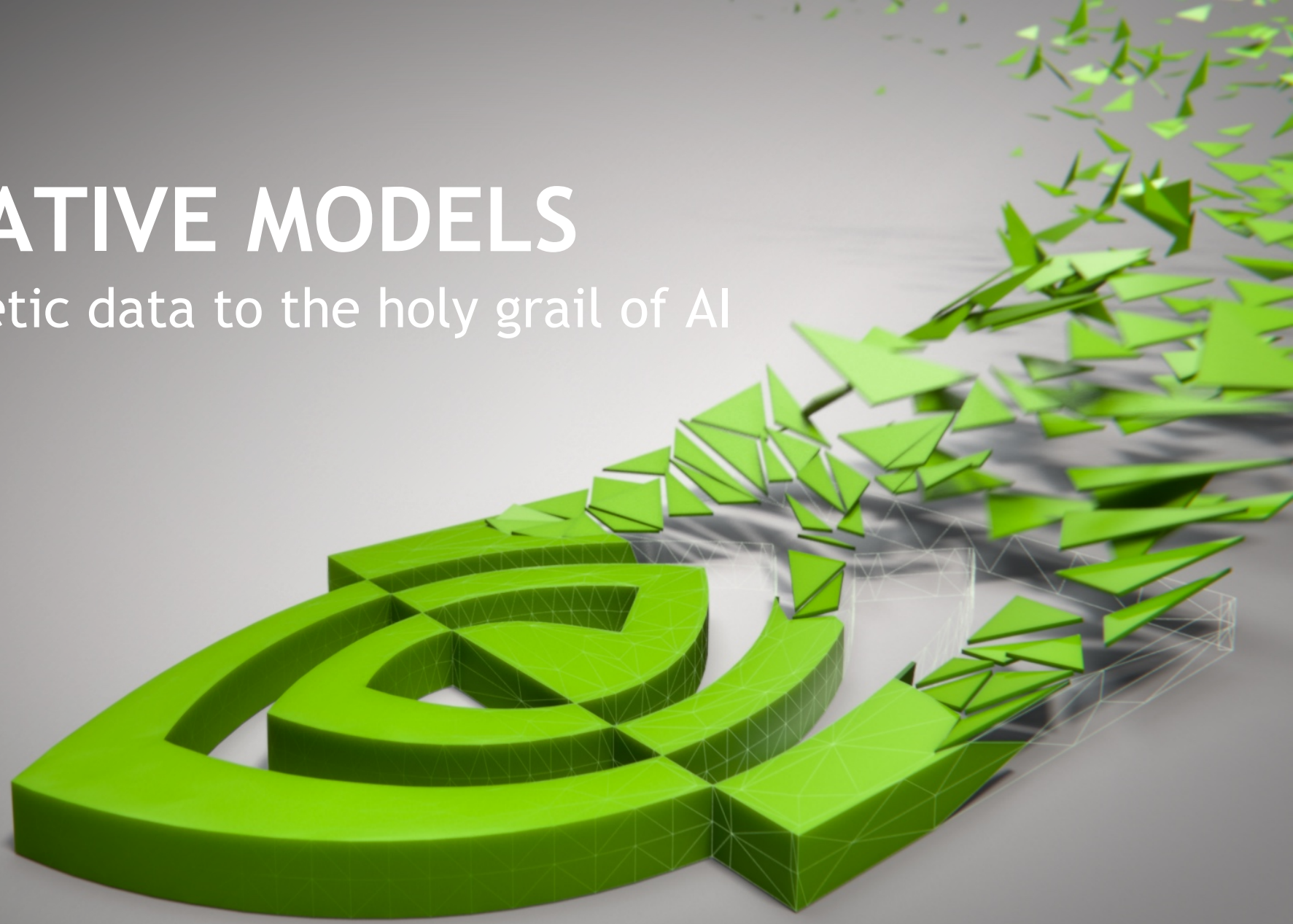
# GENERATIVE MODELS

From synthetic data to the holy grail of AI

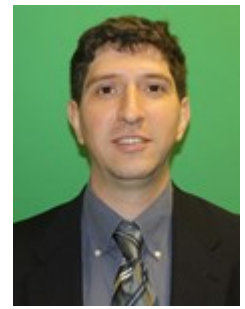
NVIDIA



# ABOUT ME

Aleksandr Volkov – avolkov@nvidia.com

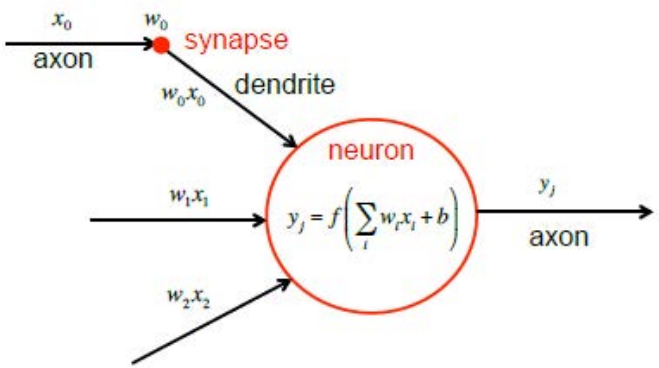
- Solution Architect @ NVIDIA - Supporting delivery of AI / Deep Learning solutions
- 10 years experience working on HPC (high performance computing) systems, signal processing, and Machine Learning
- My past experience:
  - Northrop Grumman – Systems Engineer
  - Exxon Mobil – HPC Programmer

