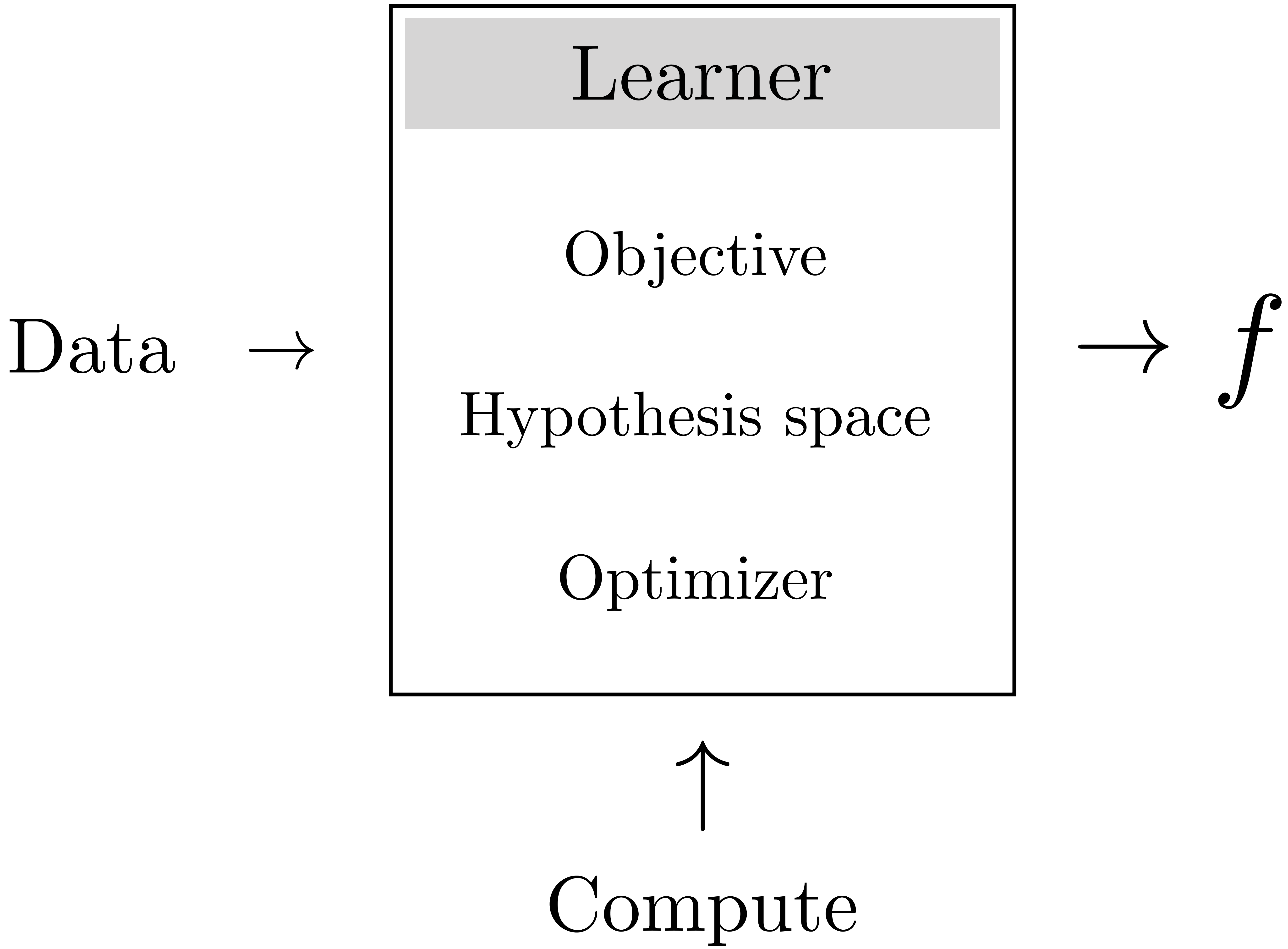
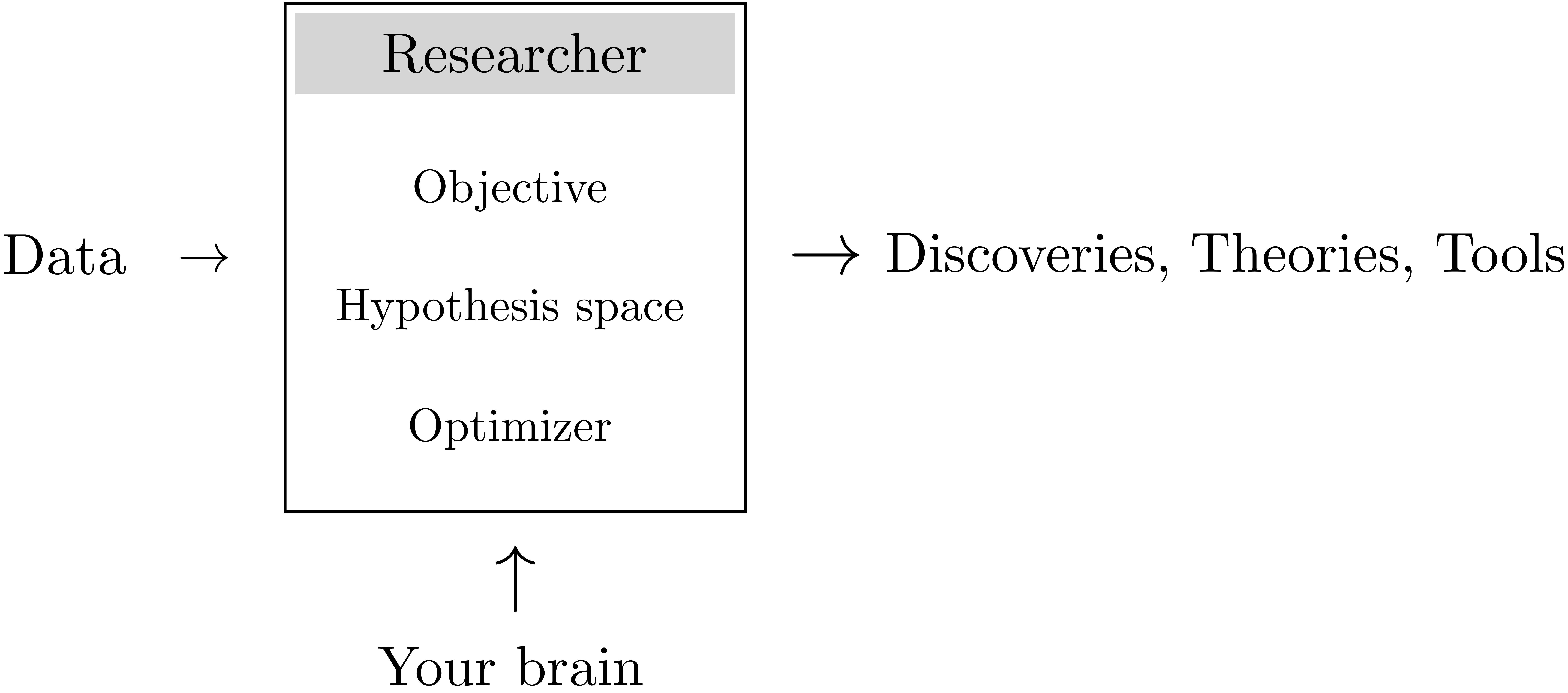
How to do research

