Learning**



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



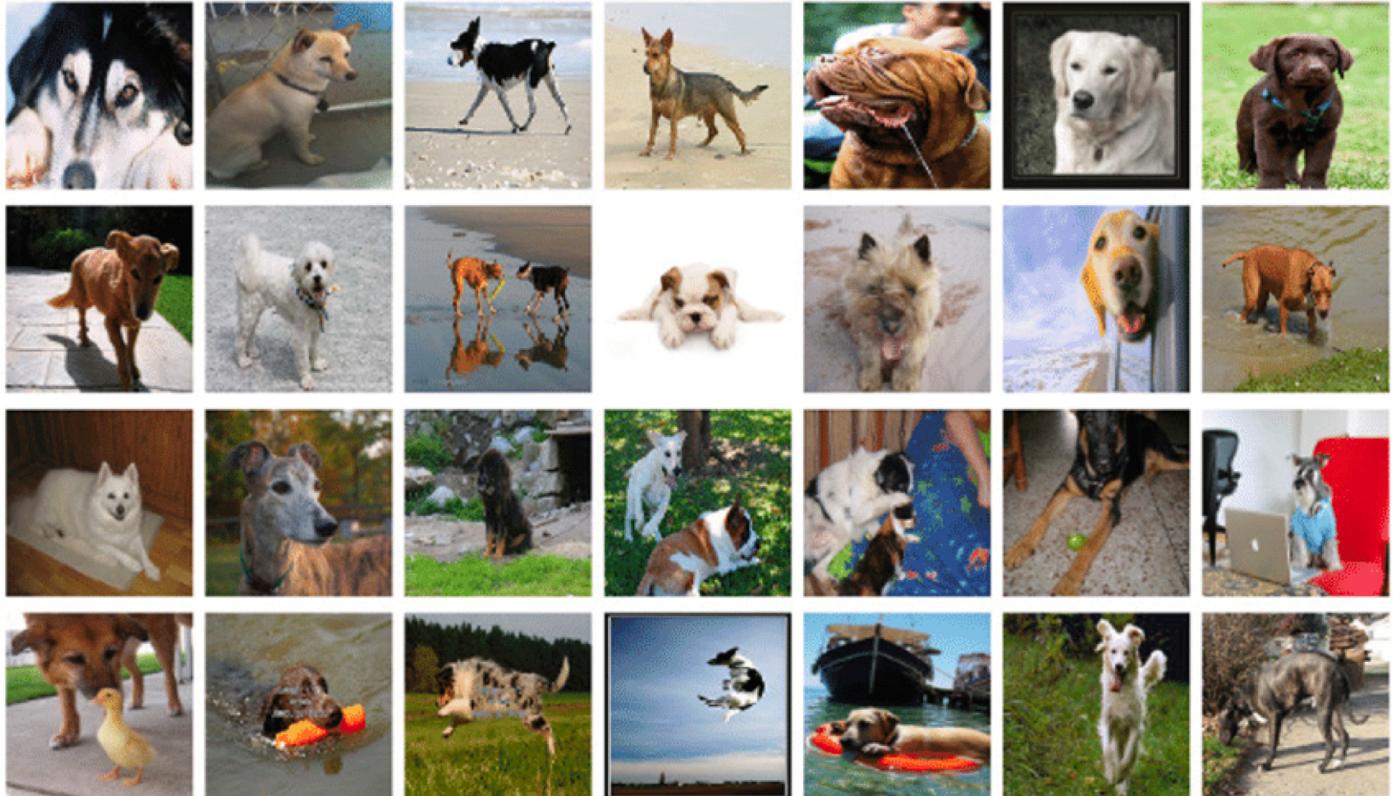
Provides various techniques that can learn from and make predictions on data

# What is Machine Learning?

- An algorithm is a sequence of steps, each of which is simple and well defined, which ends after a finite number of steps.
  - In many real-world problems, it is much easier to identify **desirable behavior** of an algorithm than to explicitly define steps. Example: identify the presence and the breed of a dog in a picture (see next slide)
  - A computer program is a **map** between input and output space
  - A computer program is a **point** in the space of such maps
  - **Search** the space of programs until desired behavior is obtained
  - Surprisingly, it works!
  - **Learning a function from examples** instead of explicit modelling
-