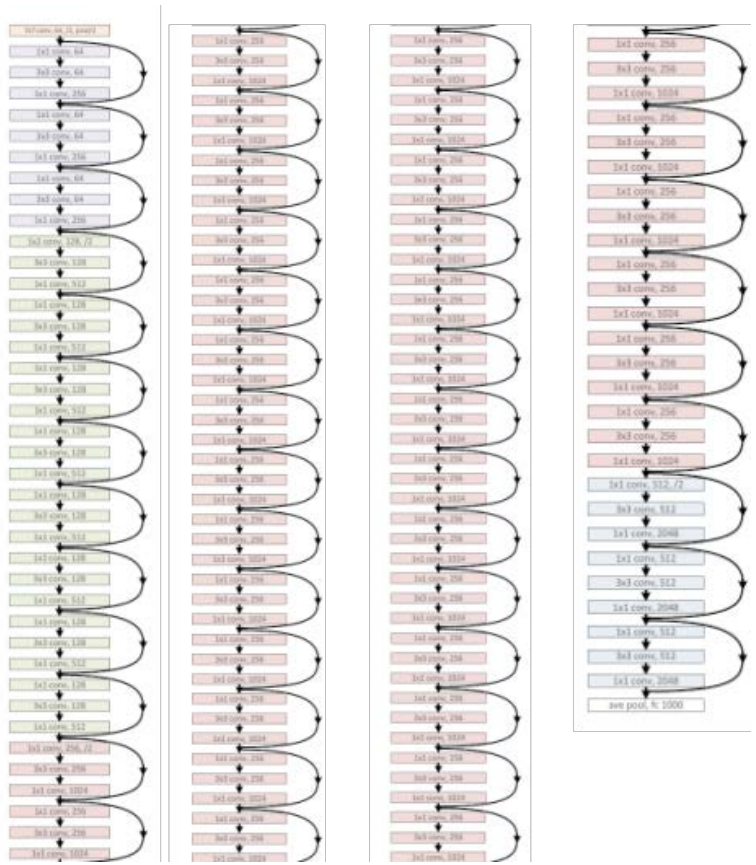


# WHY DO NEURAL NETWORKS WORK?

# NEURAL NETWORKS ARE NOT NEW

## And are disapintingly simple as an algorithm



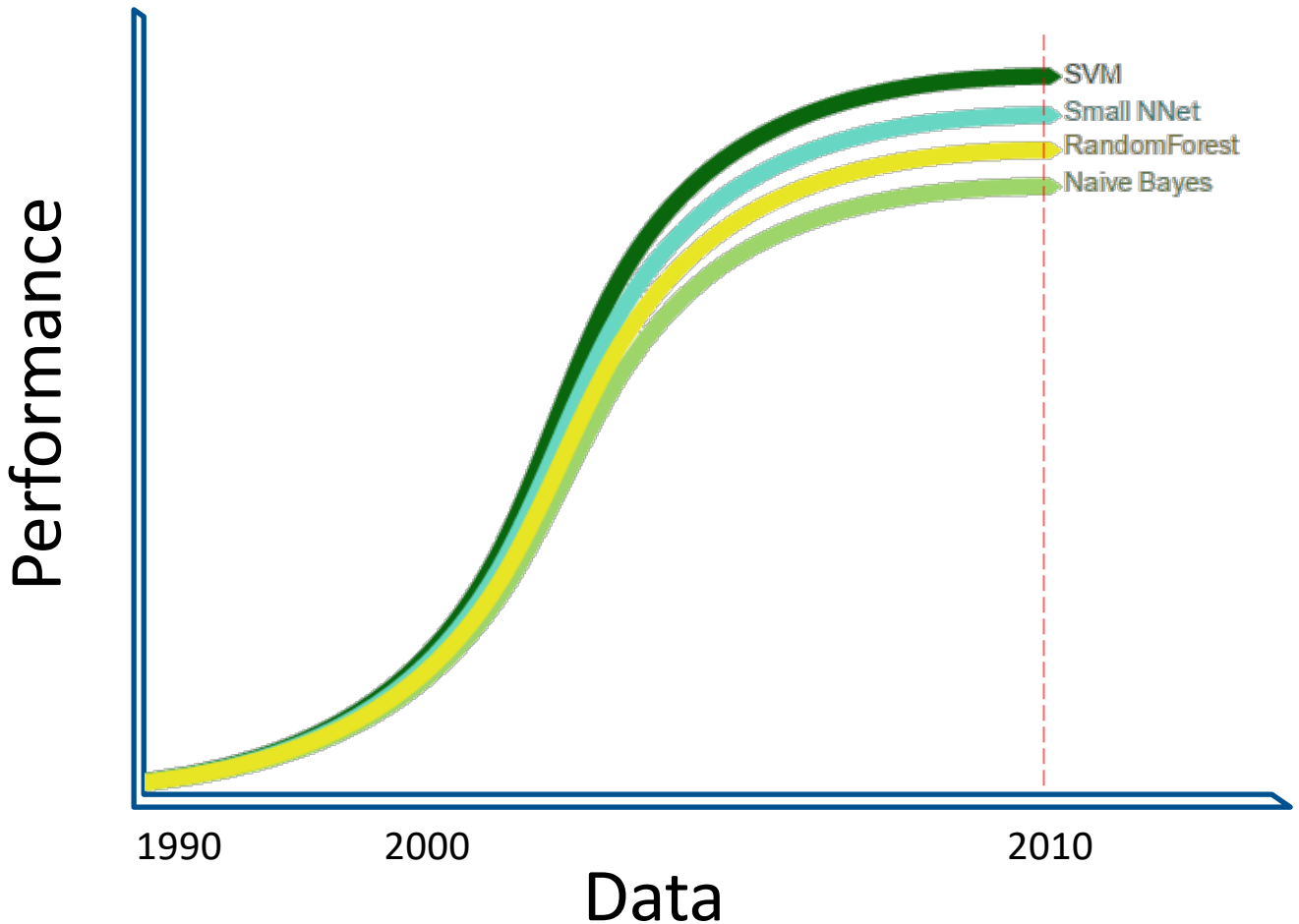
# SUPERVISED LEARNING

## Approximating complex functions

$$y = f(x)$$

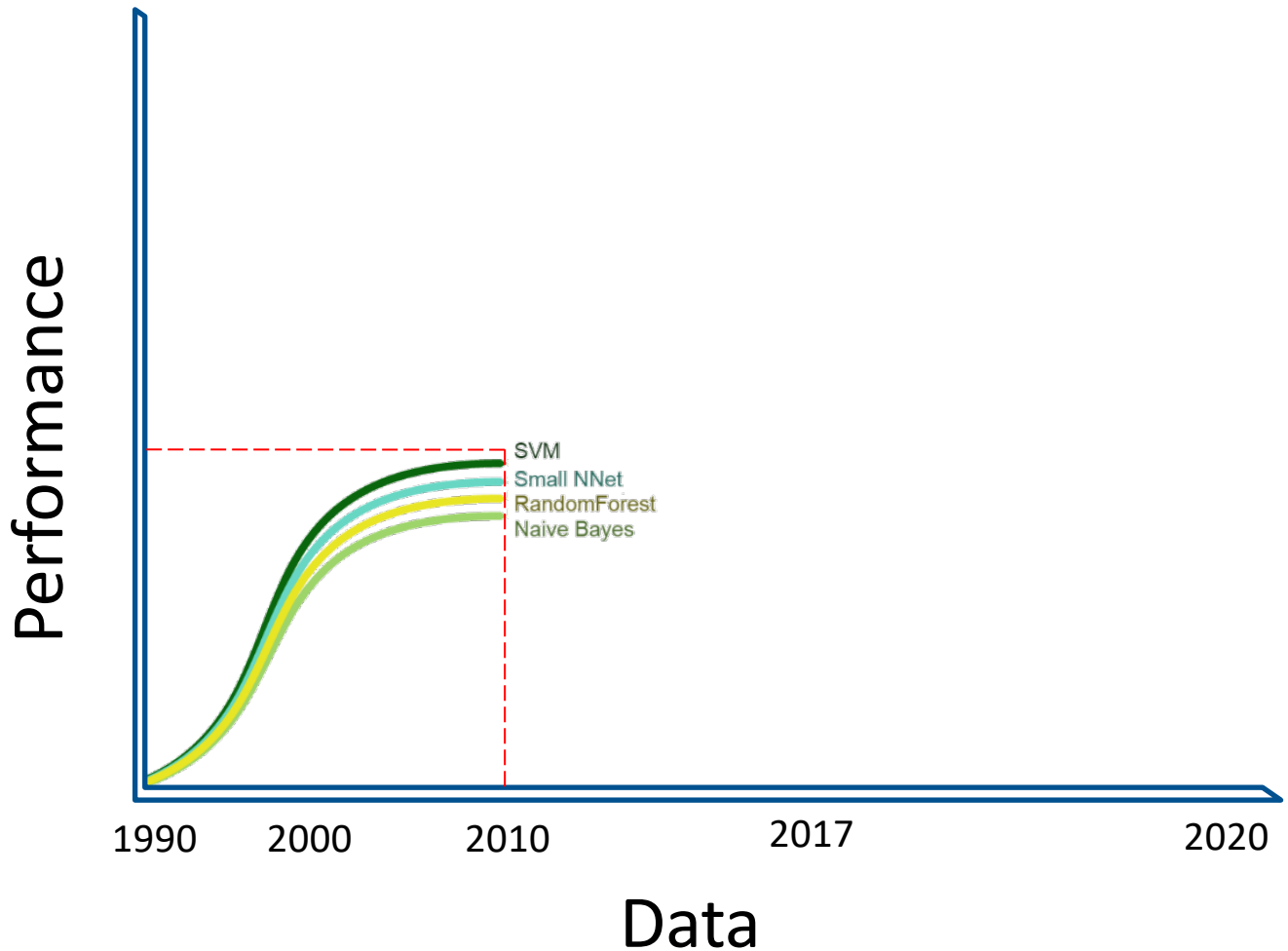
# SMALL NEURAL NETWORKS

## Underperform



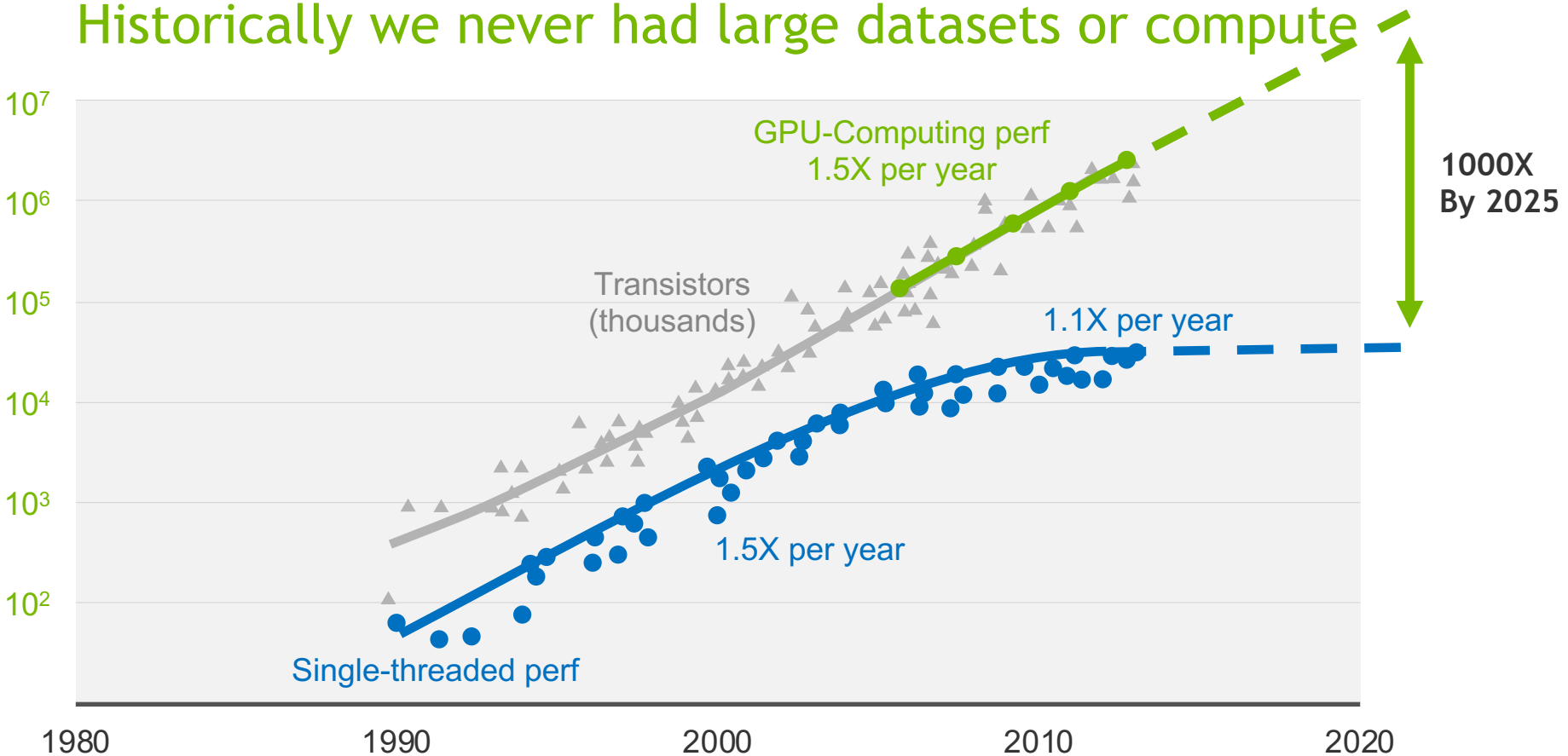
# WHY?

# DATA


Performance
SVM
Small NNet
RandomForest
Naive Bayes
1990
2000
2010
2017
2020
Data

# COMPUTE

## Historically we never had large datasets or compute



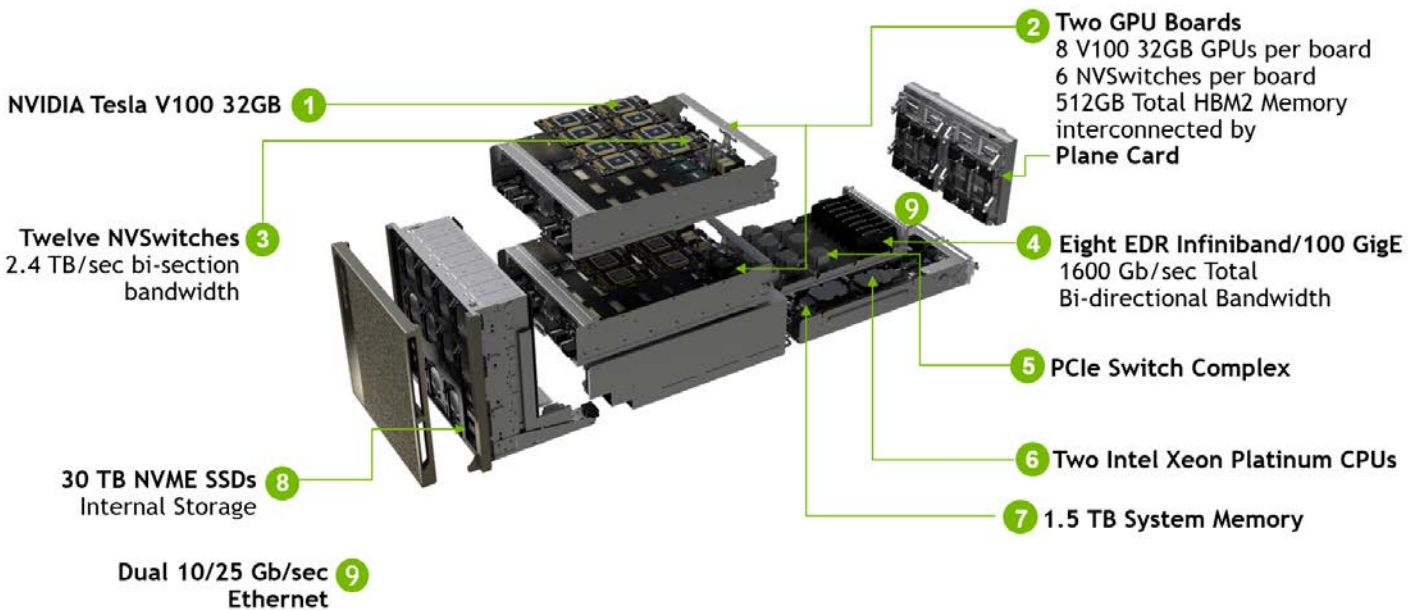
# PERSPECTIVE

# CONTEXT

## 1.759 petaFLOPs in November 2009

# CONTEXT

## 2 petaFLOPs - today



# CONTEXT TODAY

## 200 petaFLOPs in November 2018

Achitecture:

9,216 POWER9 22-core CPUs

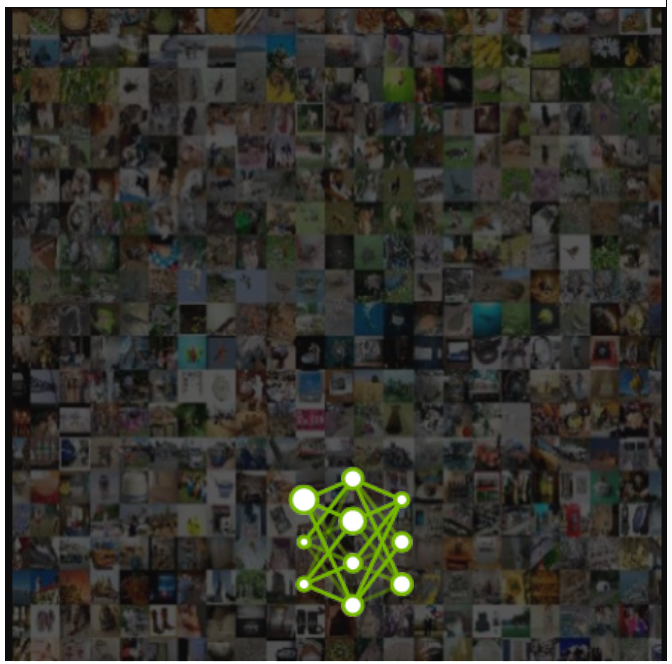
27,648 Nvidia Tesla V100 GPUs

Approaching 3.3 exaops using mixed precision

# NEURAL NETWORK COMPLEXITY IS EXPLODING

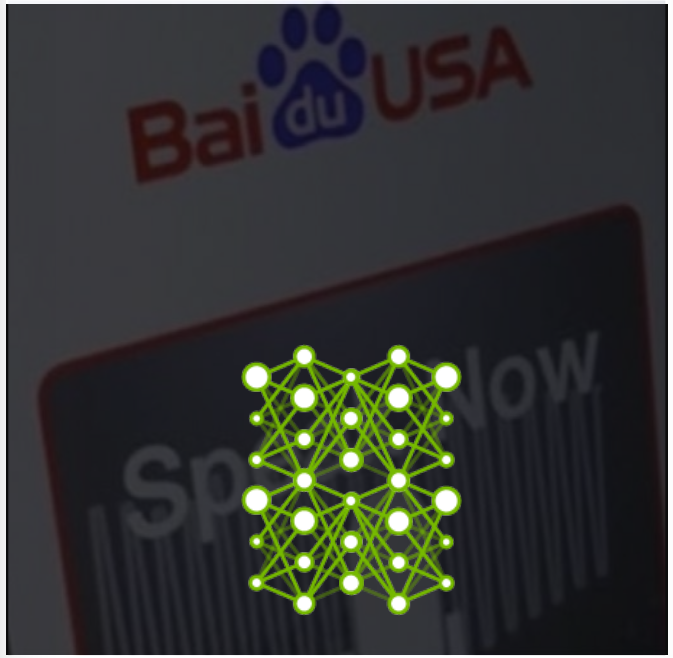
## To Tackle Increasingly Complex Challenges

