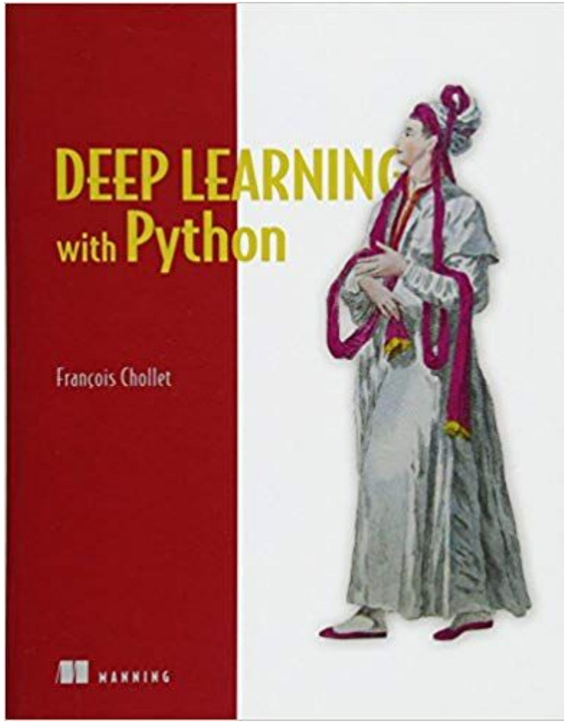
Deep learning:

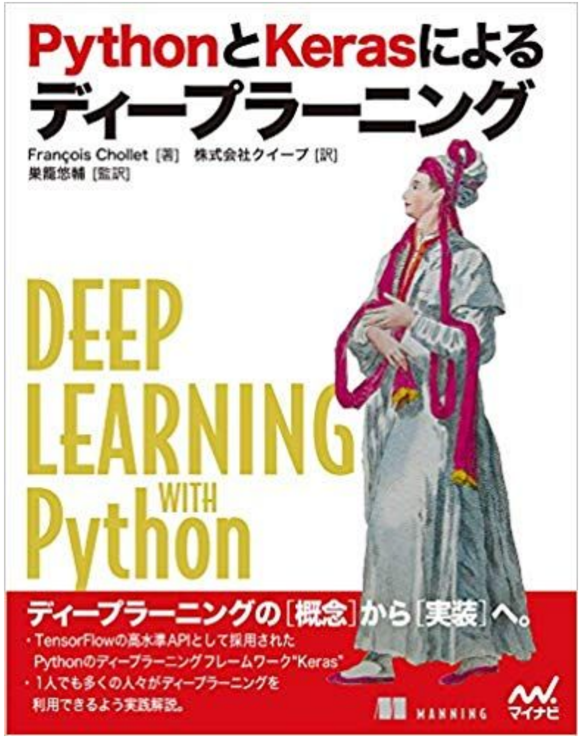
current capabilities, limitations, and future perspectives

François Chollet
@fchollet

Research at Google

Keras

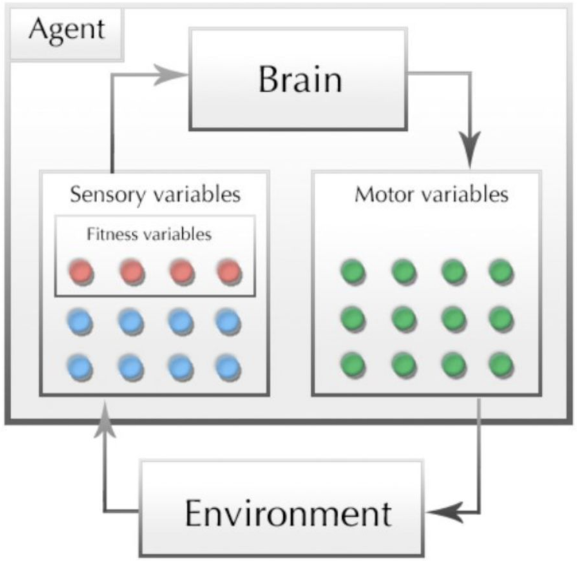

DEEP LEARNING
with Python
François Chollet
MANNING


PythonとKerasによる
ディープラーニング
François Chollet [著] 株式会社クイープ [訳]
巣籠悠輔 [監訳]
DEEP LEARNING WITH Python
ディープラーニングの[概念]から[実装]へ。
・TensorFlowの高水準APIとして採用された
Pythonのディープラーニングフレームワーク"Keras"
・1人でも多くの人々がディープラーニングを
利用できるよう実践解説。
MANNING
マイナビ

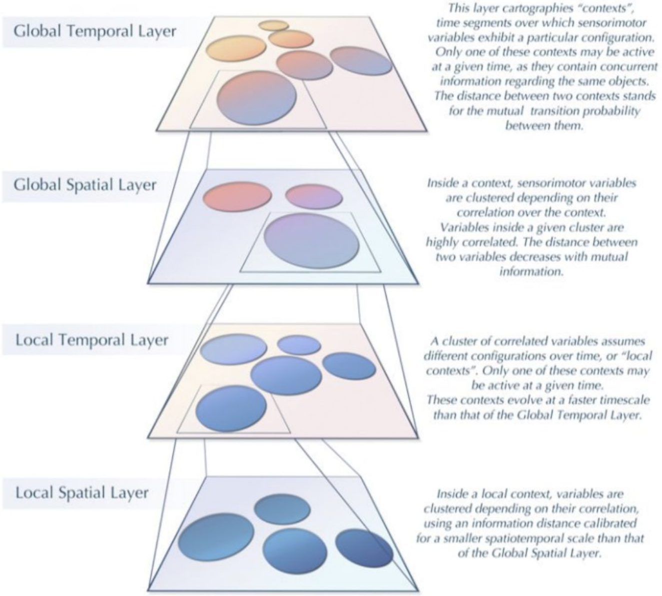

Strong AI?
2019: mostly deep learning

I’ve been thinking about human-level AI for a long time (e.g. what I was doing at Todai in 2012)