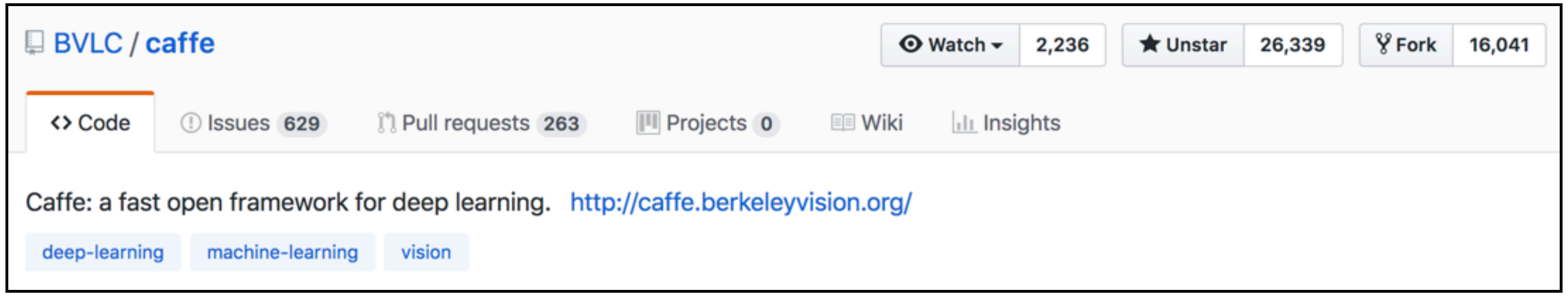
Phillip Isola, 2019


Learner
Objective
Hypothesis space
Optimizer
Data →
→ f
↑
Compute

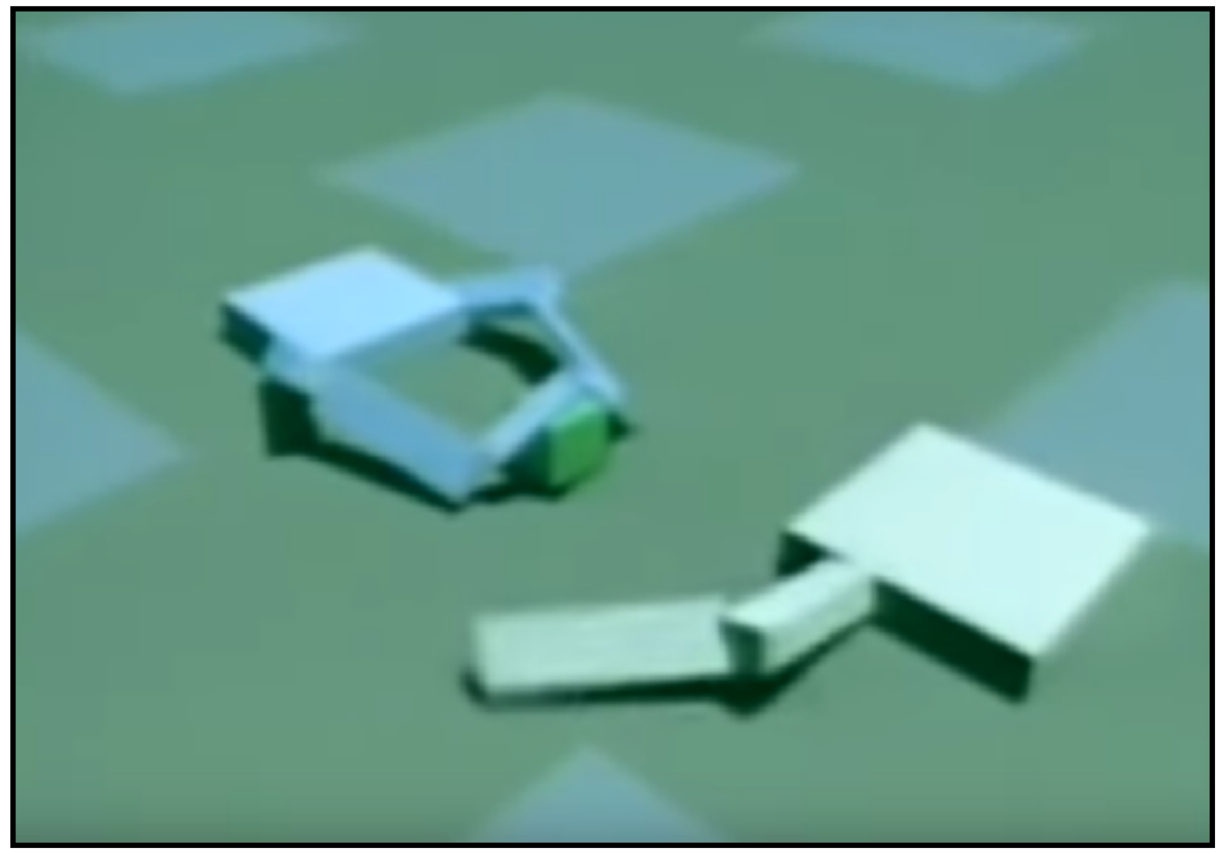

Data
Researcher
Objective
Hypothesis space
Optimizer
Discoveries, Theories, Tools
Your brain