7 ExaFLOPS
60 Million Parameters



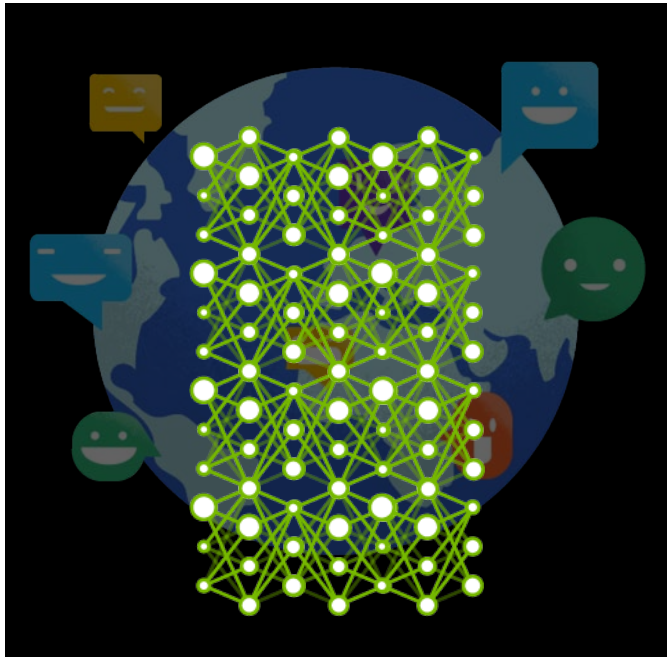
**2015 – Microsoft ResNet**
**Superhuman Image Recognition**

20 ExaFLOPS
300 Million Parameters




**2016 – Baidu Deep Speech 2**
**Superhuman Voice Recognition**

100 ExaFLOPS
8700 Million Parameters

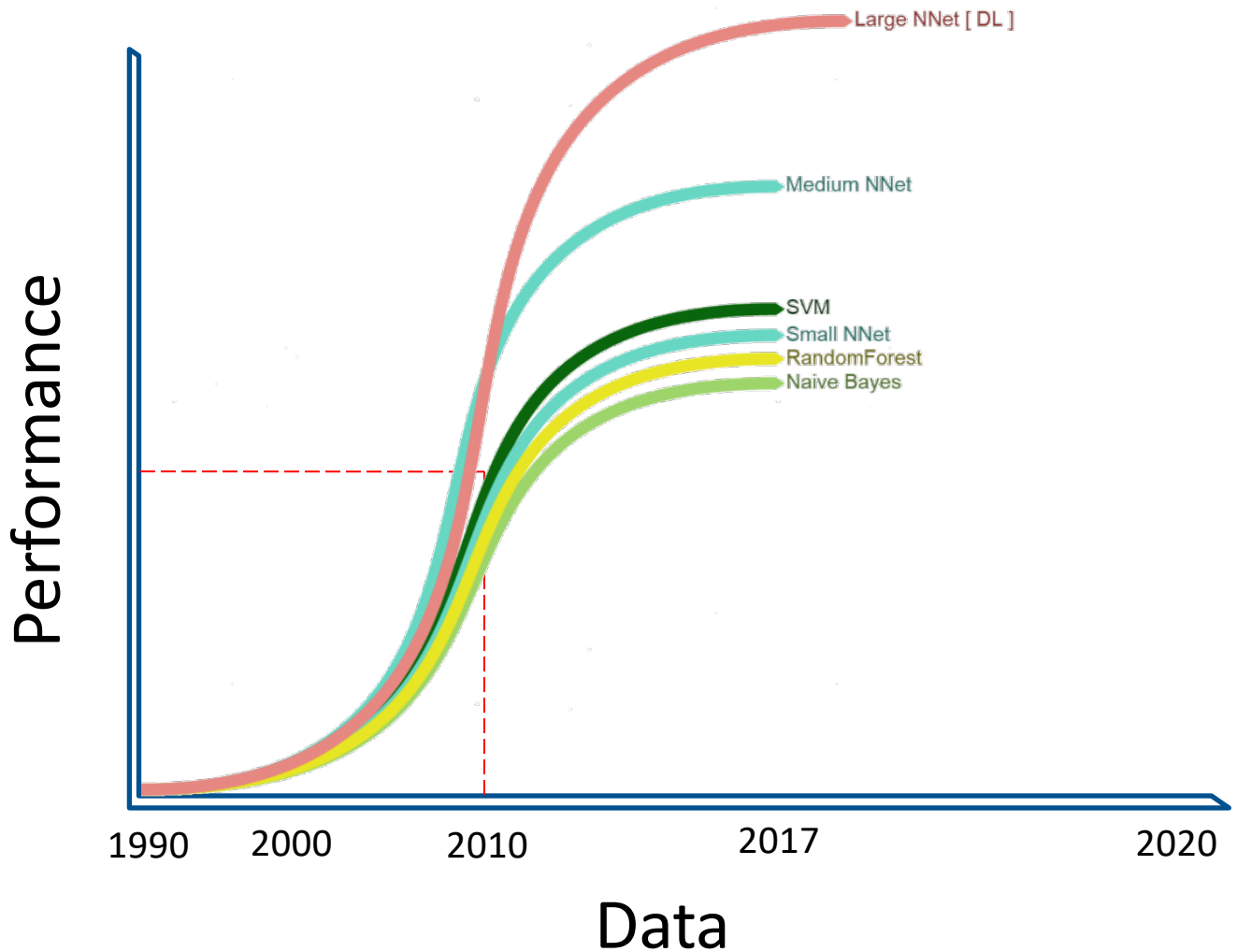


**2017 – Google Neural Machine Translation**
**Near Human Language Translation**

100 EXAFLOPS
=
2 YEARS ON A DUAL CPU SERVER

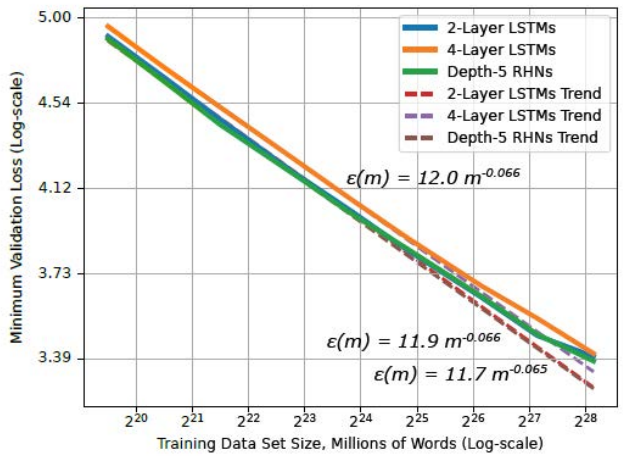
# NEURAL NETWORKS ARE NOT NEW

## Require abundance of data and compute

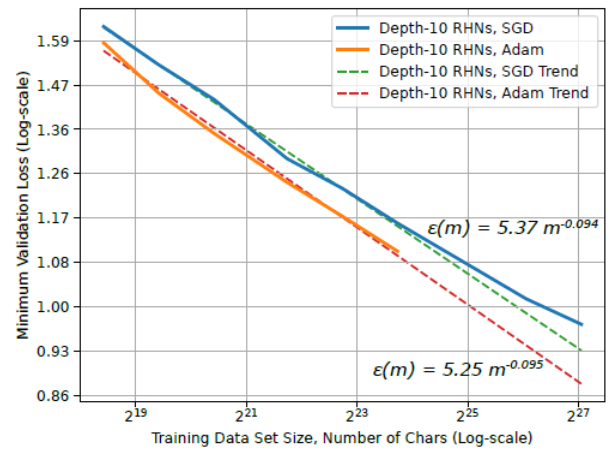


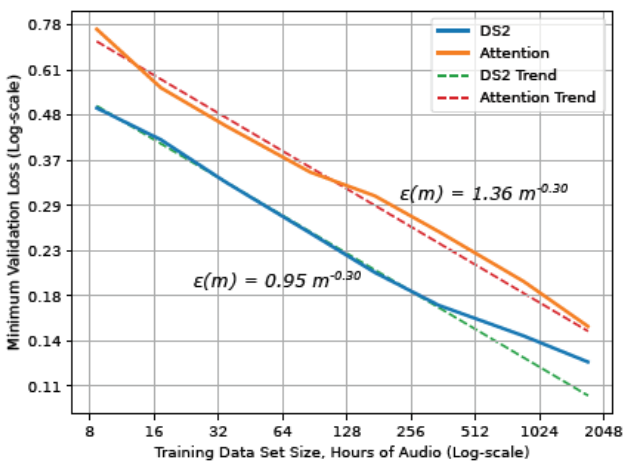
# EXPLODING DATASETS

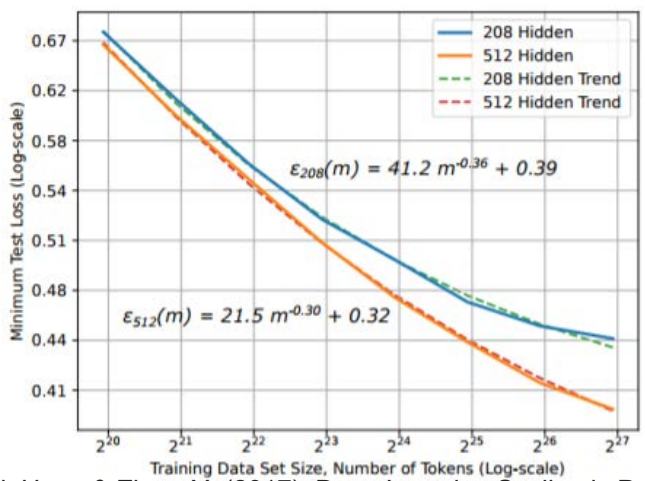
## Logarithmic relationship between the dataset size and accuracy

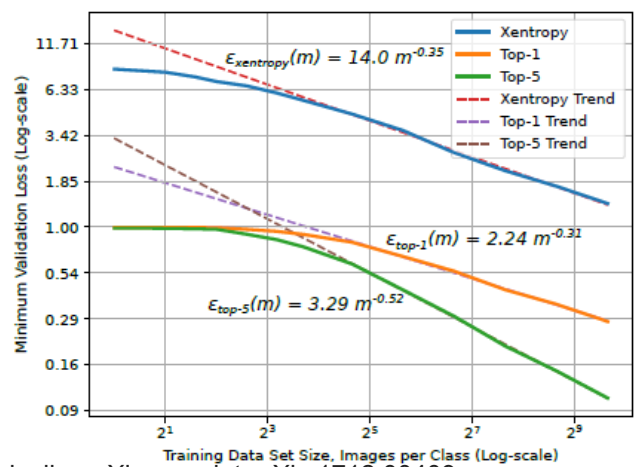



- Translation
- Language Models
- Character Language Models
- Image Classification
- Attention Speech Models



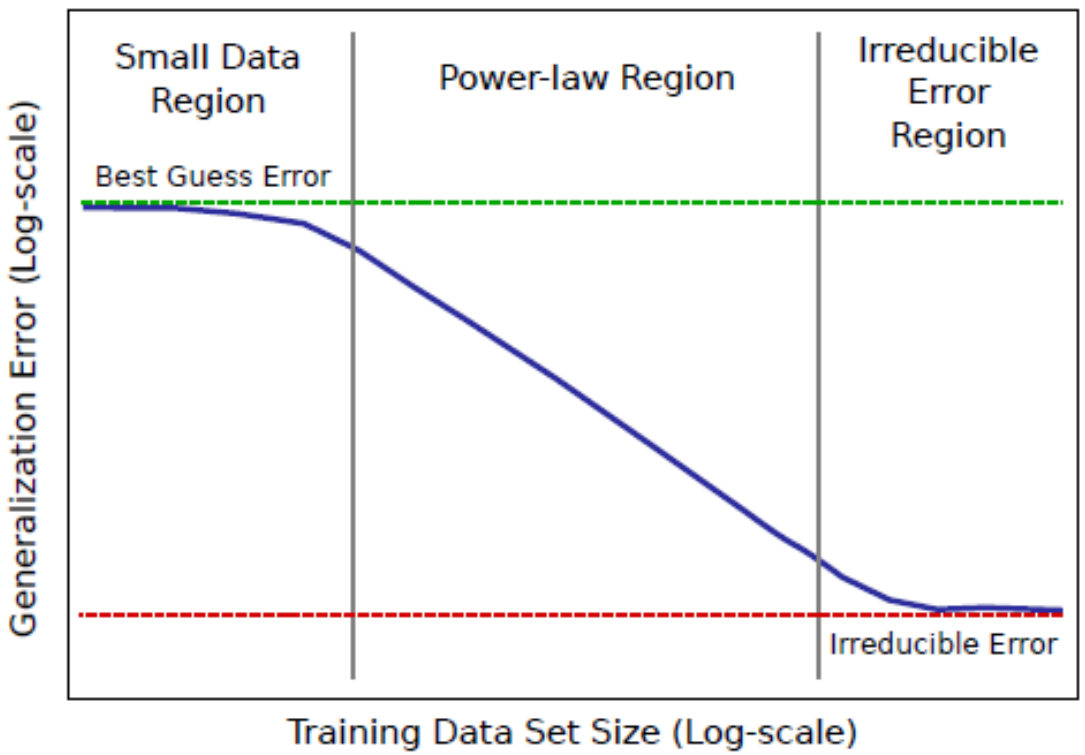




Hestness, J., Narang, S., Ardalani, N., Diamos, G., Jun, H., Kianinejad, H., ... & Zhou, Y. (2017). Deep Learning Scaling is Predictable, Empirically. arXiv preprint arXiv:1712.00409.