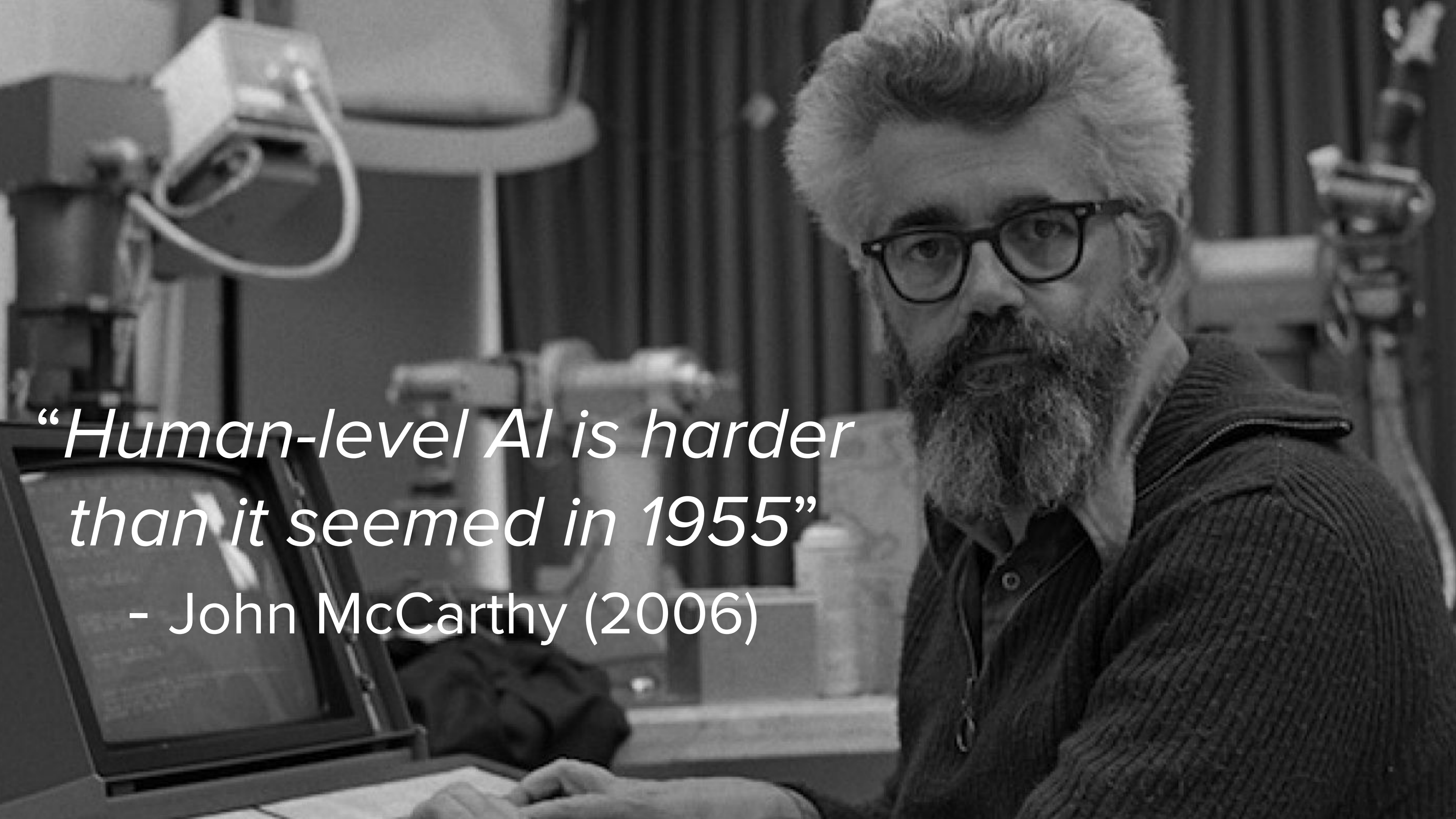
Position of the problem
Brain function?



The multiscale, massively modular structure
of the space of sensorimotor information
as described in the [redacted] architecture

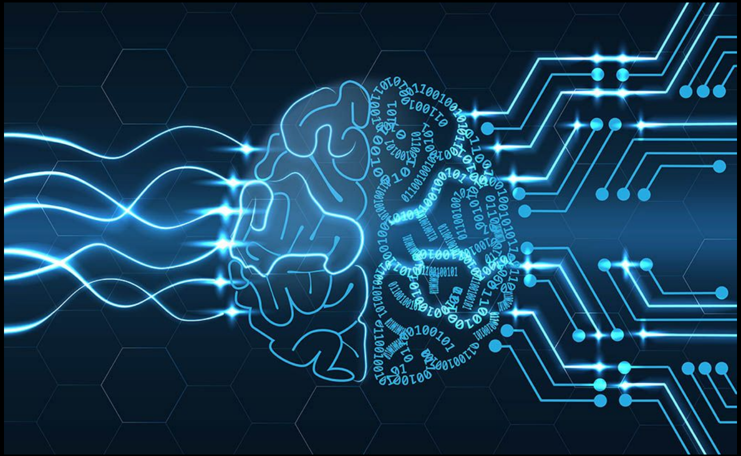
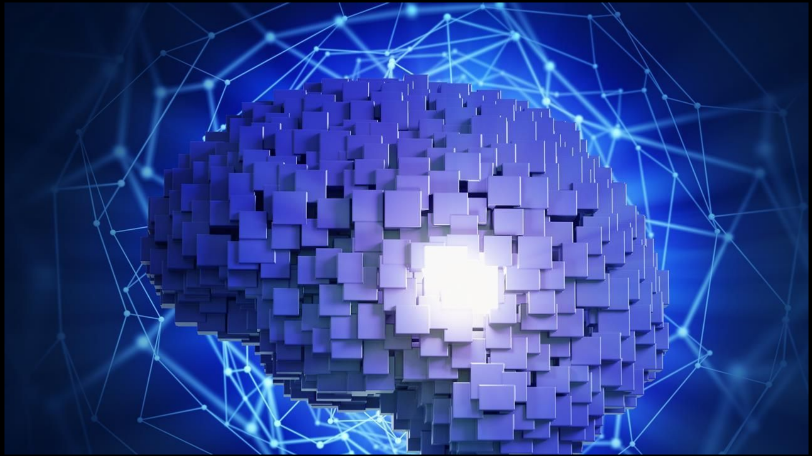


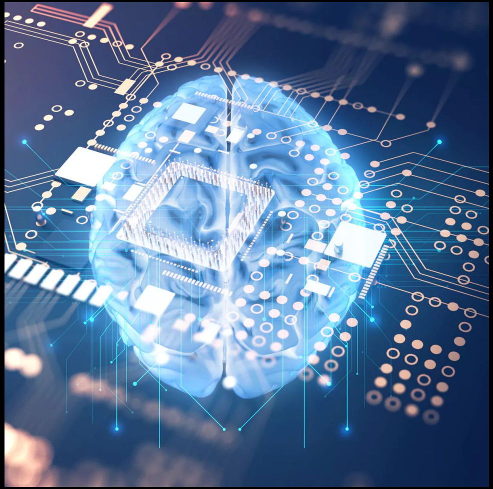
Human-level AI is harder
than it seemed in 2010