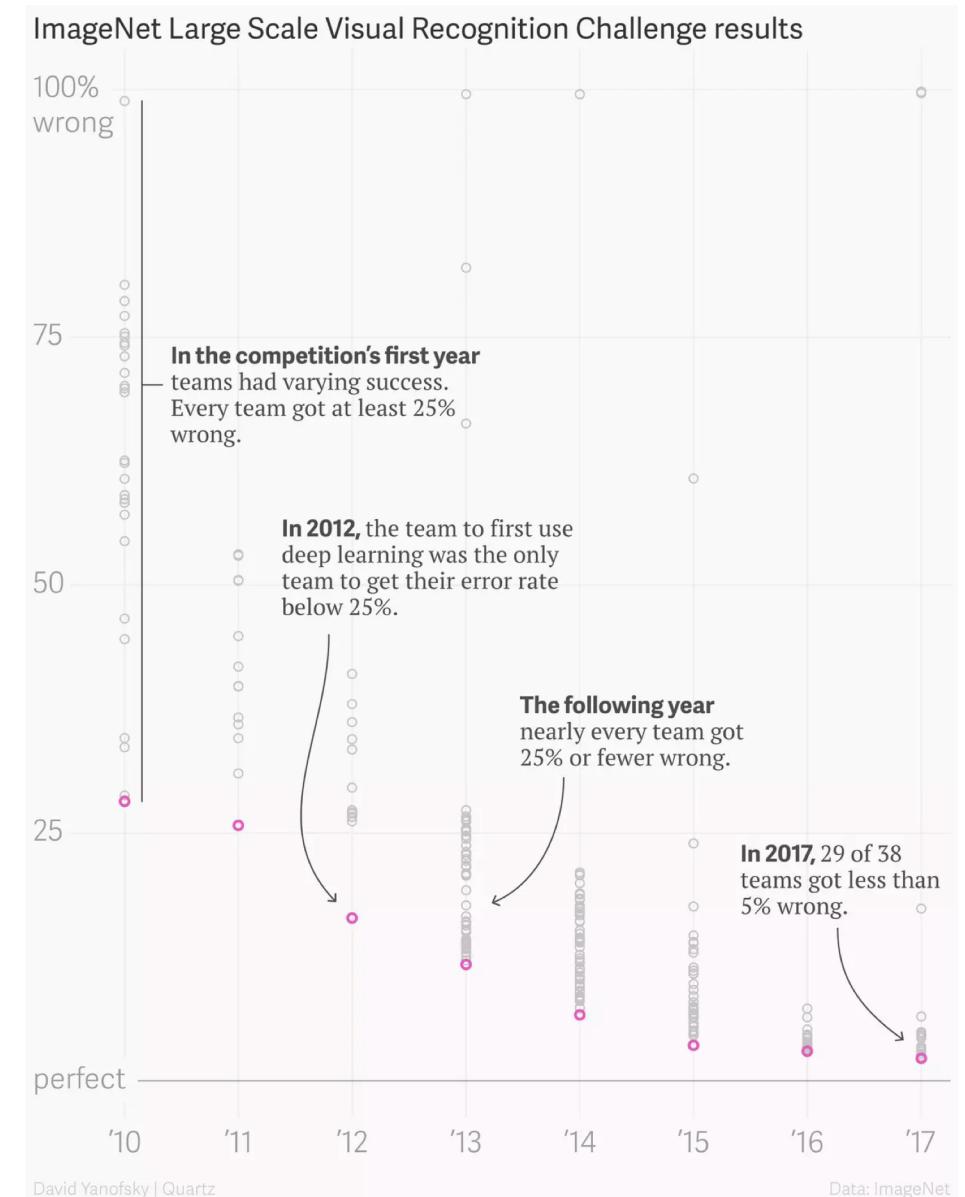
# ImageNet

- The event which led to contemporary success, global interest and hype about AI was the 2012 edition of the [ILSVRC competition](#), also known as ImageNet competition
- Classify correctly 1.2 millions of images, 1000 classes



## The “ImageNet moment”

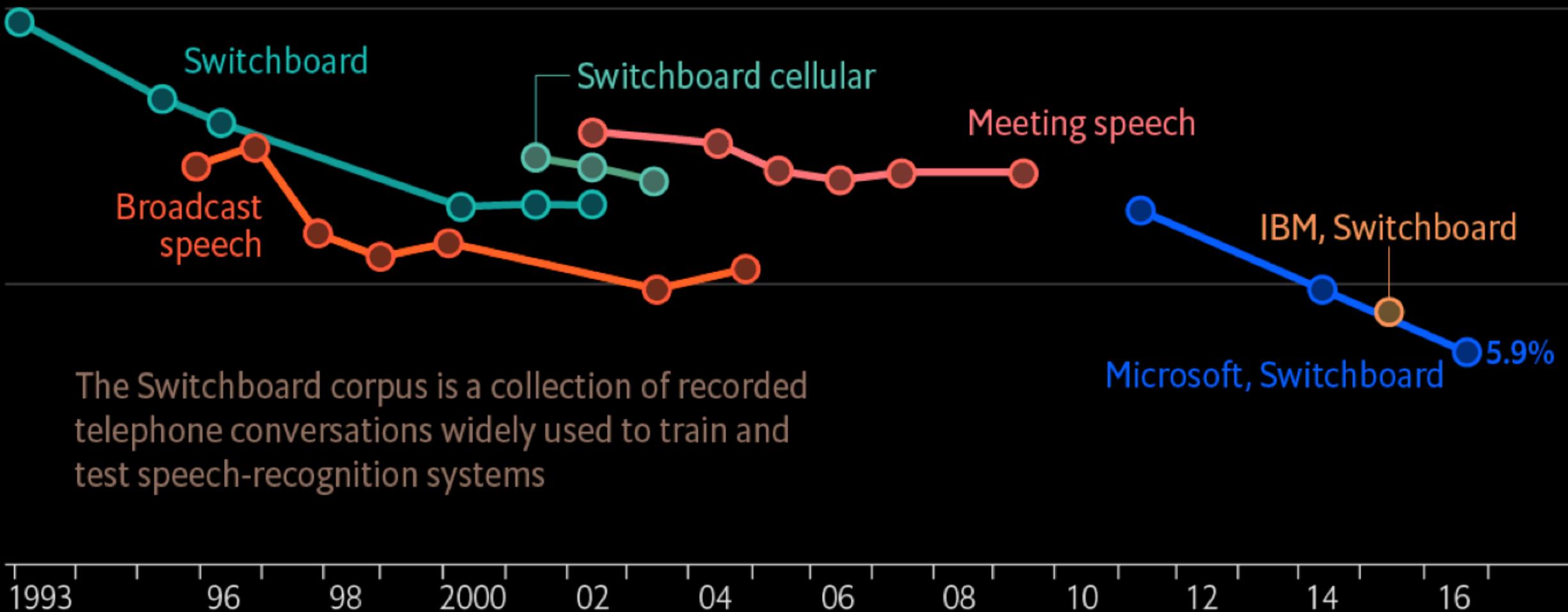
- In 2010 (ILSVRC first year), every competing team got at least 25% images wrong
- In 2012, using **Deep Learning**, team of AI legend Geoffrey Hinton reduced the error rate by nearly 50%
- This was a quantum leap, with respect to usual advance rate in AI
- Deep Learning has been used by ILSVRC winning teams since 2012, and it's now being used in all the fields seen in [slide 3](#)