# EXPLODING DATASETS

## Logarithmic relationship between the dataset size and accuracy




Hestness, J., Narang, S., Ardalani, N., Diamos, G., Jun, H., Kianinejad, H., ... & Zhou, Y. (2017). Deep Learning Scaling is Predictable, Empirically. arXiv preprint arXiv:1712.00409.

# MAKING COMPLEX PROBLEMS EASY

# TRANSFORMING IMPOSSIBLE INTO EXPENSIVE

# SUPERVISED LEARNING

## Approximating complex functions

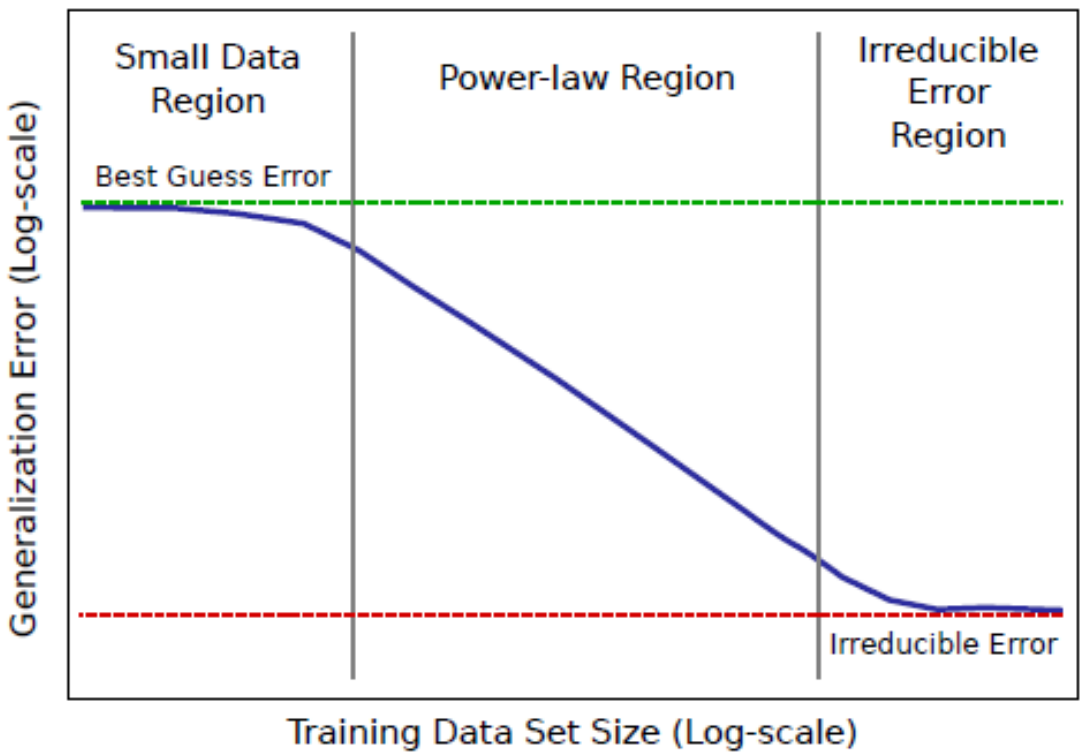
$$y = f(x)$$

# HUGE OPPORTUNITY

# HUGE CHALLENGE

# EXPLODING DATASETS

## Logarithmic relationship between the dataset size and accuracy




Hestness, J., Narang, S., Ardalani, N., Diamos, G., Jun, H., Kianinejad, H., ... & Zhou, Y. (2017). Deep Learning Scaling is Predictable, Empirically. arXiv preprint arXiv:1712.00409.

# A BETTER WAY

# GENERATIVE MODELS

## Reformulating the problem

$$y = f(x)$$

# WE NO LONGER TRAIN OUR MODEL TO SOLVE A NARROW TASK

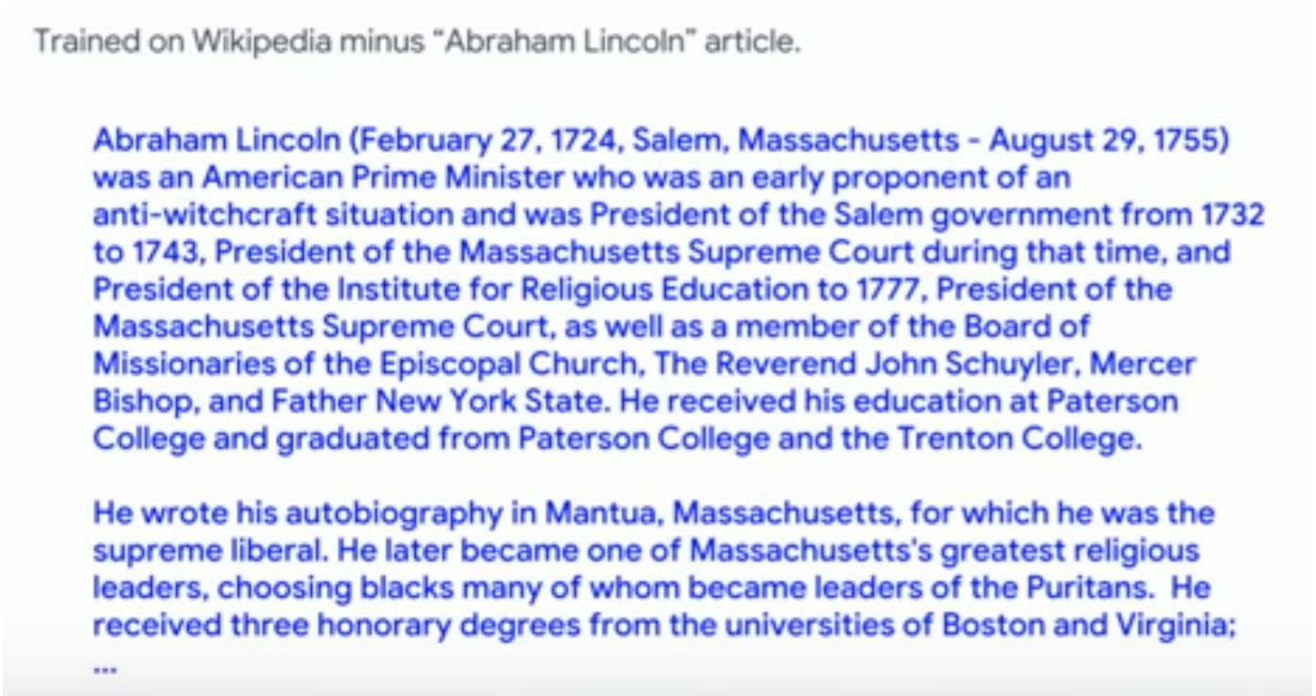
# WE TRAIN IT TO LEARN A PHENOMENON (ITS PROBABILITY DISTRIBUTION)

# TEACHING MACHINES HOW TO COMPOSE HUMAN LANGUAGE

# UNDERSTANDING LANGUAGE

## Generating text

Transformer LM – 100M Parameters



# UNDERSTANDING LANGUAGE

## Generating text

Transformer LM – 5B Parameters

Trained on Wikipedia minus "Abraham Lincoln" article.

Abraham Lincoln (March 1, 1809 - March 15, 1865) was an American lawyer, jurist, and politician. He served as the 19th President of the United States, from March 1861 until his assassination in April 1865.

Born in Newark, New Jersey, Lincoln attended Knox College of New York, studied law, and was admitted to the bar Missouri Territory in 1826. He relocated to Illinois, and was elected to the Illinois House of Representatives in 1827, 1822, 1823, and 1826. He was elected Speaker of the Illinois House in 1825, served as Attorney General of the Illinois Territory in 1829, and was elected to the Illinois Senate in 1831. Lincoln was elected to the United States House of Representatives as a Democrat in 1833, and he was elected as a representative in the United States House of Representatives, serving two terms from 1839 to 1843. He was elected as a Whig in 1842, but he only served one term in Congress, and returned to the House of Representatives, serving

...

*“Due to concerns about large language models being used to generate deceptive, biased, or abusive language at scale, we are only releasing a much smaller version of GPT-2 along with sampling code. We are not releasing the dataset, training code, or GPT-2 model weights.”*

*“Better Language Models and Their Implications”, OpenAI*

# TEACHING MACHINES HOW TO COMPOSE MUSIC

# UNDERSTANDING BEAUTY

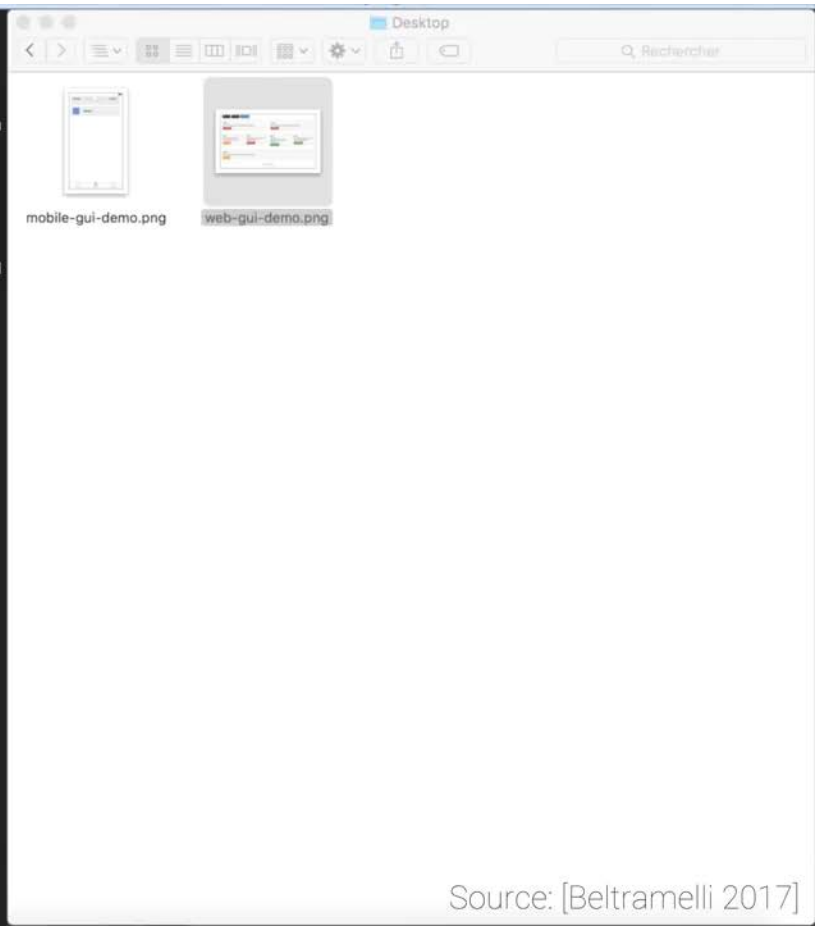
## Composing music

# TEACHING MACHINES HOW TO GENERATE CODE

# UNDERSTANDING CODE

## Generating computer code

```
pix2code:model computer$ ./pix2code
Error: no arguments supplied
Usage:
        -i|--input <input image path>
        -t|--target <target platform (ios|android|web)>
pix2code:model computer$ ./pix2code --target ios --input /Users/computer/Desktop/mobile-gu
i-demo.png
Input image: /Users/computer/Desktop/mobile-gui-demo.png
Target platform: ios
Generating code with pre-trained model...
Compiling output for target platform...
Code generated successfully
pix2code:model computer$ ./pix2code --target web --input /Users/computer/Desktop/web-gui-d
emo.png
Input image: /Users/computer/Desktop/web-gui-demo.png
Target platform: web
Generating code with pre-trained model...
```