There are many ways to conribute

Tools



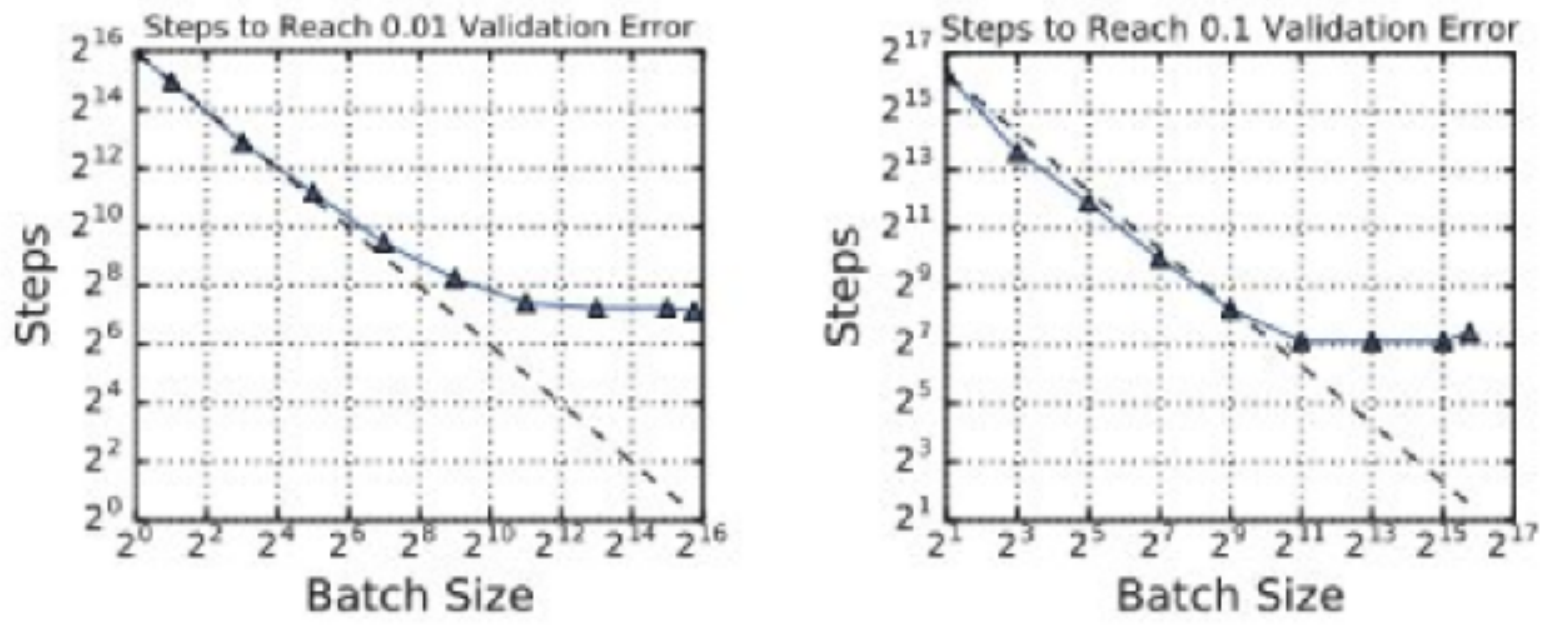
Creativity



Theory

Theorem 4. *(weak* topology) Let* $\{\mathbb{P}_n\}$ *be a sequence of distributions. Considering* $n \to \infty$, *under mild Assumption,* $\max_\phi M_{f_\phi}(\mathbb{P}_\mathcal{X}, \mathbb{P}_n) \to 0 \iff \mathbb{P}_n \xrightarrow{D} \mathbb{P}_\mathcal{X}$*, where* $\xrightarrow{D}$ *means* converging in distribution *[3]*.