# Speech-recognition word-error rate, selected benchmarks, %

Log scale

100

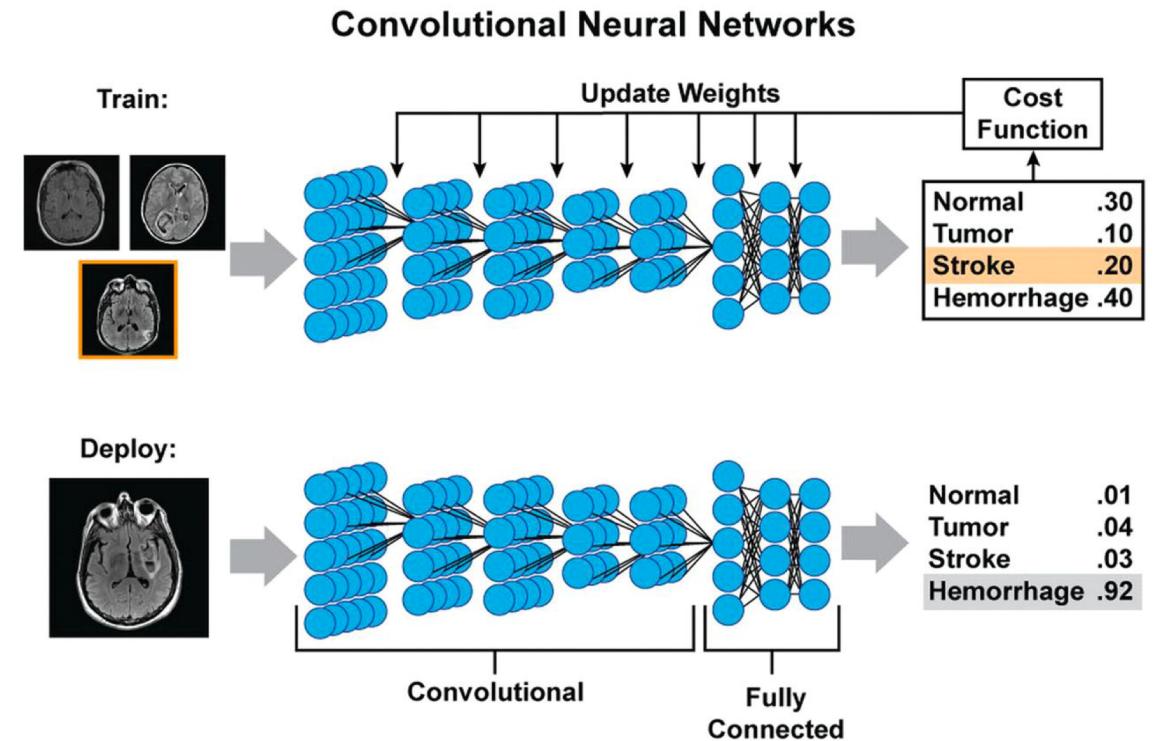


The Switchboard corpus is a collection of recorded telephone conversations widely used to train and test speech-recognition systems