“Human-level AI is harder than it seemed in 1955”
- John McCarthy (2006)

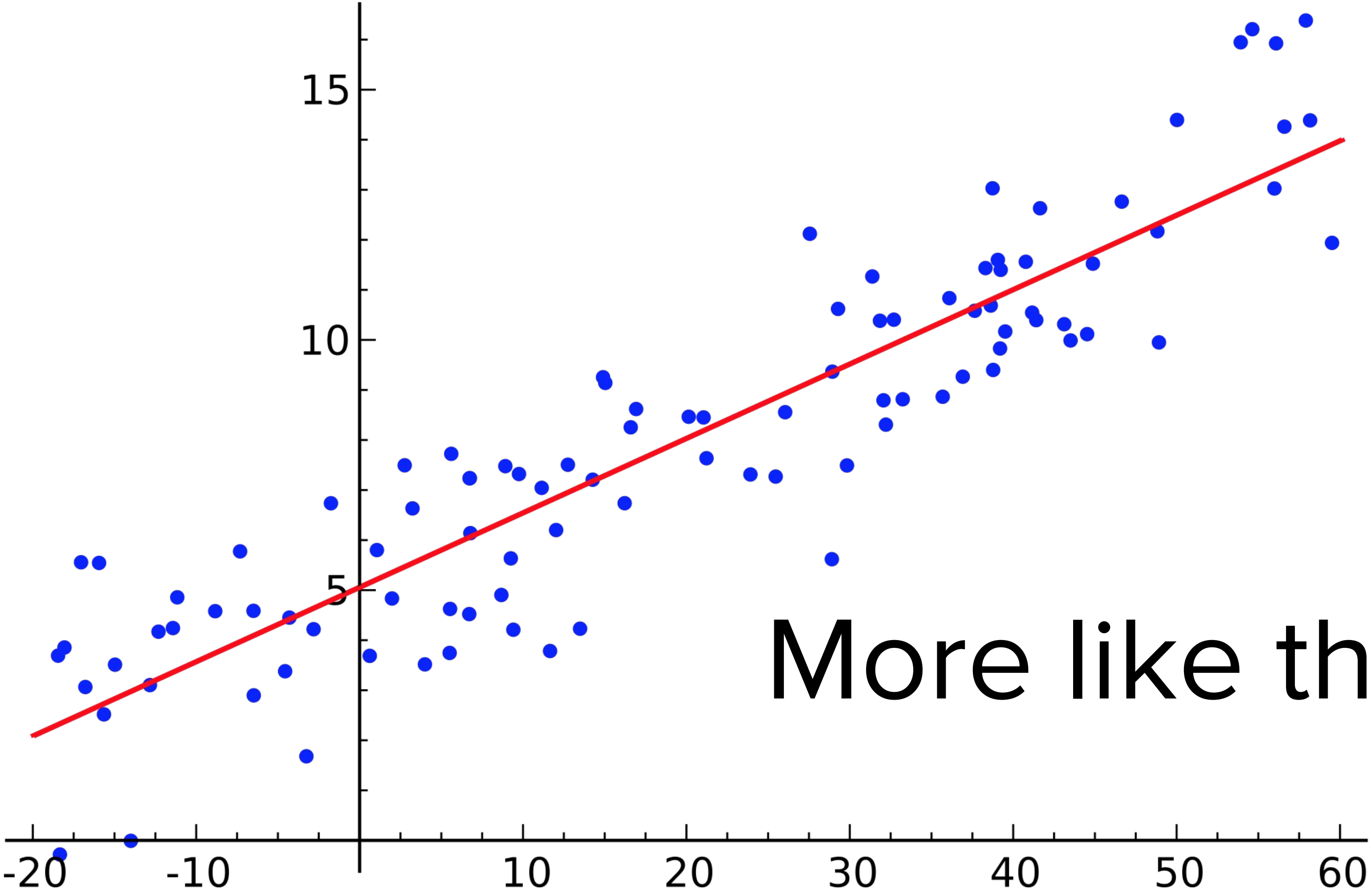

Strong AI?
2019: mostly deep learning

What are we talking about when we talk about deep learning?