Empiricism

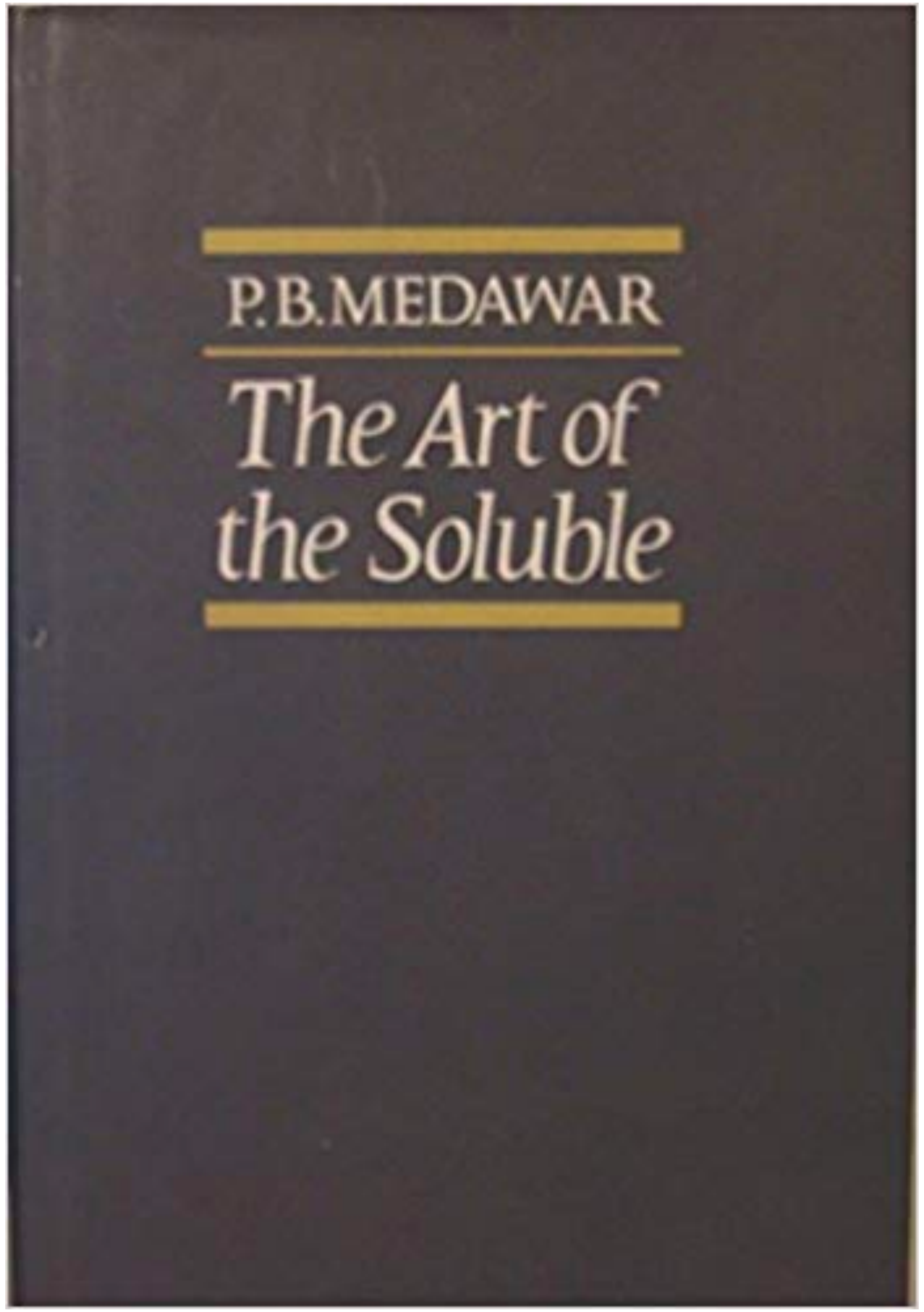


(a) Simple CNN on MNIST (b) Simple CNN on Fashion MNIST

Communication

Picking a topic

Science is the "Art of the Soluble"



"'Good scientists study the most important problems they think they can solve. It is, after all, their professional business to solve problems not to grapple with them.' —Peter Medawar"
— Jitendra Malik