Sources: Microsoft; research papers

# What is Deep Learning?

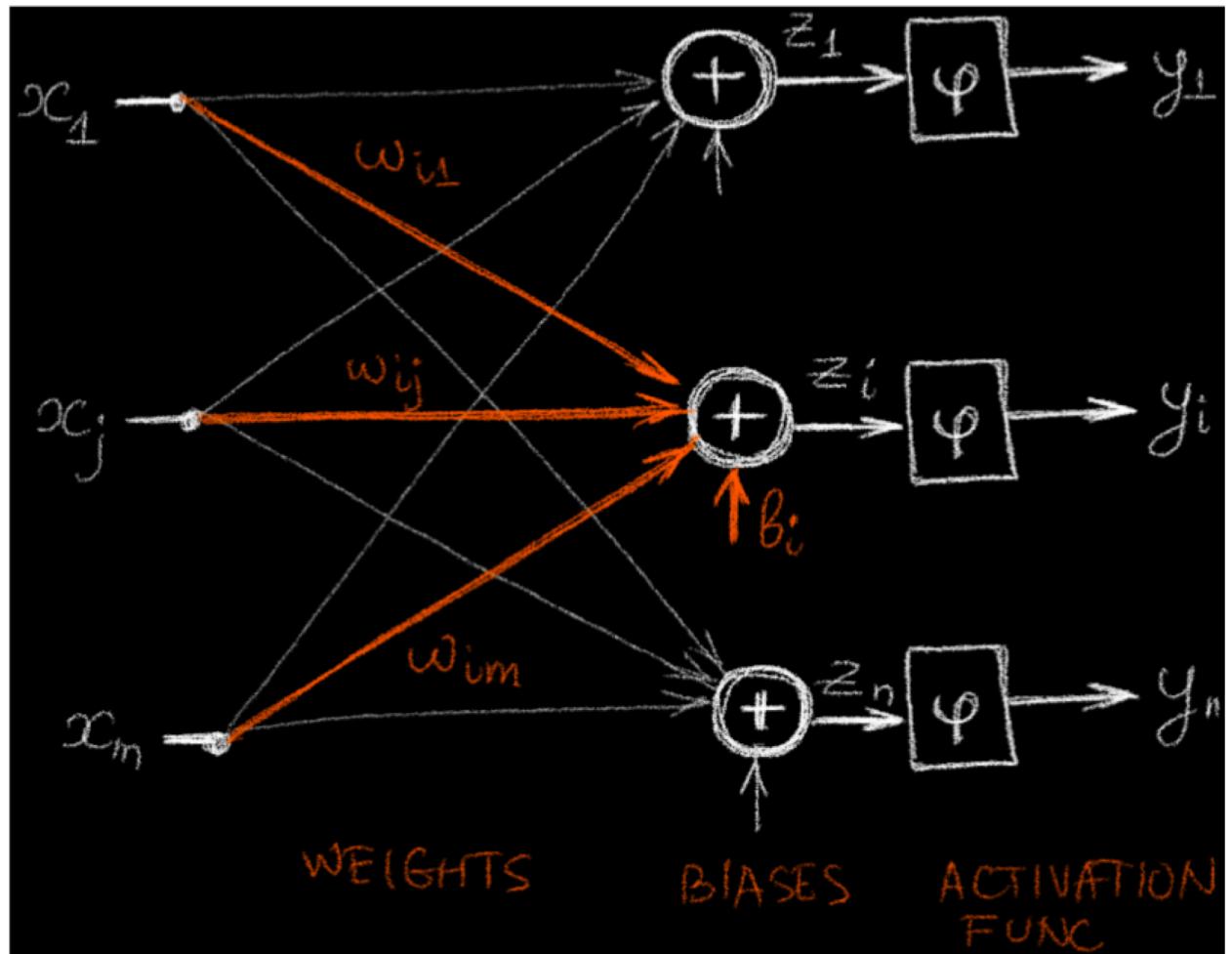
- Using **huge** ( $O(10^6) \div O(10^9)$ ) of parameters) multilayer neural networks (NN) as the trainable model in Machine Learning
- Backpropagation-trained NN have been around since at least 1986, however the «old» NN were **much** smaller (~ 25 parameters)
- Only later, hardware (GPU), large amount of data **and** algorithmic innovations (SGD, dropout, BatchNorm, residual connections, Fixup initialization, etc.) made possible to train **Deep** (multiple layers) neural networks



The huge amount of money spent by FAANG on Deep Learning is making the field progress as never before (BigGAN, VQ-VAE-2, BERT, GPT-2, XLNet, RoBERTa, etc.)