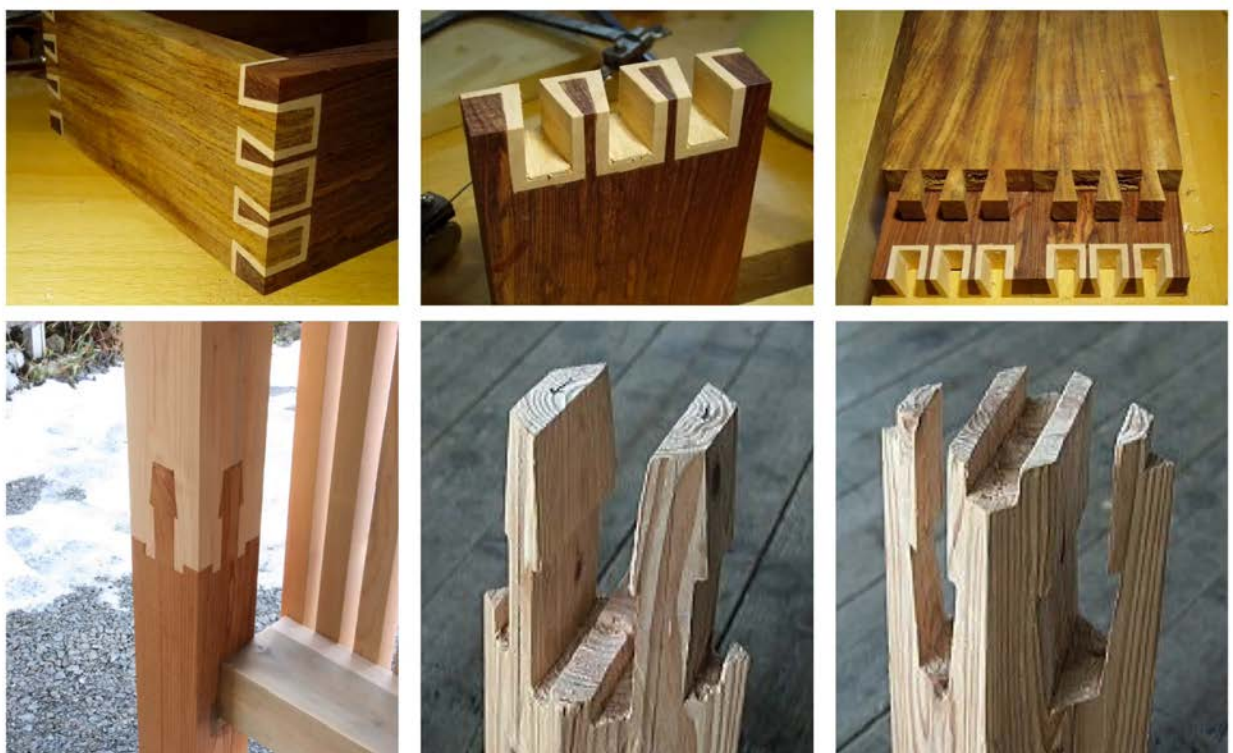



Source: [Beltramelli 2017]

# TEACHING MACHINES HOW TO GENERATE OTHER FORMS OF HUMAN LANGUAGE

# UNDERSTANDING DESIGN

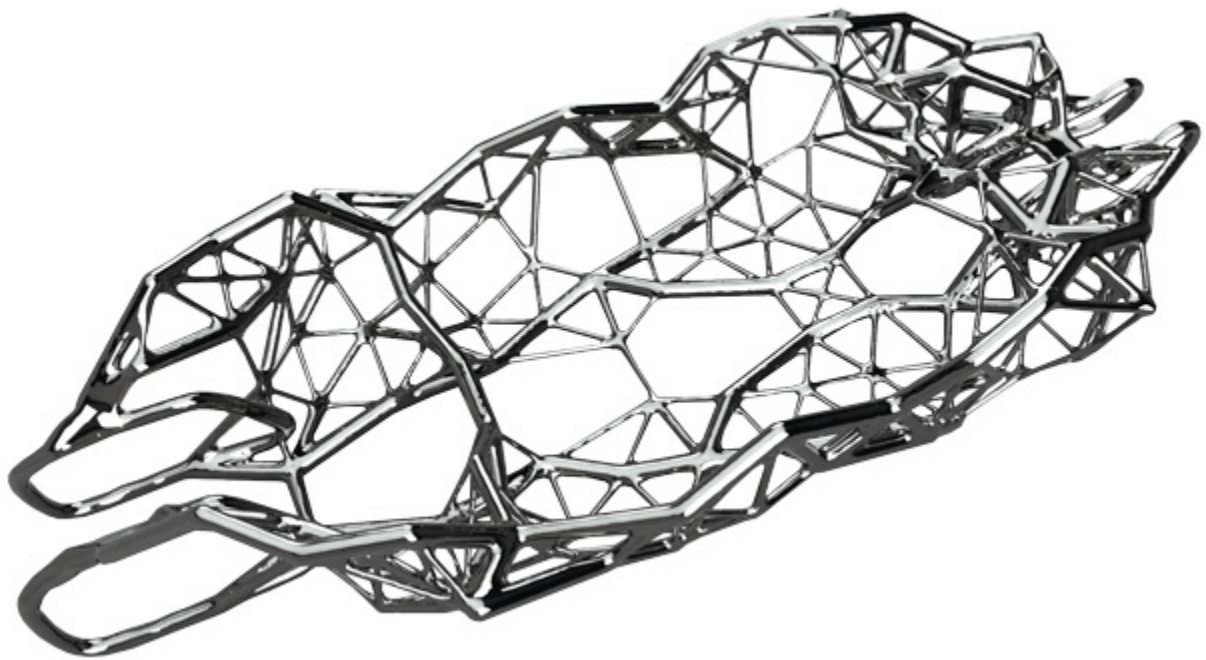
## 2D drawings to 3D sketches



Yao, Jiaxian, et al. "Interactive Design and Stability Analysis of Decorative Joinery for Furniture." *ACM Transactions on Graphics (TOG)* 36.2 (2017): 20.

# UNDERSTANDING DESIGN

## Models from specification



Autodesk research

# TEACHING MACHINES HOW TO SPEAK

# UNDERSTANDING SPEECH

## The presence of speech generation

http://deepzen.io/

# TEACHINGS MACHINES HOW TO SING

# UNDERSTANDING BEAUTY

## Performing music

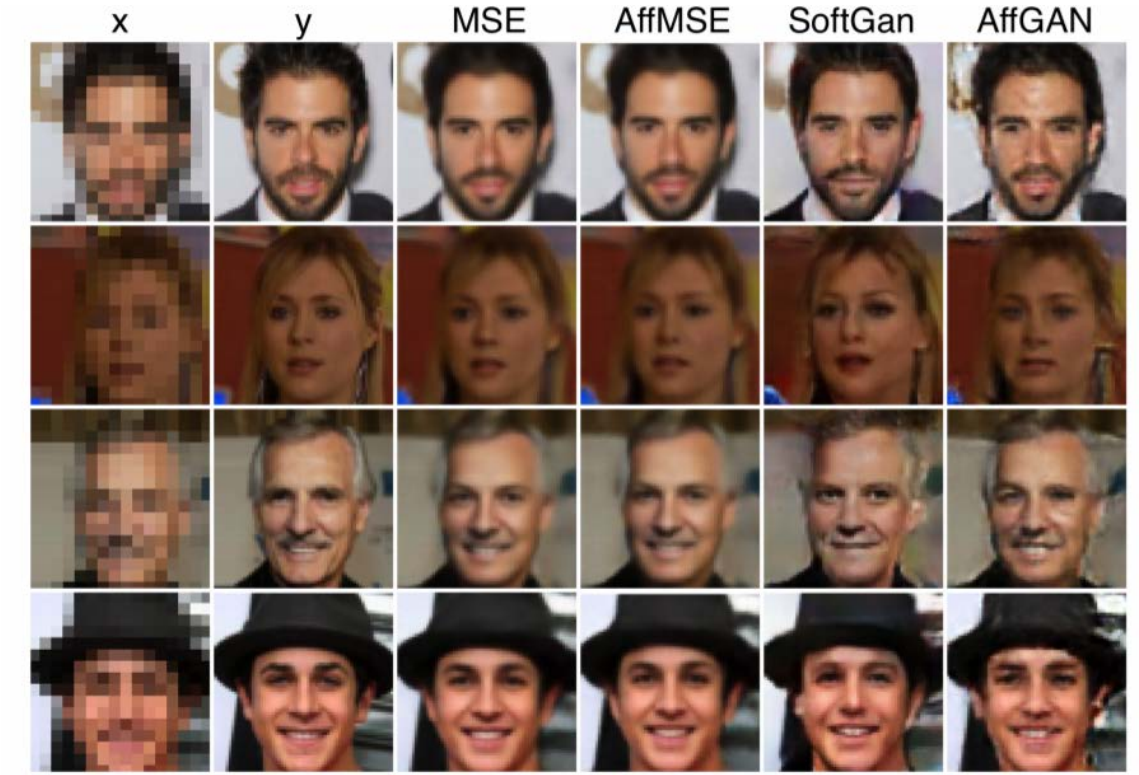
# UNDERSTANDING BEAUTY

## Performing music

# TEACHINGS MACHINES ABOUT HUMANS

# UNDERSTANDING IMAGE

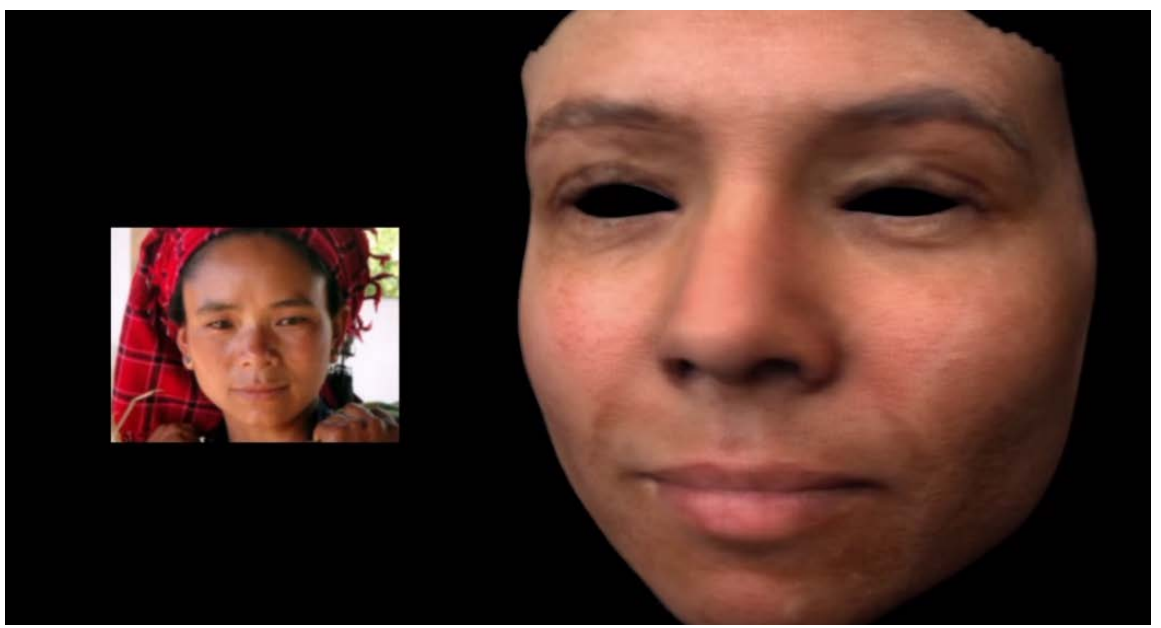
## Neural Network is computing most likely value of pixels




Sønderby, Casper Kaae, et al. "Amortised map inference for image super-resolution." *arXiv preprint arXiv:1610.04490* (2016).