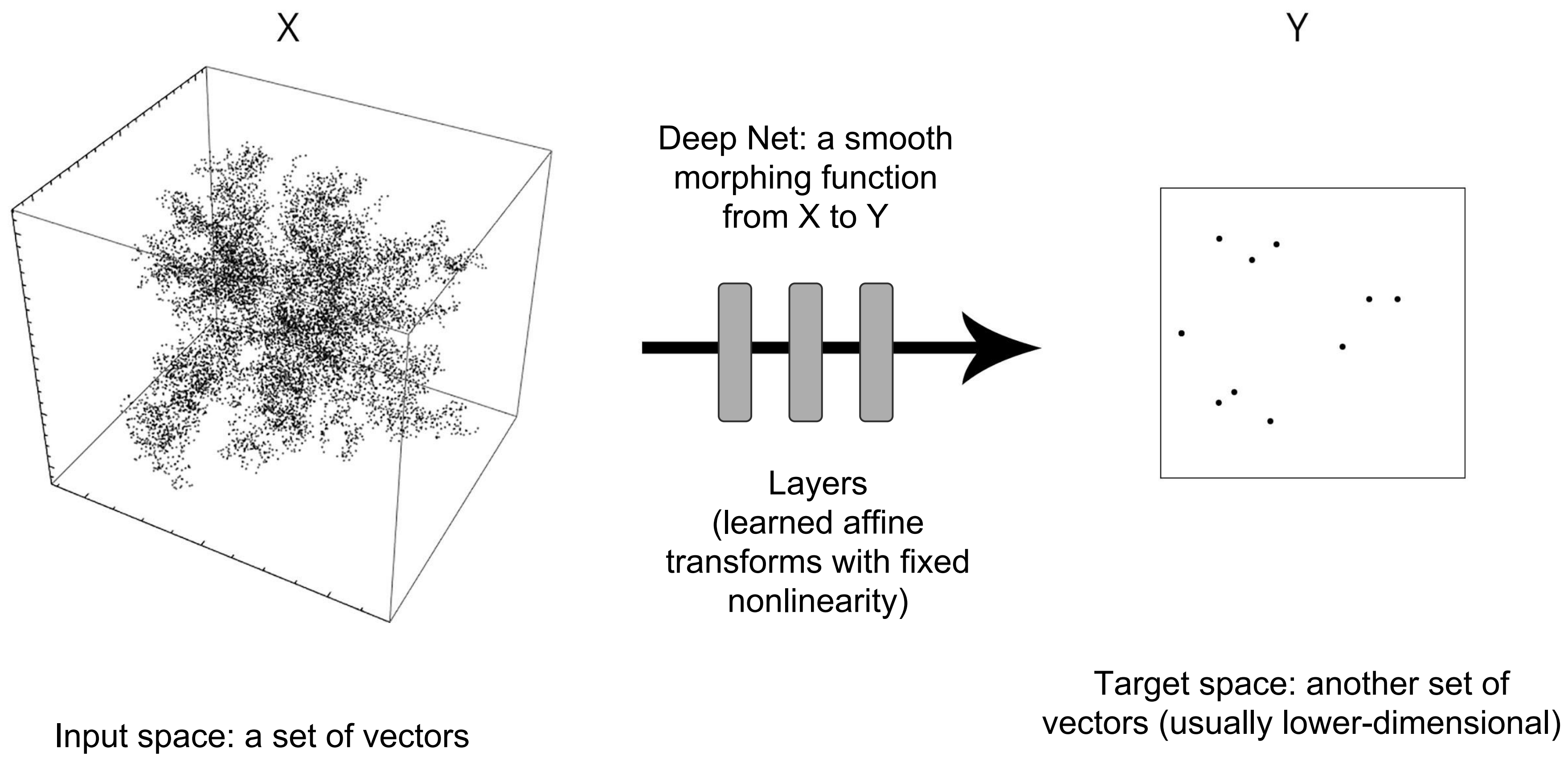


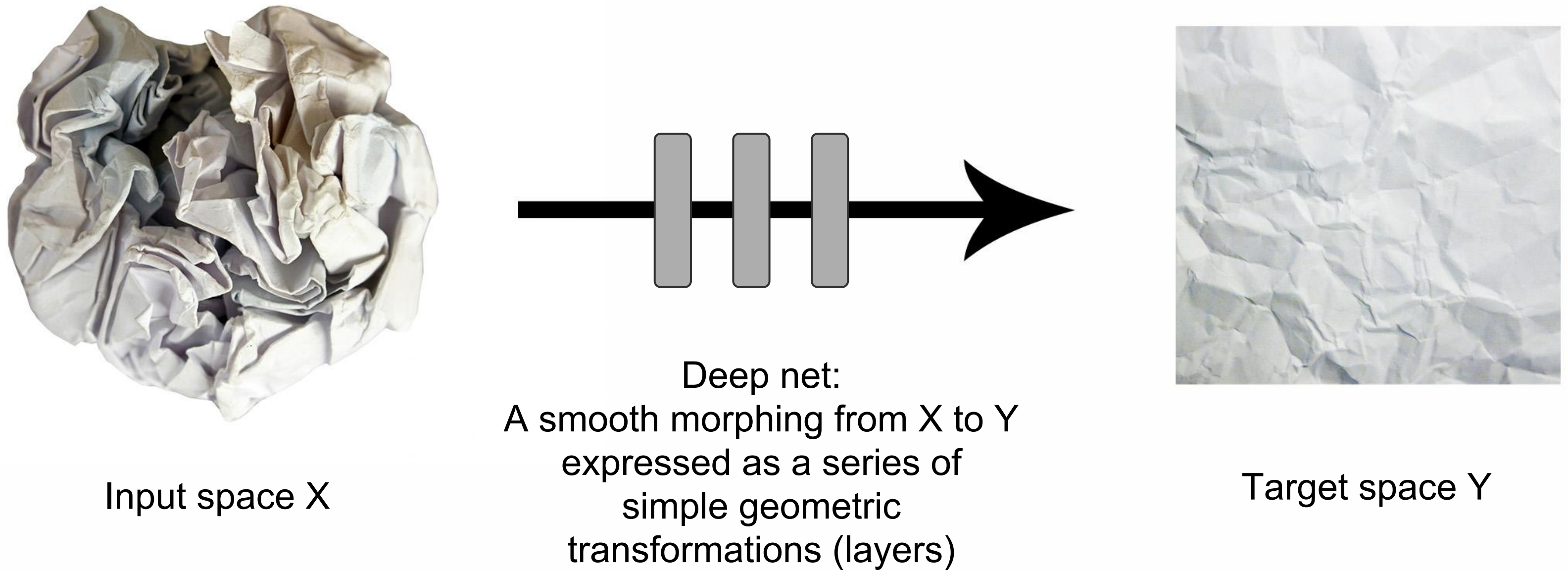
Not this



More like this

The most surprising thing about deep learning is how **simple** it is


X
Y
Deep Net: a smooth morphing function from X to Y
Layers (learned affine transforms with fixed nonlinearity)
Input space: a set of vectors
Target space: another set of vectors (usually lower-dimensional)