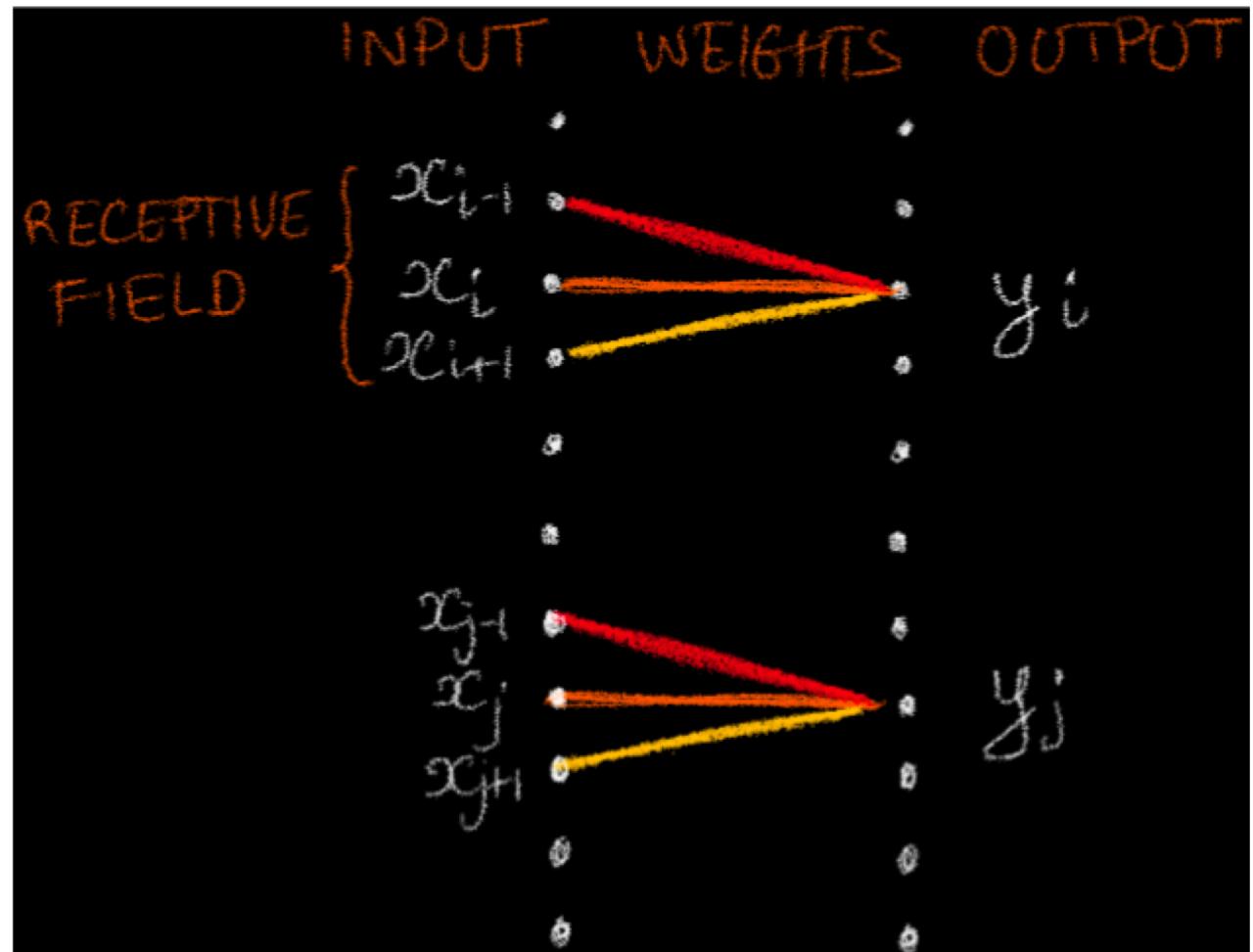
# Image classification

- An image of width  $W$ , height  $H$  and with  $C$  channels can be represented as a  $W \times H \times C$  tensor
- Images are **BIG**: a low-res  $256 \times 256$  RGB image is a tensor of  $3 \times 256 \times 256 \approx 2 \cdot 10^5$  components
- A classical MLP (fully connected, feed-forward neural network) with 512 neurons in the first hidden layer, would use  $(3 \times 256 \times 256 + 1) \times 512 \approx 10^8$  weights only in the first hidden layer, and we want to go deep!!



# Weight sharing

- Deep Learning is **HUNGRY**:  $N = 1.2 \cdot 10^6$  for ILSVRC!\*
- It turns out we can reduce the # of weights as well as the size of the dataset with the same idea, **translation equivariance**
- $\tau_p \mathcal{H} = \mathcal{H} \tau_p \quad \forall p \in \mathbb{Z}^d$
- **Translation-equivariant** representations simplify the construction of **translation-invariant** classifiers
- Also, representations should be sensitive to **local** features: the pixels identifying an eye in a face are adjacent
- Q: how do we get **locality** and **translation-equivariance**?
- A: weight sharing, a.k.a **filters**