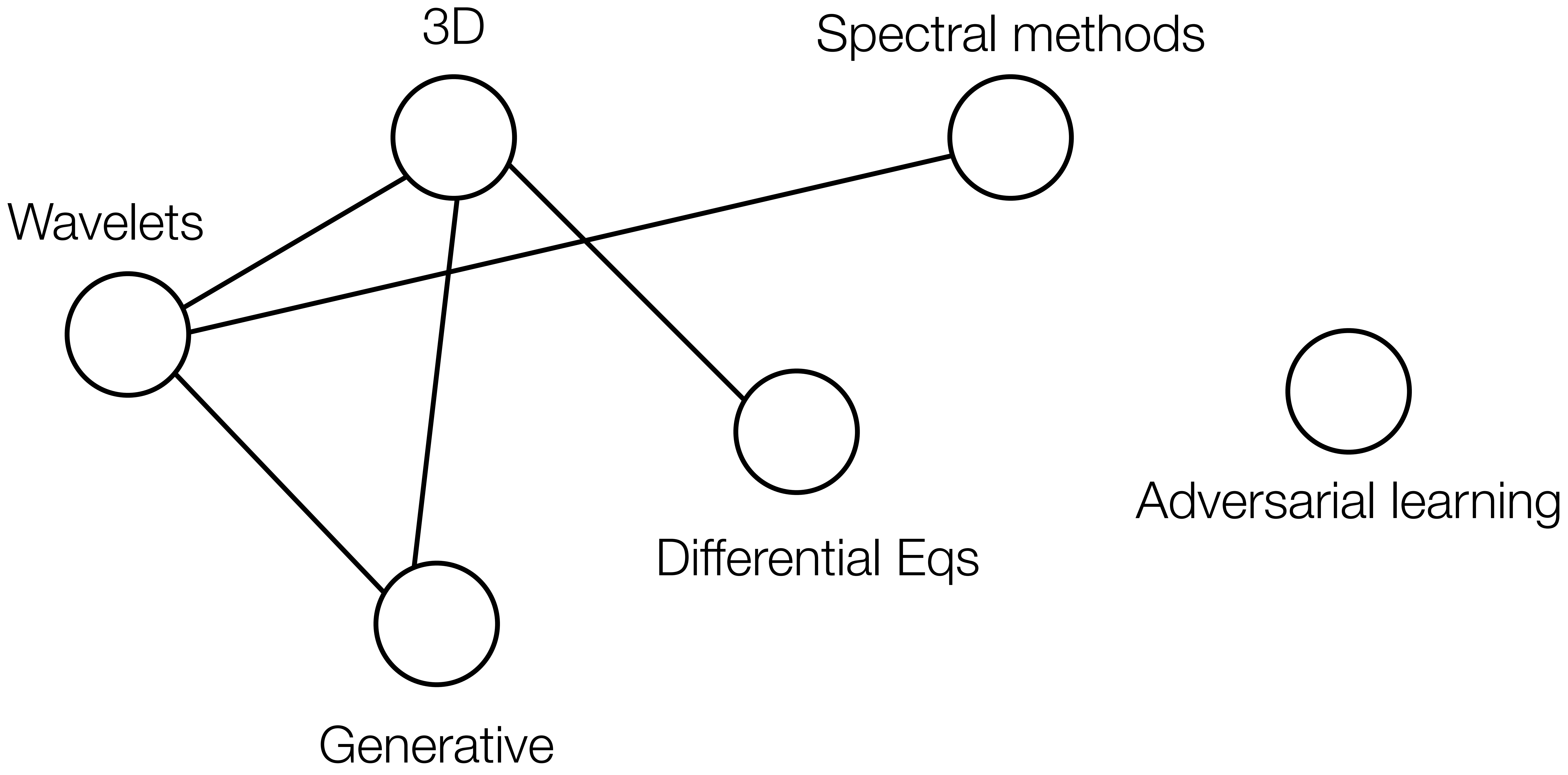
Identify limitations

	Method	Train on data	One-pass Sampling	Exact log-likelihood	Free-form Jacobian
	Variational Autoencoders	✓	✓	✗	✓
	Generative Adversarial Nets	✓	✓	✗	✓
	Likelihood-based Autoregressive	✓	✗	✓	✗
Change of Variables	Normalizing Flows	✗	✓	✓	✗
	Reverse-NF, MAF, TAN	✓	✗	✓	✗
	NICE, Real NVP, Glow, Planar CNF	✓	✓	✓	✗
	FFJORD	✓	✓	✓	✓