# UNDERSTANDING THE SHAPE OF A FACE

## From a single image





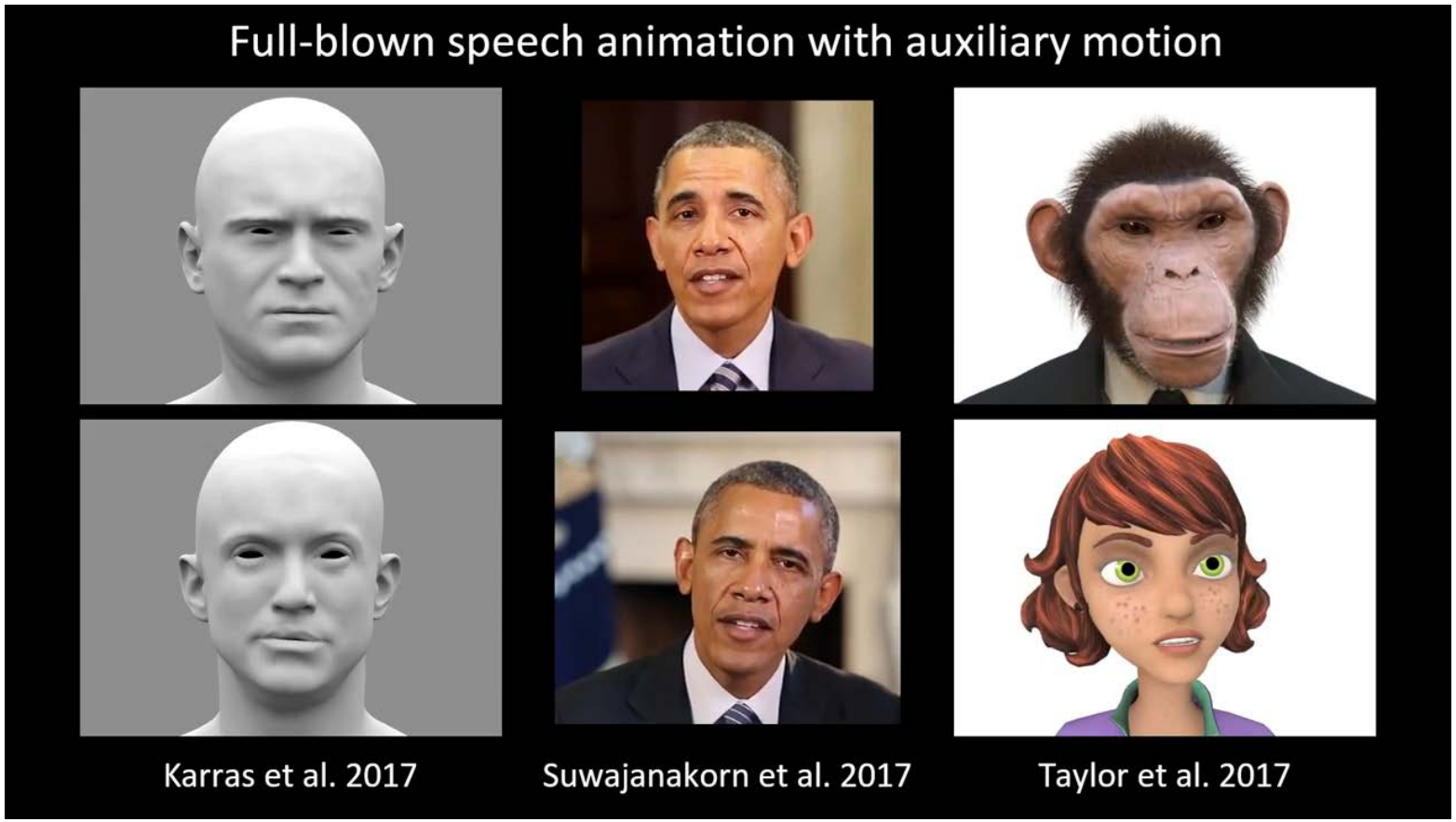
Saito, Shunsuke, et al. "Photorealistic Facial Texture Inference Using Deep Neural Networks." arXiv preprint arXiv:1612.00523 (2016).

# UNDERSTANDING A CONCEPT OF A FACE

## Generation

# ANIMATING A FACE

## And more




Karras, Tero, et al. "Audio-driven facial animation by joint end-to-end learning of pose and emotion." *ACM Transactions on Graphics (TOG)* 36.4 (2017): 94.

# ANIMATING A FACE

## Teaching the network to animate facial expressions

# TEACHING MACHINES ABOUT MOTION

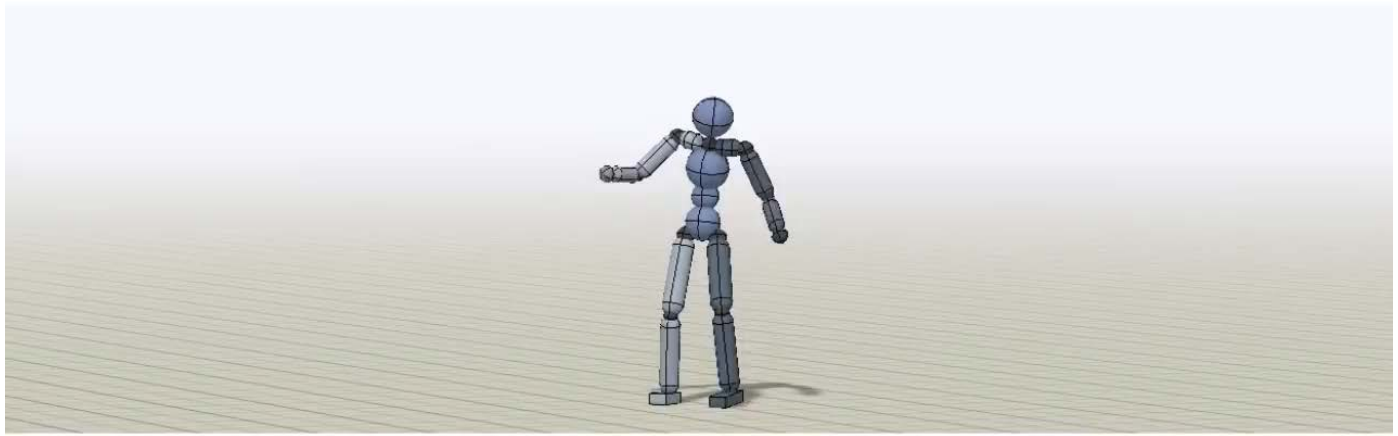
# CHARACTER ANIMATION

## Teaching the network to animate movement

# CHARACTER ANIMATION

## Teaching the network to animate movement

DeepMimic: Example-Guided Deep Reinforcement Learning of Physics-Based Character Skills



Xue Bin Peng[1], Pieter Abbeel[1], Sergey Levine[1], Michiel van de Panne[2]

[1] University of California Berkeley

[2] University of British Columbia

# TEACHINGS MACHINES ABOUT THE VISUAL WORLD

# UNDERSTANDING IMAGE

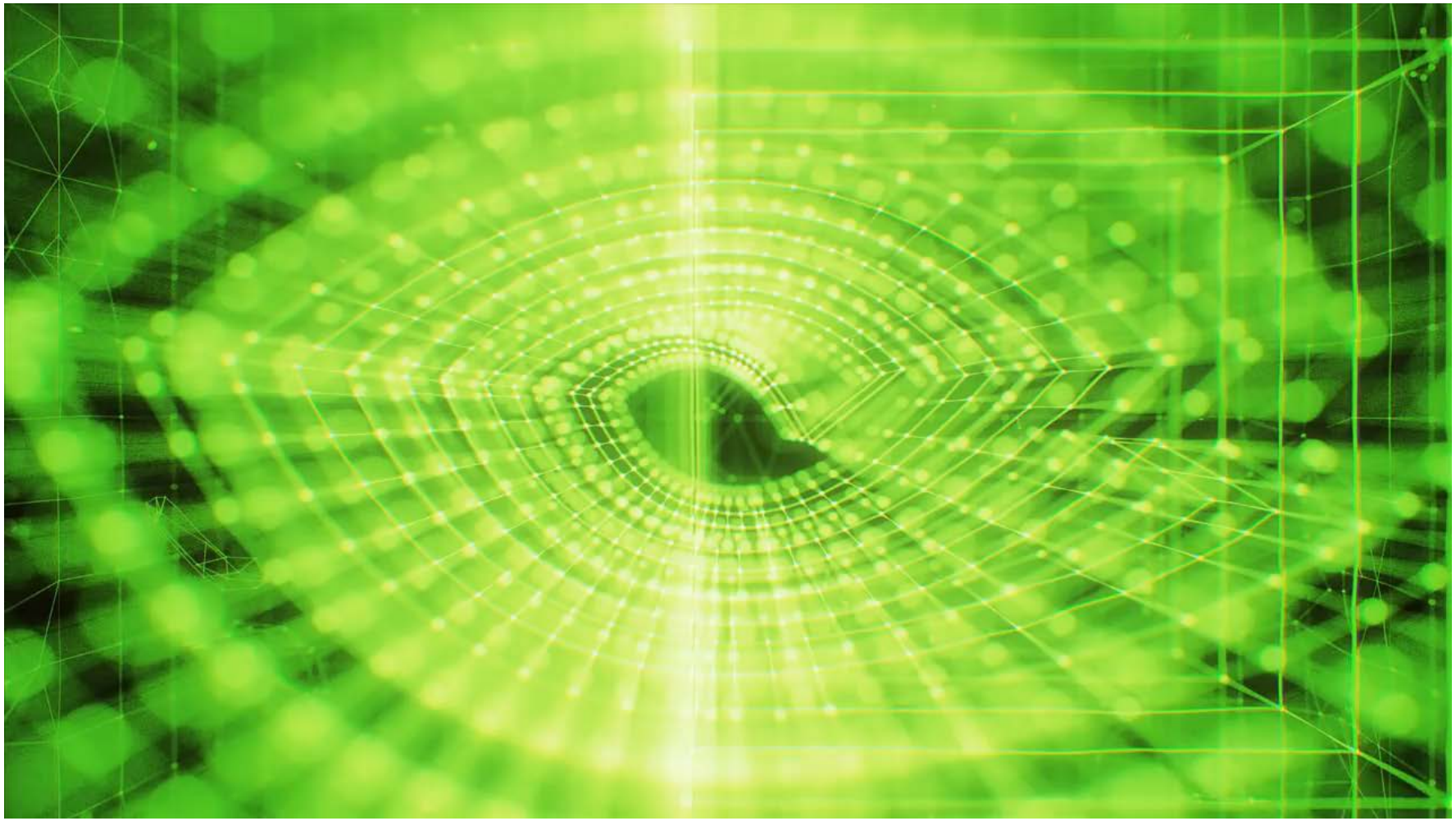
## Teaching the network the physics of the world