# Convolutional layer

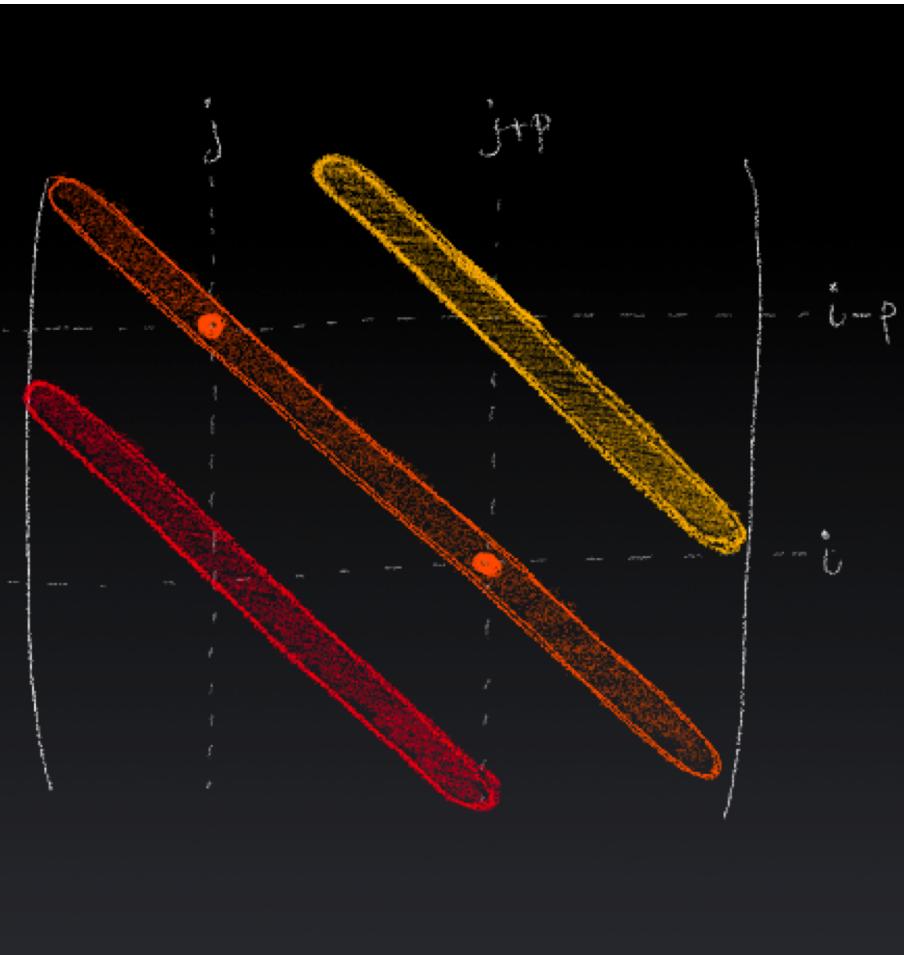
## Toeplitz operators

$$y_{\mathbf{i}}^k = \sum_{l=1}^m \sum_{\mathbf{j} \in \mathbb{Z}^d} w_{\mathbf{i}-\mathbf{j}}^{kl} x_{\mathbf{j}}^l + b^k$$

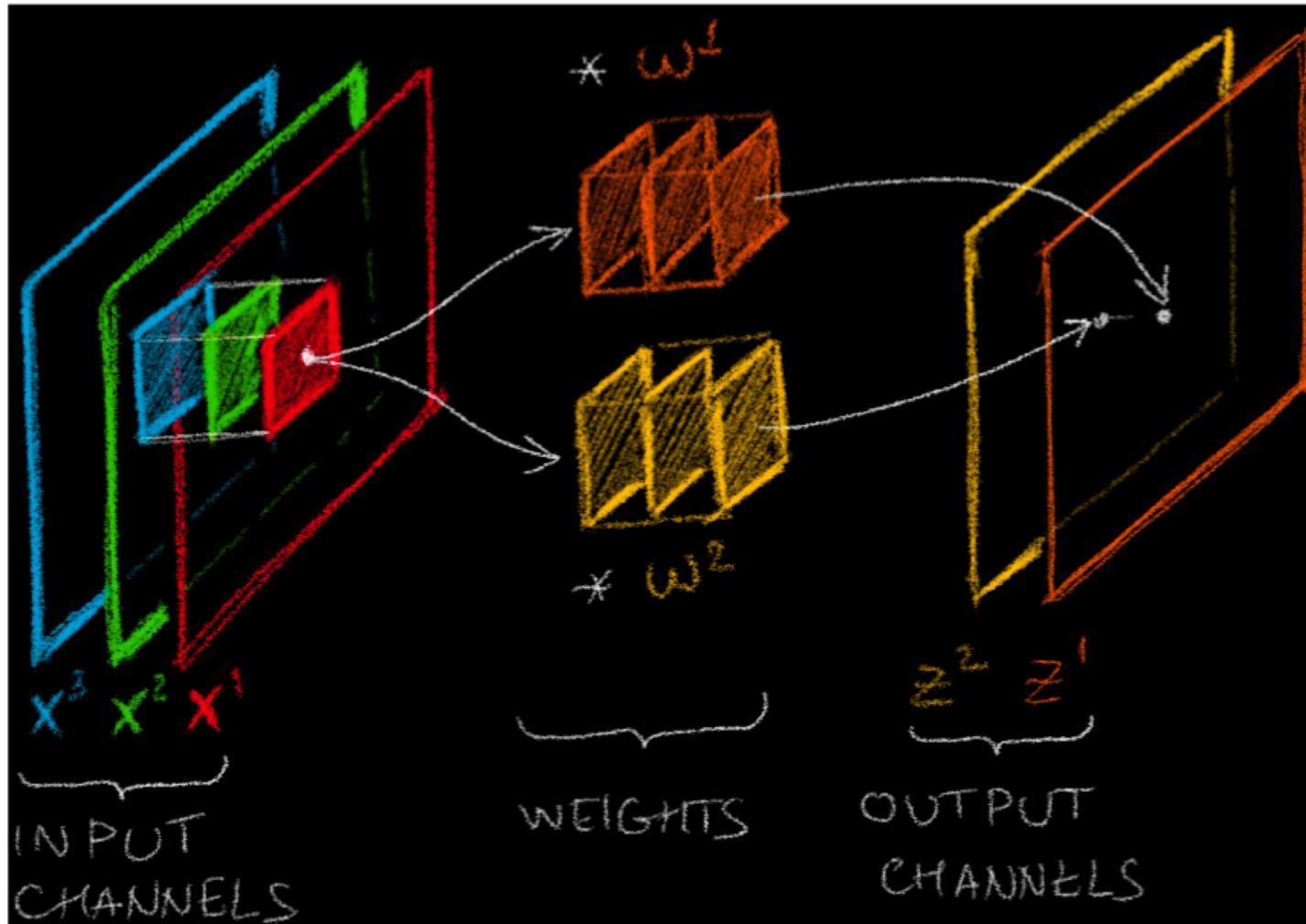
$$= \sum_{l=1}^m \sum_{\mathbf{j} \in \mathbb{Z}^d} w_{\mathbf{j}}^{kl} x_{\mathbf{i}-\mathbf{j}}^l + b^k$$