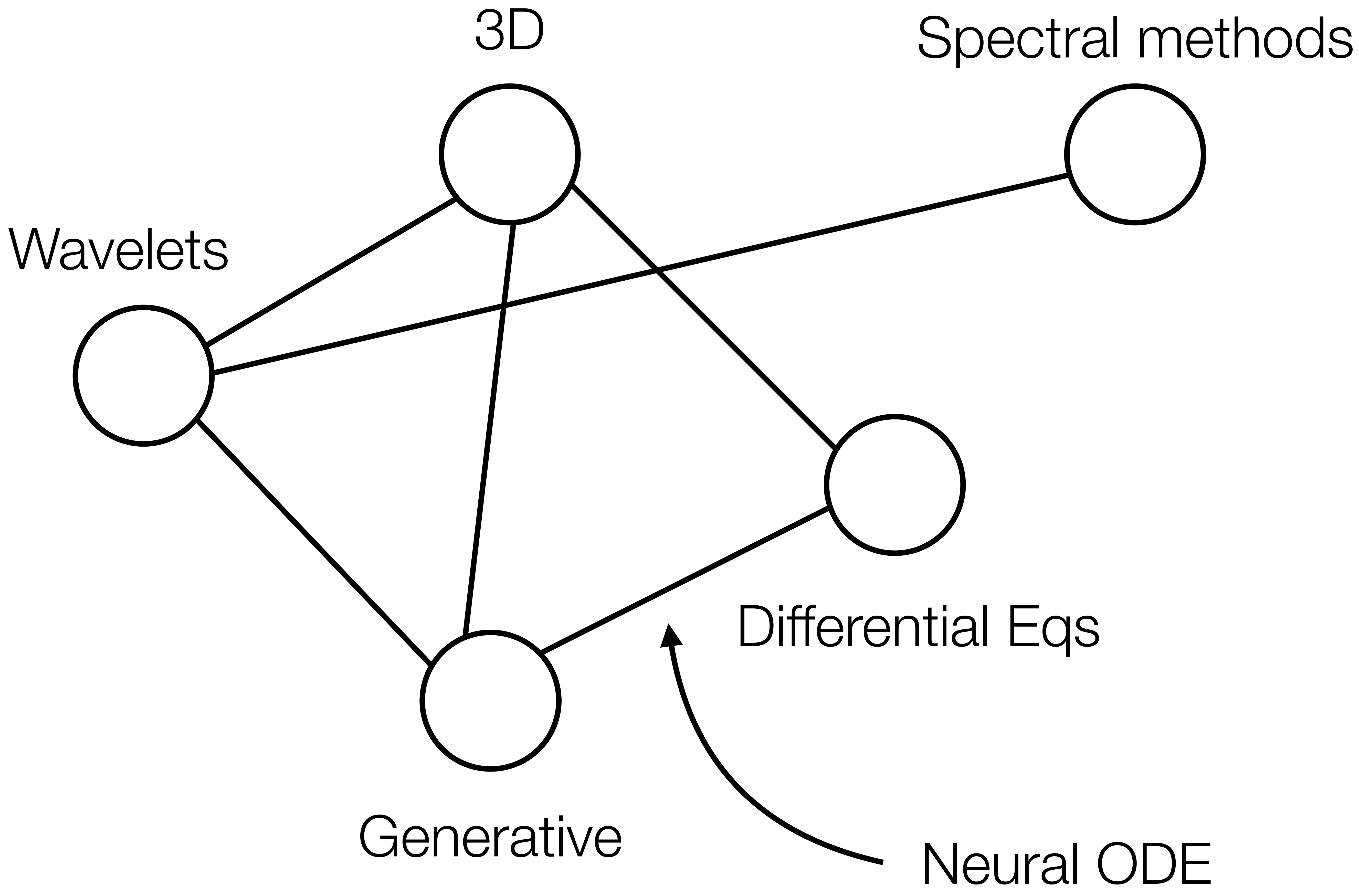
Table 1: A comparison of the abilities of generative modeling approaches.

[“FFJORD”, Grathwohl et al., 2018]