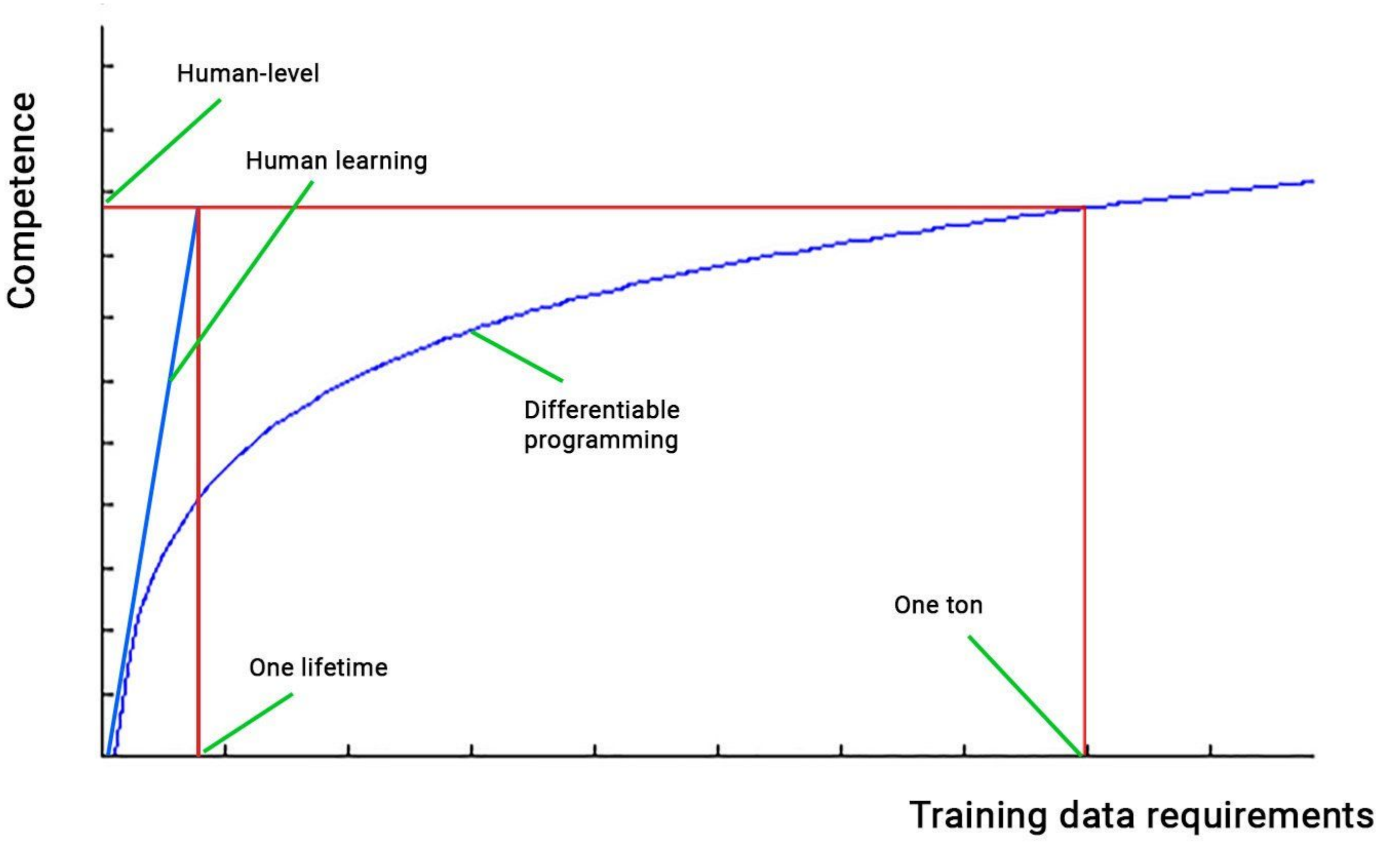

Input space X
Deep net:
A smooth morphing from X to Y
expressed as a series of
simple geometric
transformations (layers)
Target space Y

a DENSE SAMPLING of the
`input X output` space

With “enough data”, such a mapping fitted with SGD can solve arbitrarily complex problems

- Mapping pixels to labels
- Mapping soundwaves to speech transcripts
- Mapping situation to action
- Anything to anything, really

So it doesn’t scale very well



The limitations of deep learning

Extreme sensitivity to **adversarial perturbations**
Extreme sensitivity to **any input change** not seen in the training data
It can only make sense of **what it has seen before**

Measuring the tendency of CNNs to Learn Surface Statistical Regularities

Jason Jo
MILA, Université de Montréal
IVADO
jason.jo.research@gmail.com

Yoshua Bengio
MILA, Université de Montréal
CIFAR
yoshua.umontreal@gmail.com

Why do deep convolutional networks generalize so poorly to small image transformations?

Aharon Azulay
ELSC
Hebrew University of Jerusalem
aharon.azulay@mail.huji.ac.il

Yair Weiss
CSE, ELSC
Hebrew University of Jerusalem
yweiss@cs.huji.ac.il

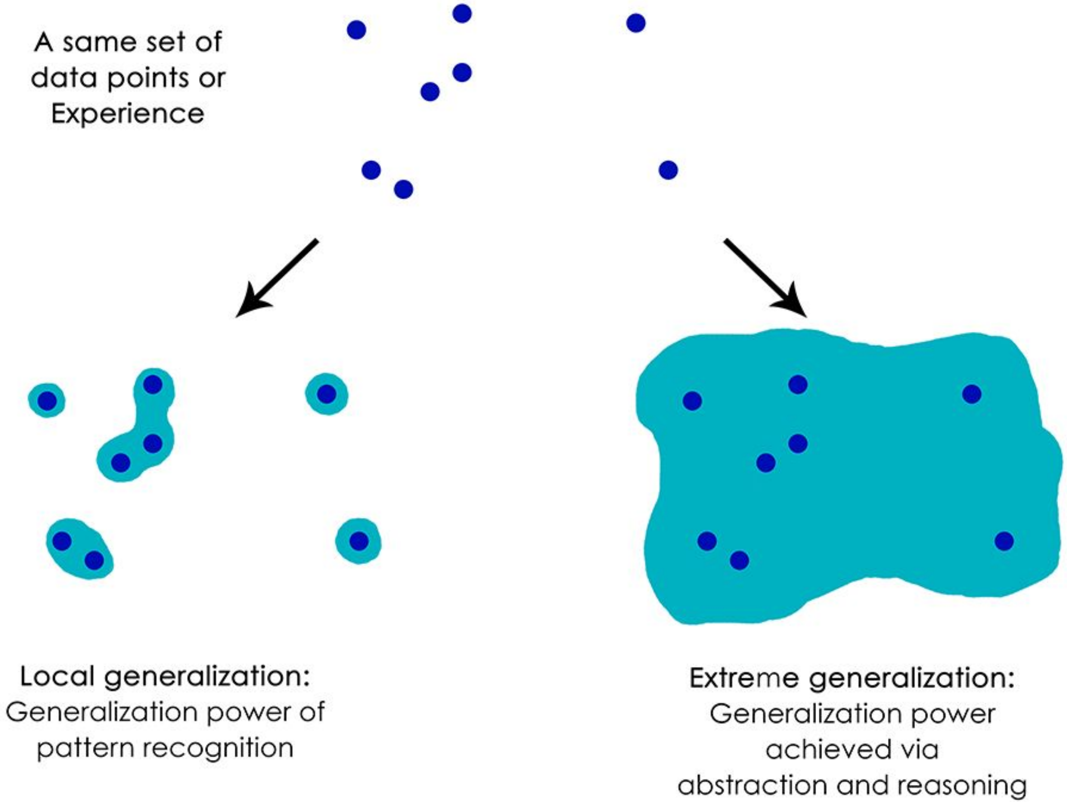
Don’t anthropomorphize deep learning models

Source:
Samim Winiger

Deep learning is **pattern recognition**

Models **memorize** a data manifold and
perform **local generalization** around training examples

Local generalization vs. extreme generalization



Task: finding launch parameters for a moon rocket

Deep learning:

- Hard-code space of launch parameters
- Learn point-by-point mapping between launch parameters and rocket outcome
- Launch millions of times

Humans:

- Develop abstract model of the problem (rocket science)
- Adapt it empirically (few launches)

Task: avoiding getting hit by a car

Deep learning:

- Learn point-by-point mapping between sensorimotor space and vital outcome
- Die millions of times
- Need to re-learn most of it in a new environment

Humans:

- Learn from others (imitation, spoken instructions)
- Model the world in an abstract way (e.g. understand physics to predict collisions in a new context)
- Die ~0 times

What does it take to achieve **extreme generalization**?