### Convolution

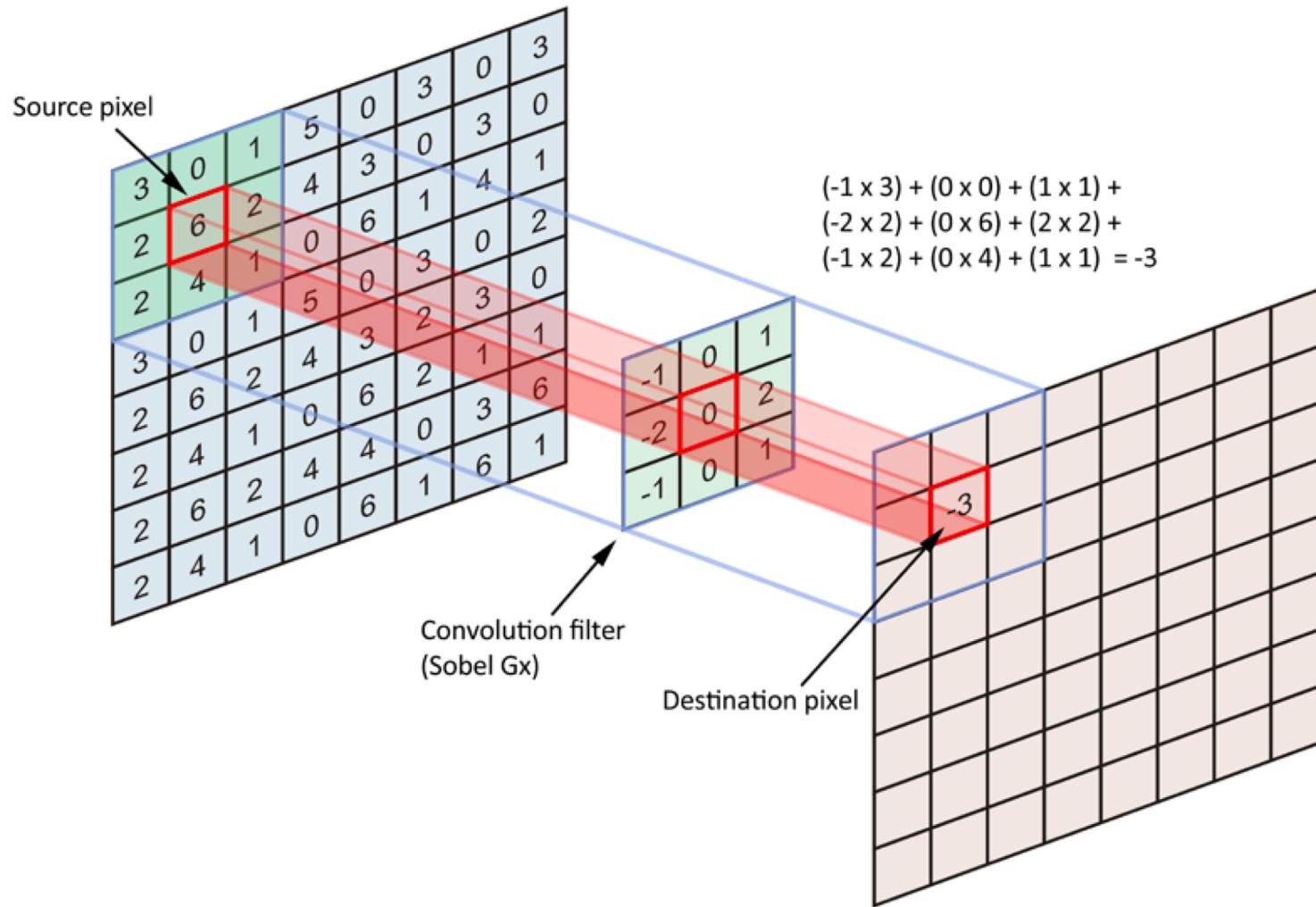
$$\mathbf{y}^k = \sum_{l=1}^m \mathbf{w}^{kl} * \mathbf{x}^l + b^k$$