# UNDERSTANDING IMAGE QUALITY

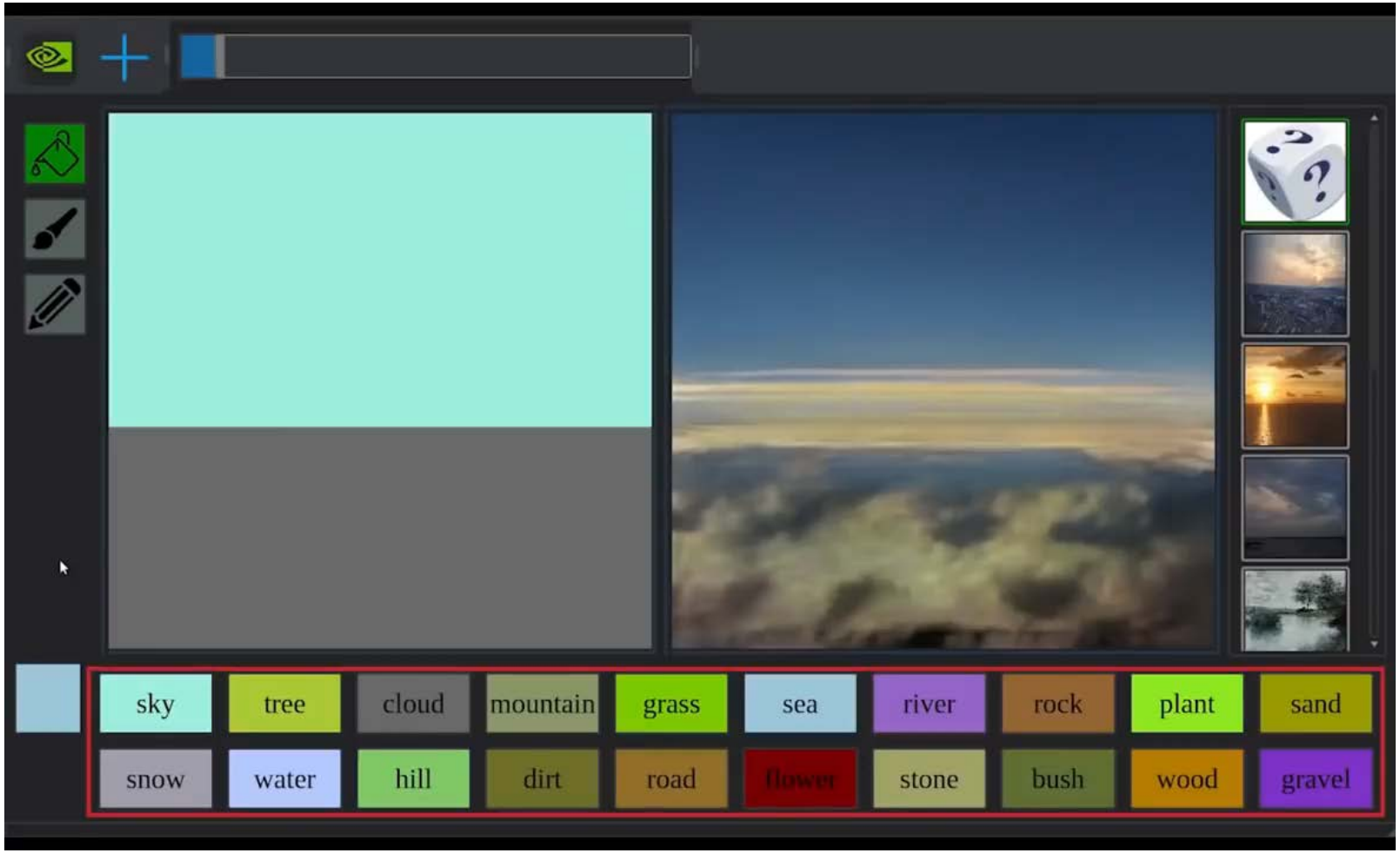


http://people.ee.ethz.ch/~ihnatova/

# UNDERSTANDING THE VISUAL PROPERTIES OF THE WORLD



# UNDERSTANDING THE VISUAL PROPERTIES OF THE WORLD



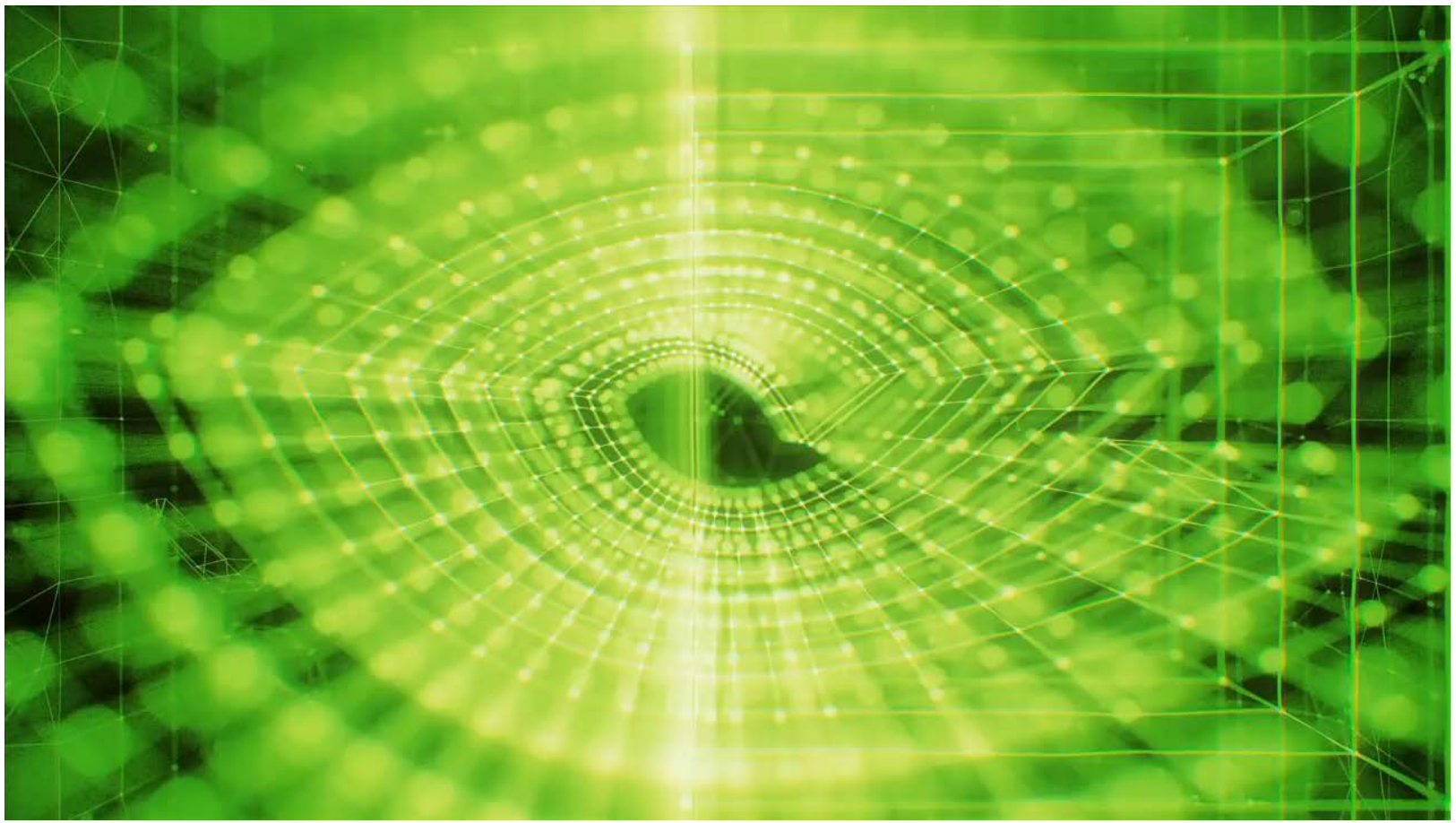
# UNDERSTANDING IMAGE GENERAL PROPERTIES

How should the image look like from a different angle

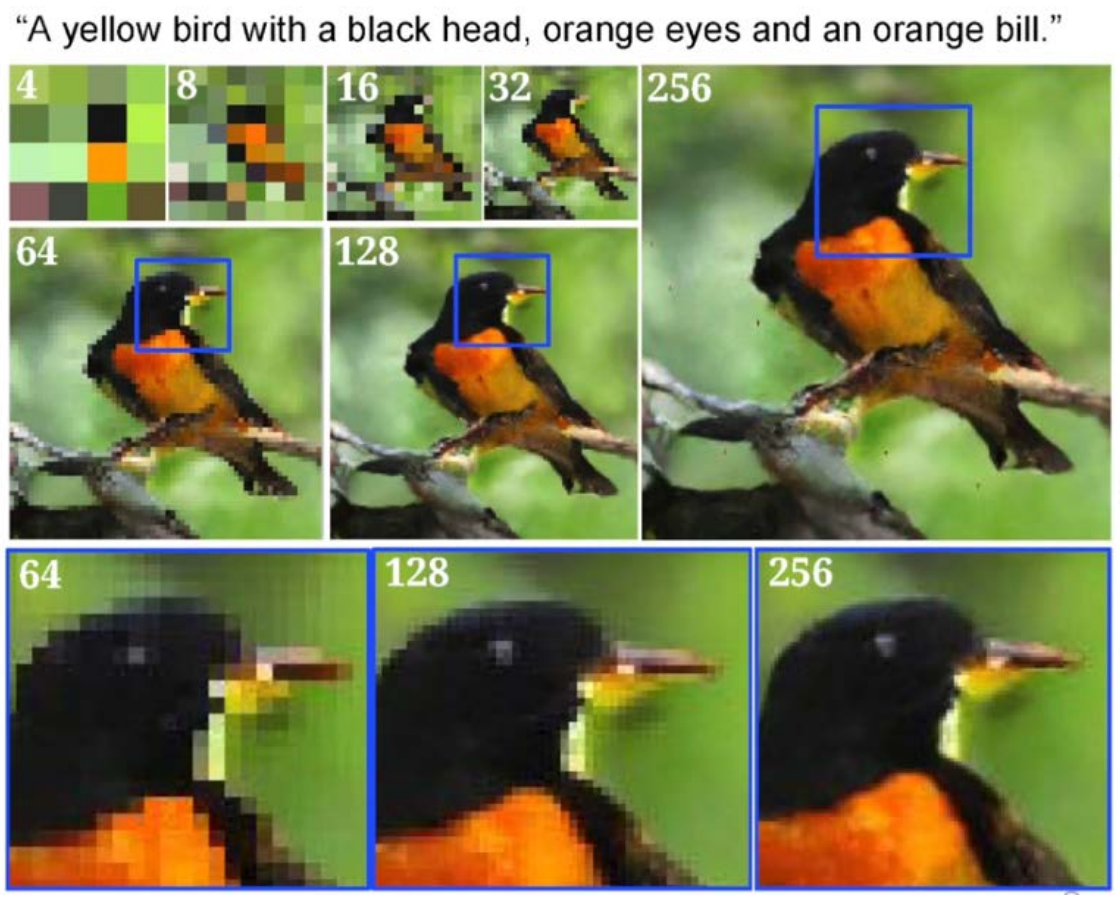
# TEACHING MACHINES ABOUT STYLE AND BEAUTY

# SEMANTIC STYLE TRANSFER

## Understanding common semantic part of the image



# GENERATING IMAGES FROM DESCRIPTION





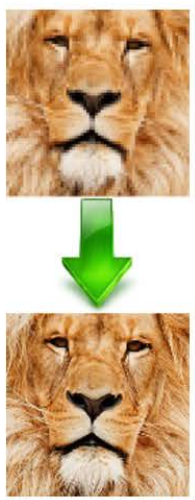
Reed, Scott, et al. "Parallel Multiscale Autoregressive Density Estimation." arXiv preprint arXiv:1703.03664 (2017).

# NOT JUST THEORETICAL RESEARCH

## Nvidia GameWorks

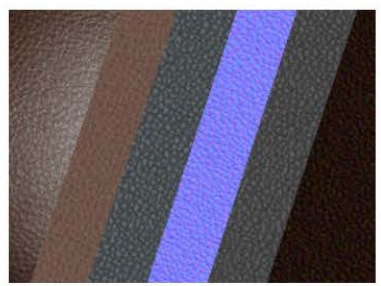
### Super-Resolution

Allows the user to get a 2x, 4x or 8x increase in resolution of textures and photos. The user can also toggle "photorealistic hallucination" which uses a novel deep learning technique to infer increased detail during upscaling.



### Photo To Material: 2shot

Using two photos of a real-world surface, generates diffuse, normals, specular and gloss maps.



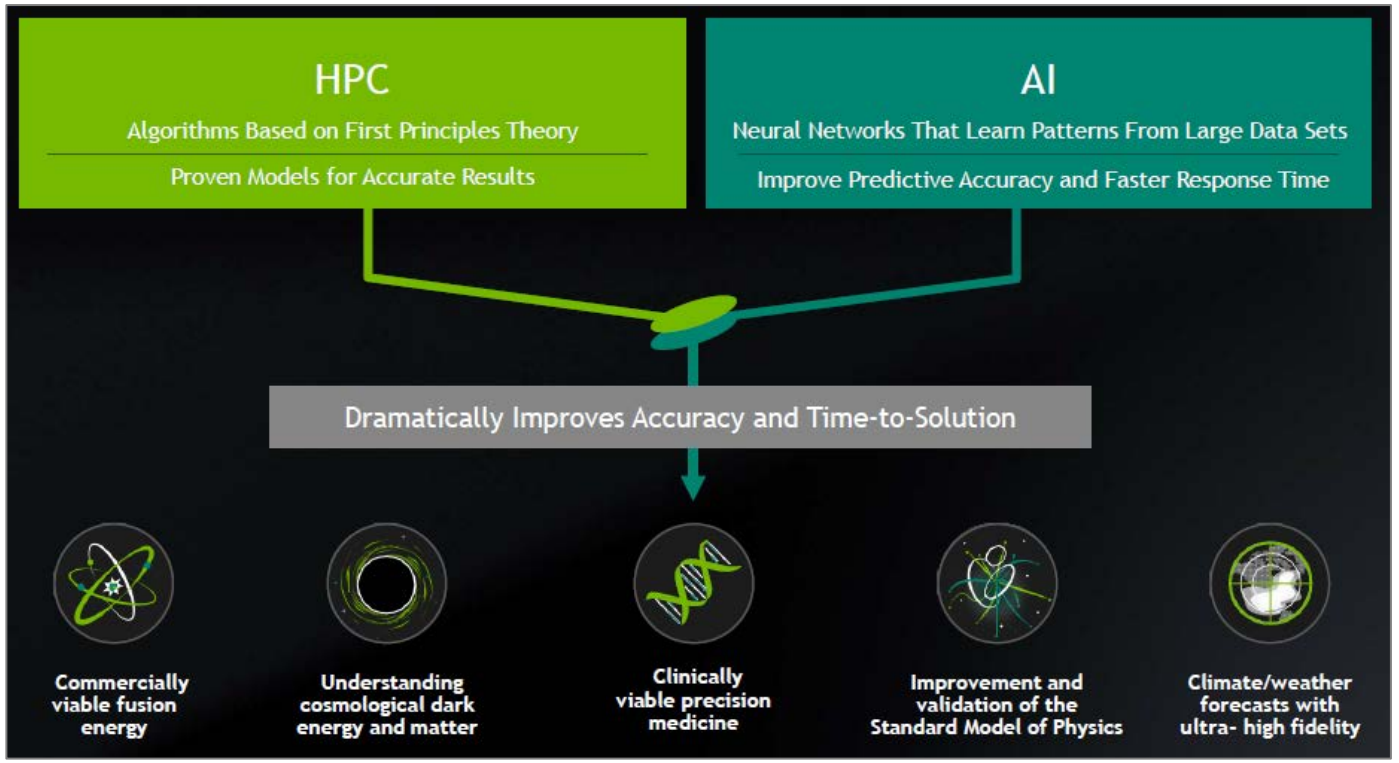
### Texture Multiplier

Takes a texture or photo as input and provides the user with an organic variation.