Extreme generalization via strong abstraction:
the **next frontier** for AI

Better metrics.
Richer models.
Stronger priors.

We'll need **better metrics**

First step: an ambitious new benchmark to measure progress towards extreme generalization intelligence skill experience (data) priors

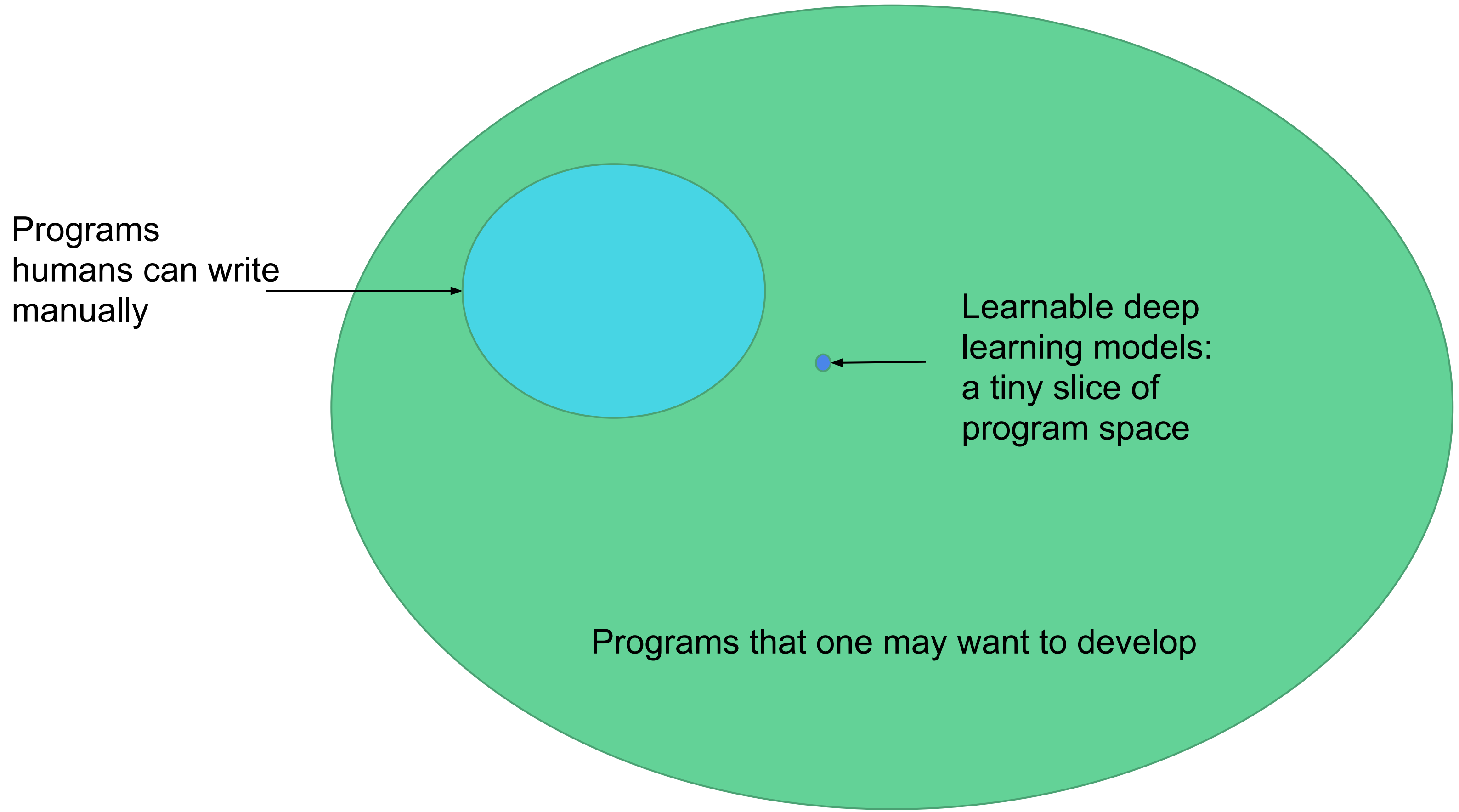
Posit: $I = S / (E + P)$

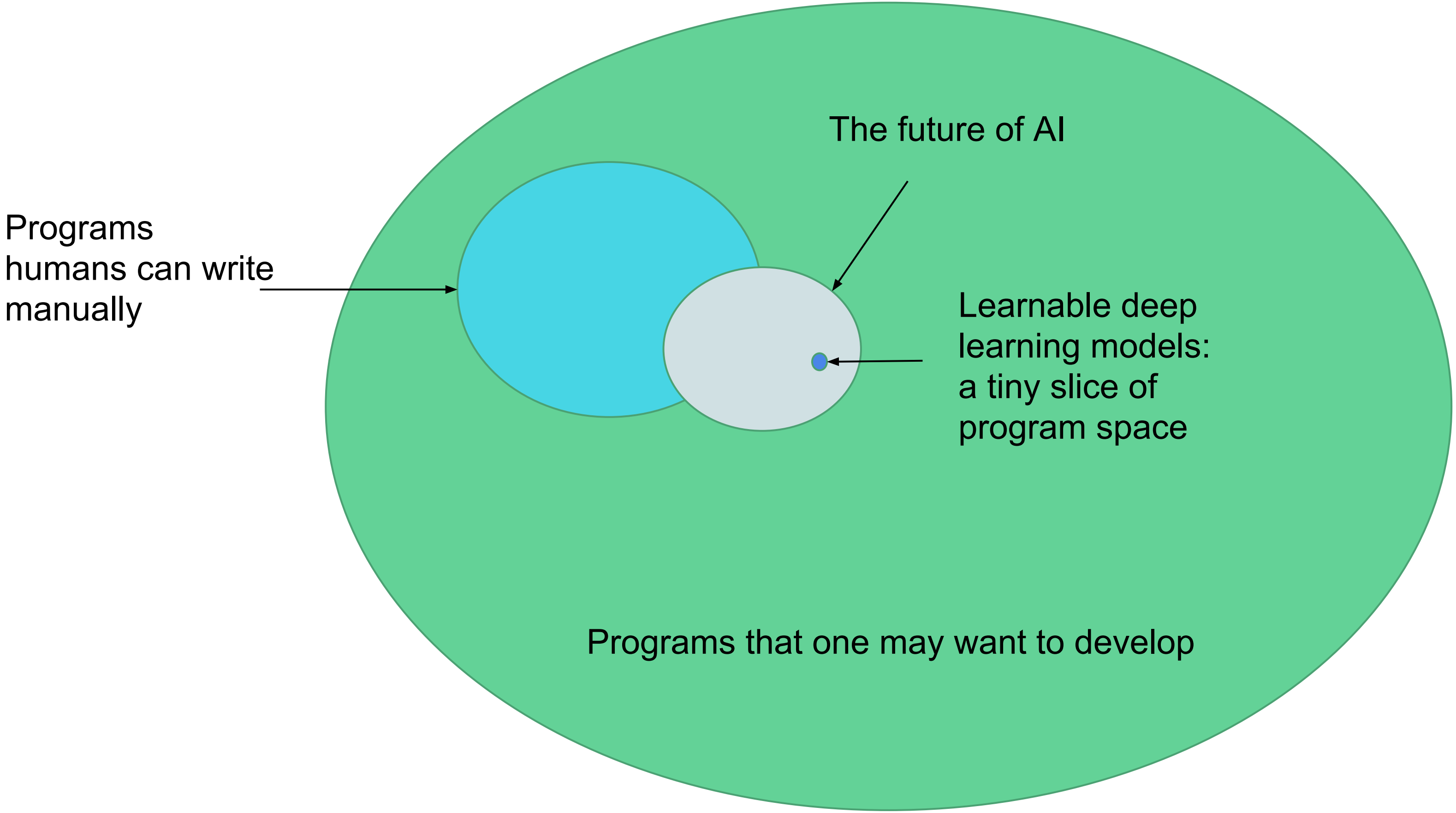
Measure I instead of S:

- Control for experience
- Control for priors

We'll need **richer models**

More like current **computer programs**,
less like paper ball uncrumpling


Programs
humans can write
manually
Learnable deep