## Convolutional layer



# Convolutional layer

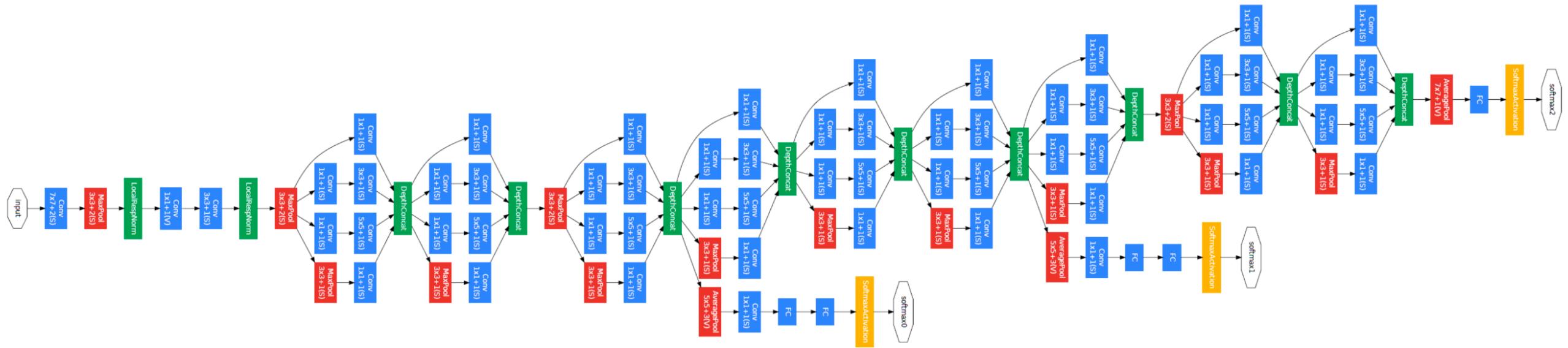


# Convolutional Neural Networks

- Stacking multiple convolutional layers one after the other gives us a **Convolutional Neural Network (CNN)**
  - In practice, we need many other layers to build a good image classifier:
    - batch normalization/spectral normalization
    - max pooling, global pooling
    - strided convolutions
    - atrous convolutions
    - skip/residual connections, which enhance the CNN **topological expressivity**
    - dropout (rapidly becoming obsolete)
    - new initializations, training schedules, label modifications, data augmentation tricks, etc.
  - Listing them all would take way too much time
  - Let's just mention two of the most famous/robust CNN architectures in use today
-

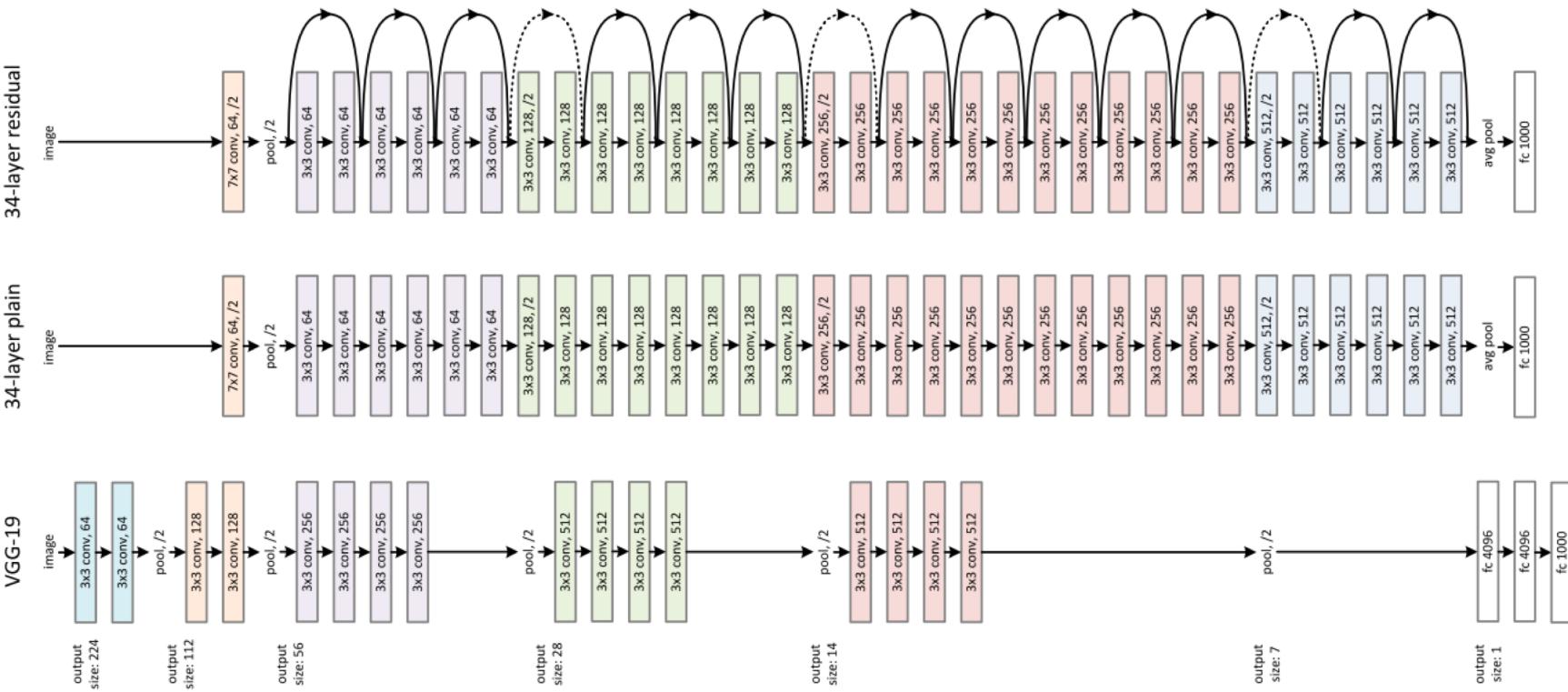
# Inception

- From Google AI
  - various version, figure shows v1 (2014) but v3 is the most commonly used today



# ResNet

- a family of architectures from Microsoft Research, which introduced the residual connection
- ResNet-50 (50 layers) is the most commonly used today, but ResNet-34 (picture) and ResNet-101 are also widely adopted



## Why should we care?

- Machine Learning and in particular Deep Learning are fueling a revolution in many businesses, including Oil & Gas
  - Customers know about AI, and are asking O&G companies to provide AI-powered services
  - OEMs are already using it to strengthen their offering (Siemens GT low NOx emissions)
  - DL will be key to building a technology competitive advantage for TPS products and services
-

# Conclusions

- AI is changing the world we live in, including the way we do business
- In order to build a technology competitive advantage for their products and services, O&G companies need to build strong AI skills now
- By developing focused AI demonstrators, promoting a culture of AI and fostering external collaborations and partnership, our team is making sure that we stay on top of this rapidly evolving technology





That's all Folks!