Add a node


3D
Spectral methods
Wavelets
Adversarial learning
Differential Eqs
Generative

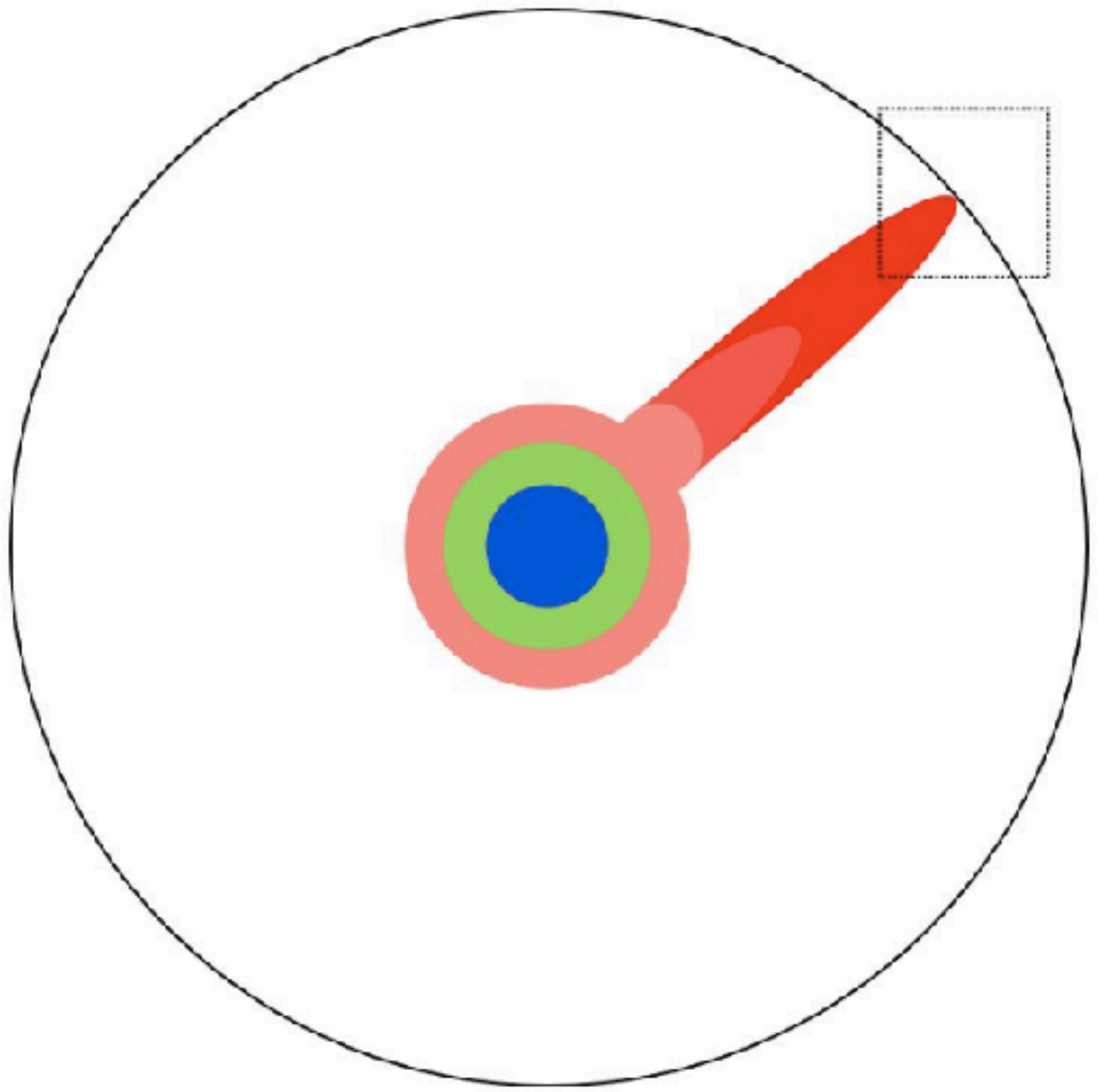
Add an edge

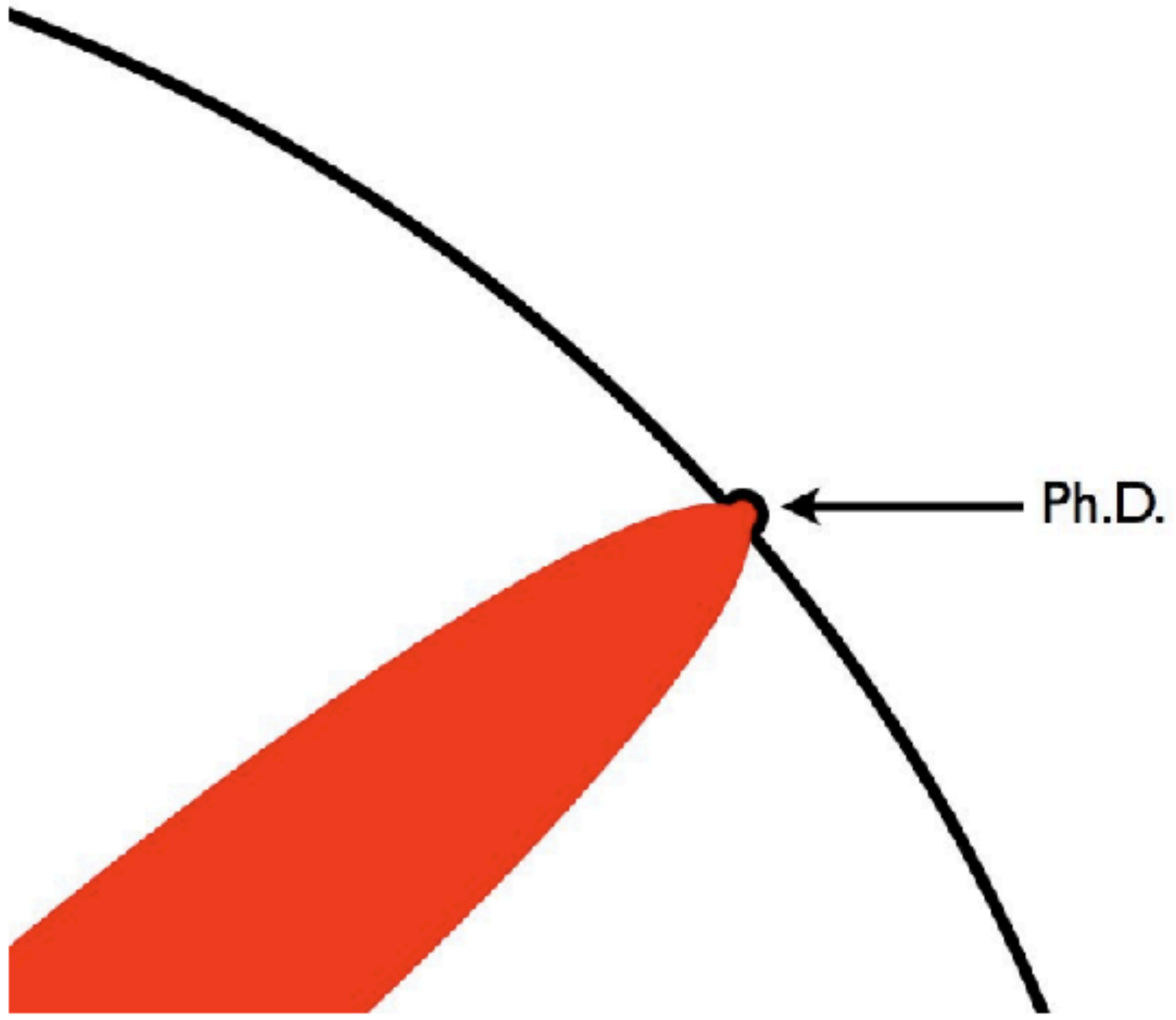

3D
Spectral methods