learning models:
a tiny slice of
program space
Programs that one may want to develop


Programs
humans can write
manually
The future of AI
Learnable deep
learning models:
a tiny slice of
program space
Programs that one may want to develop

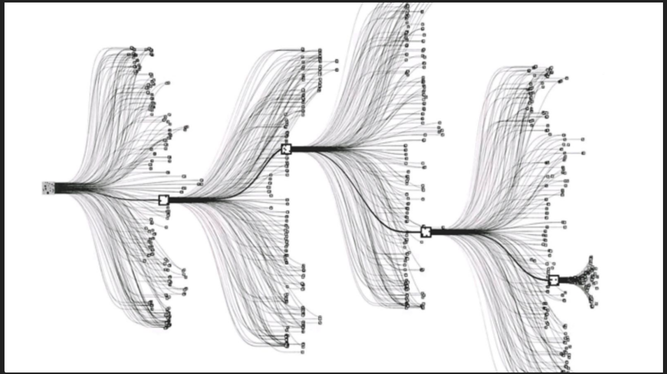
“Learning” in “Machine Learning” will be more **program synthesis** than tuning the parameters of a hard-coded geometric transform

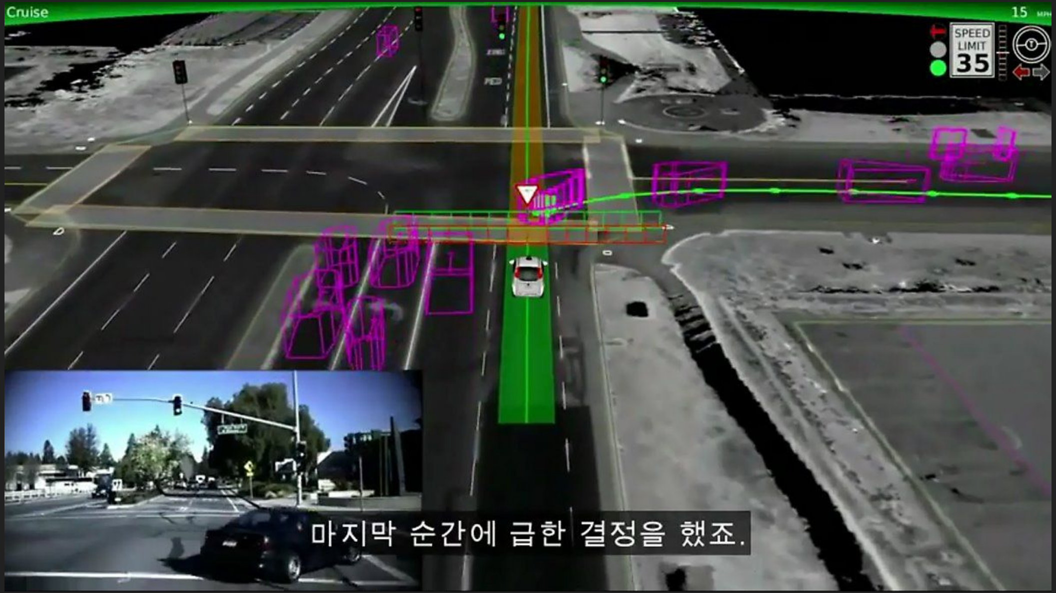
But we’re not going to
throw away deep learning

Future AI systems will blend
pattern recognition (geometric intelligence)
with **abstraction & reasoning**
(symbolic intelligence)