# TEACHINGS MACHINES ABOUT THE PHYSICAL WORLD

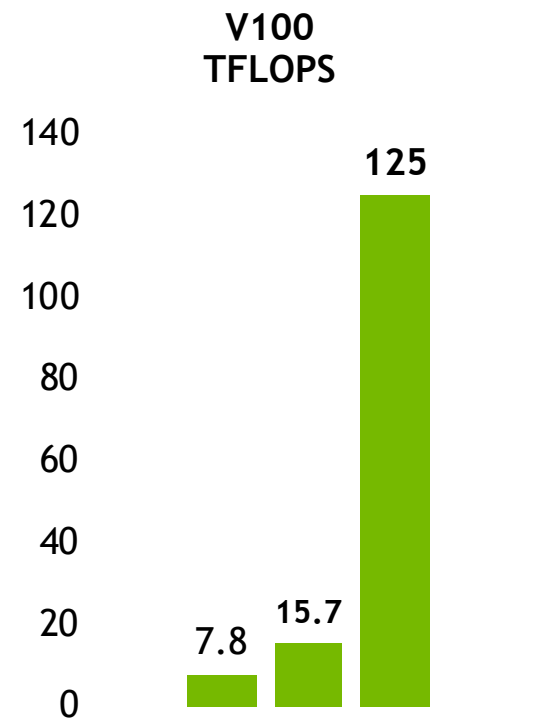
# AI – THE NEW INSTRUMENT FOR SCIENCE

## Need for general purpose compute

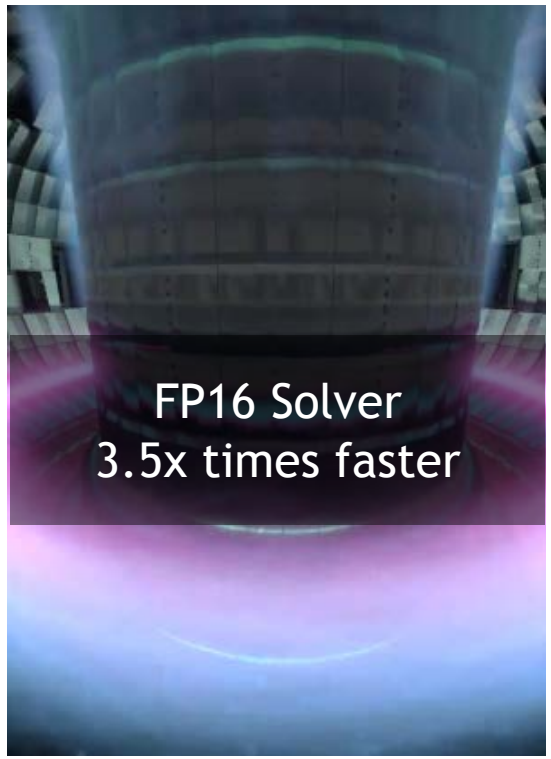


# NEURAL NETWORKS FOR SCIENCE

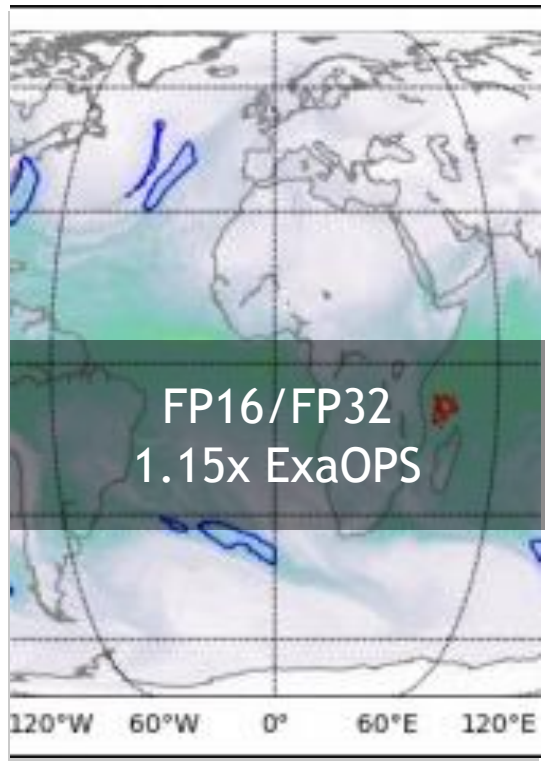
## Multi-precision computing



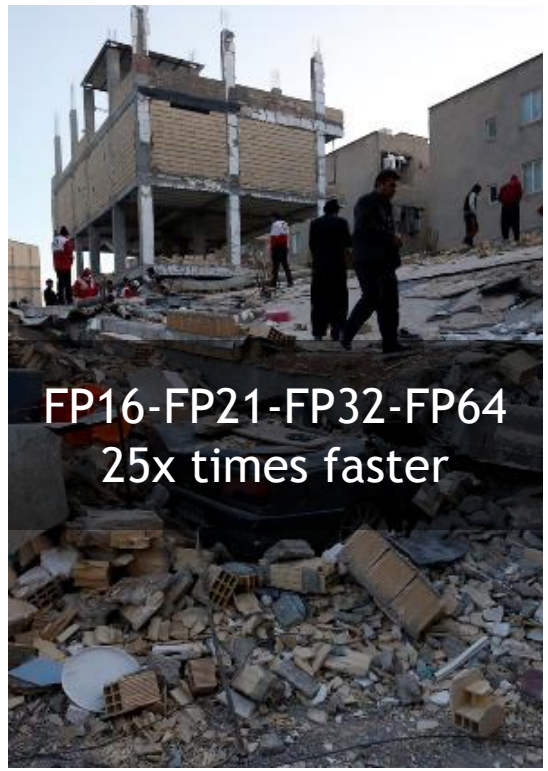

FP64+ MULTI-PRECISION




PLASMA FUSION APPLICATION




AI-POWERED WEATHER PREDICTION




EARTHQUAKE SIMULATION

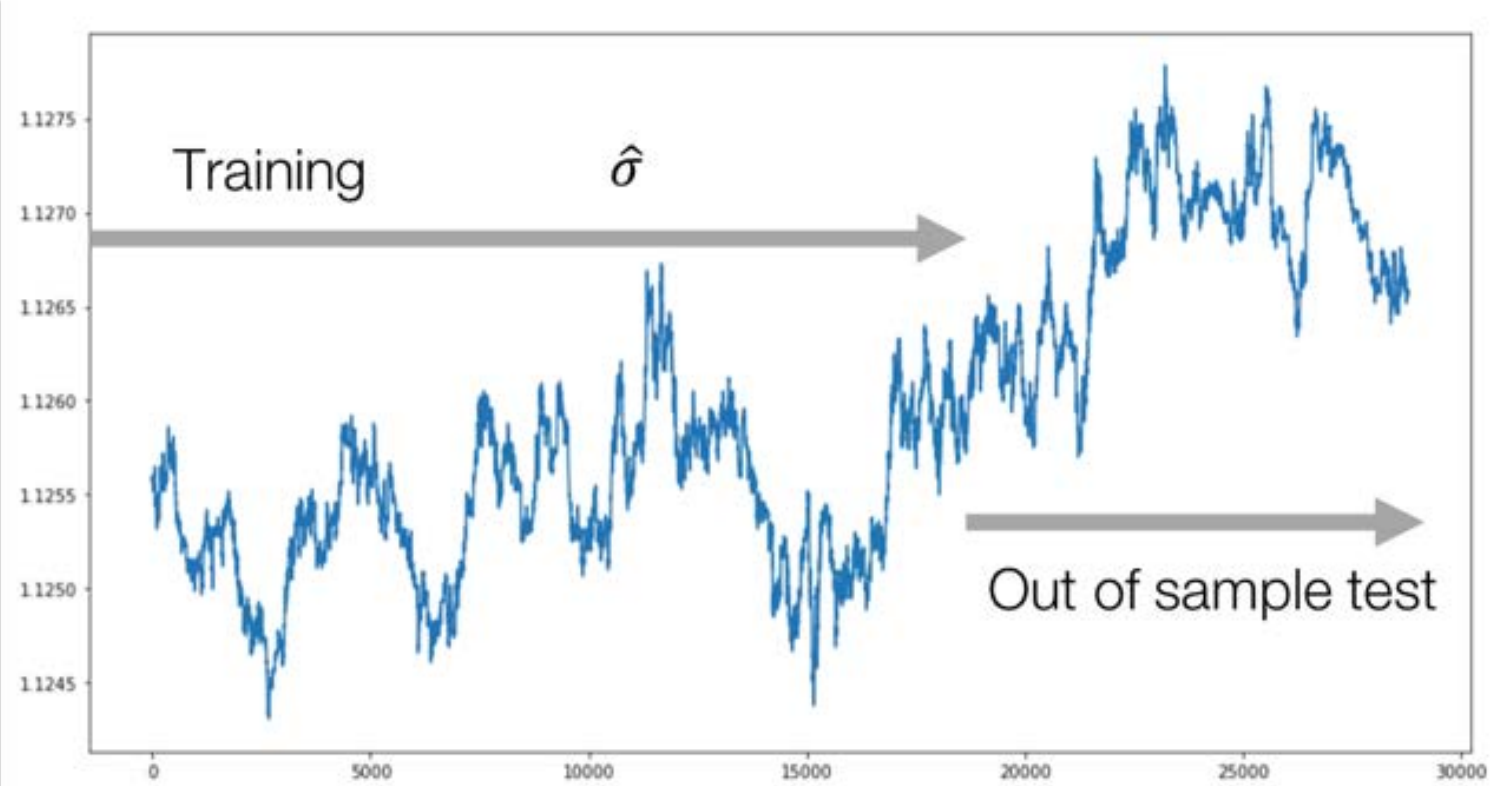
# AI – THE NEW INSTRUMENT FOR SCIENCE

## Fluid dynamics

# TEACHINGS MACHINES ABOUT ECONOMY

# UNDERSTANDING ECONOMICS

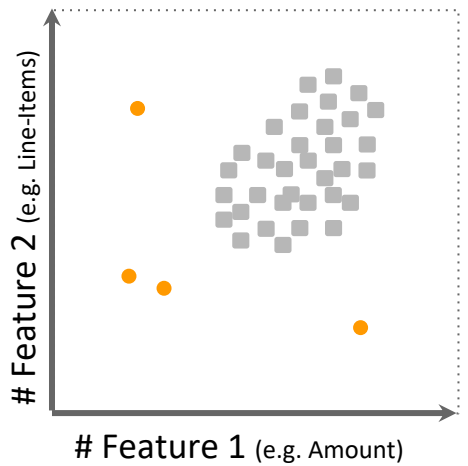
## Generating stock movement



# Understanding Business

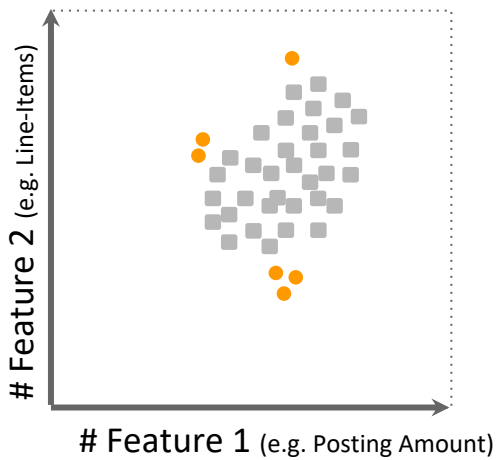
## Differentiating mistakes from fraud

### „Global“ Accounting Anomalies




- Usually postings or documents that exhibit an unusual or **rare attribute values**, such as:
  - Seldom used user accounts,
  - Reverse postings, corrections

### „Local“ Accounting Anomalies



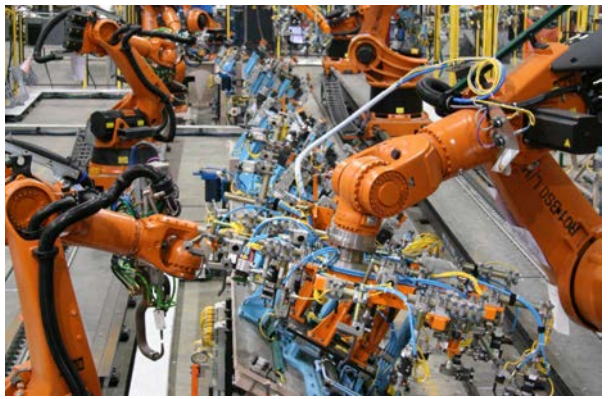

- Usually postings or documents that exhibit an unusual or **rare attribute combination**, for example:
  - Unusual posting activities
  - Deviating user behavior

# TEACHINGS MACHINES ABOUT MACHINES

# AI FOR INDUSTRIAL APPLICATIONS

Wide range of applications

## FACTORY INSPECTION



Quality Inspection
Fault Detection & Classification
Inventory Inspection

## FIELD INSPECTION

Condition Based Maintenance
Remaining Useful Life

## PREDICTIVE MAINTENANCE

Sensor Time Series Analysis
Failure Prediction