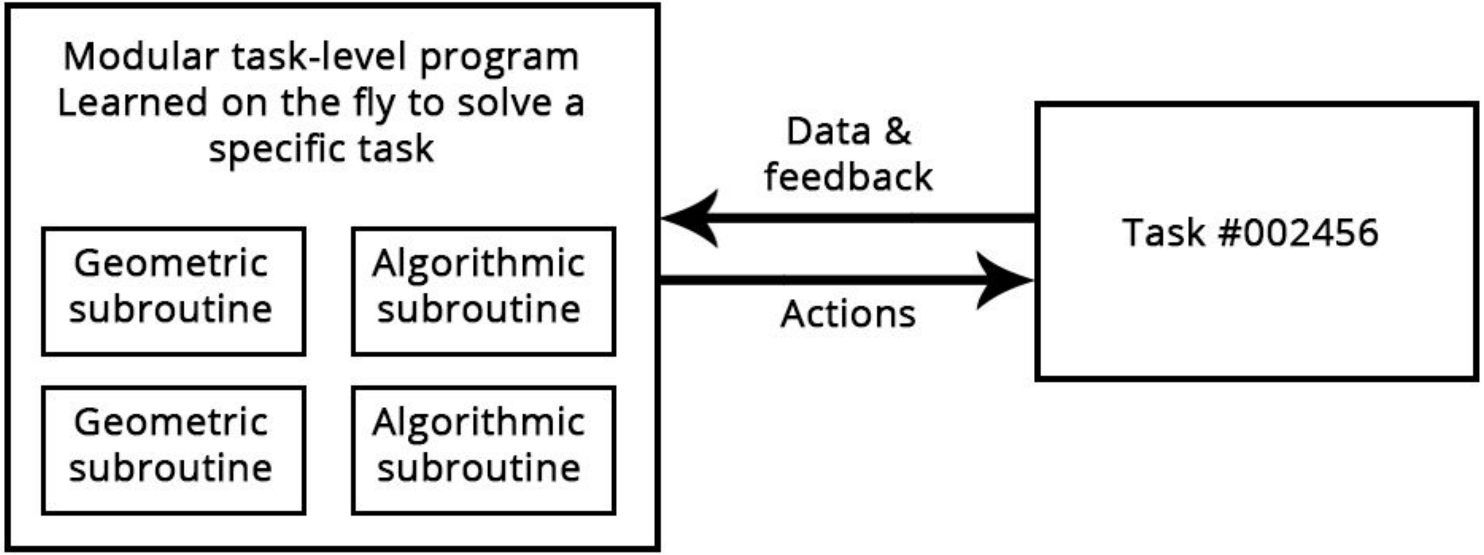
Early examples (with handcrafted symbolic modules)

AlphaGo





Blending symbolic AI (programming) & geometric AI (deep learning)



We can’t learn these modules on every new task (too complex, too little data)

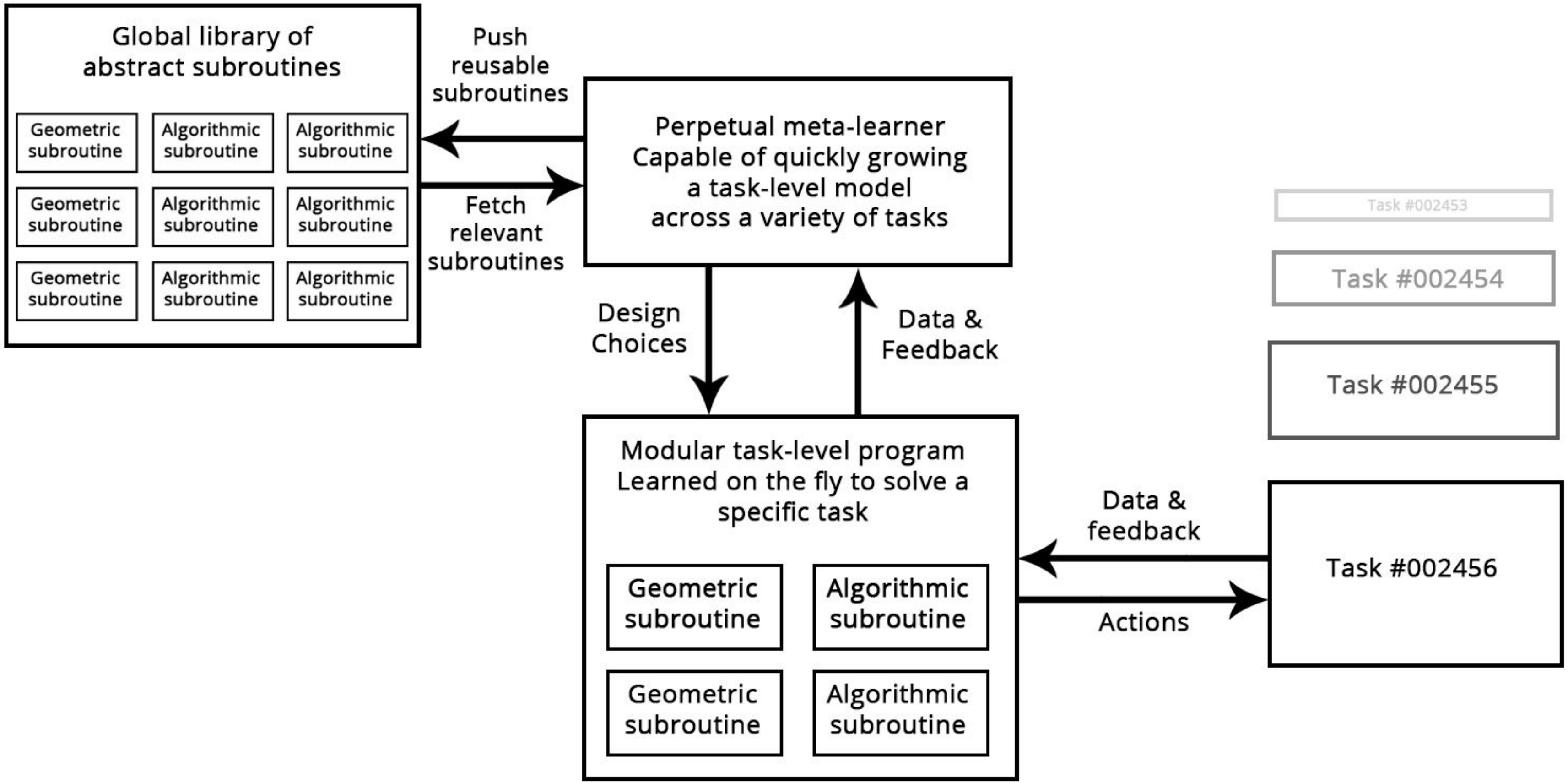
We’ll need a **library** of reusable symbolic & geometric modules, shared across many tasks, many systems

We'll need **stronger priors**

Lifelong learning across many tasks and many domains

Learn subroutines reusable across diverse problem domains

Lifelong learning and multi-task learning



The long-term vision:
AI resembling current human-driven
software development workflows

You can also compare it to **science**, which itself is
a form of non-biological Strong AI

Thank you <3

François Chollet
@fchollet