Wavelets
Differential Eqs
Generative
Neural ODE

Novelty

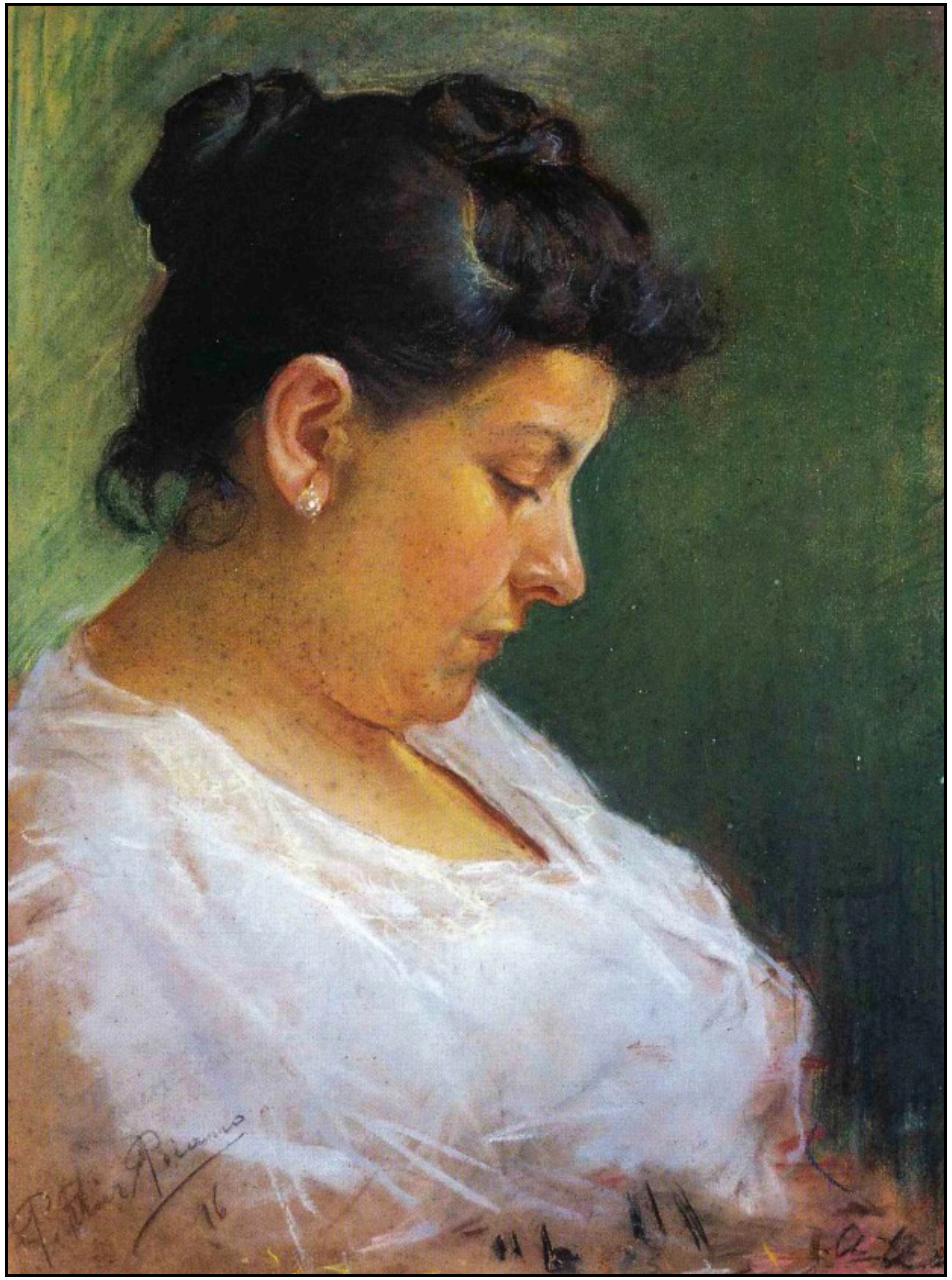
What matters is novelty w.r.t. humanity.

Very hard to achieve without knowing what has already been done.

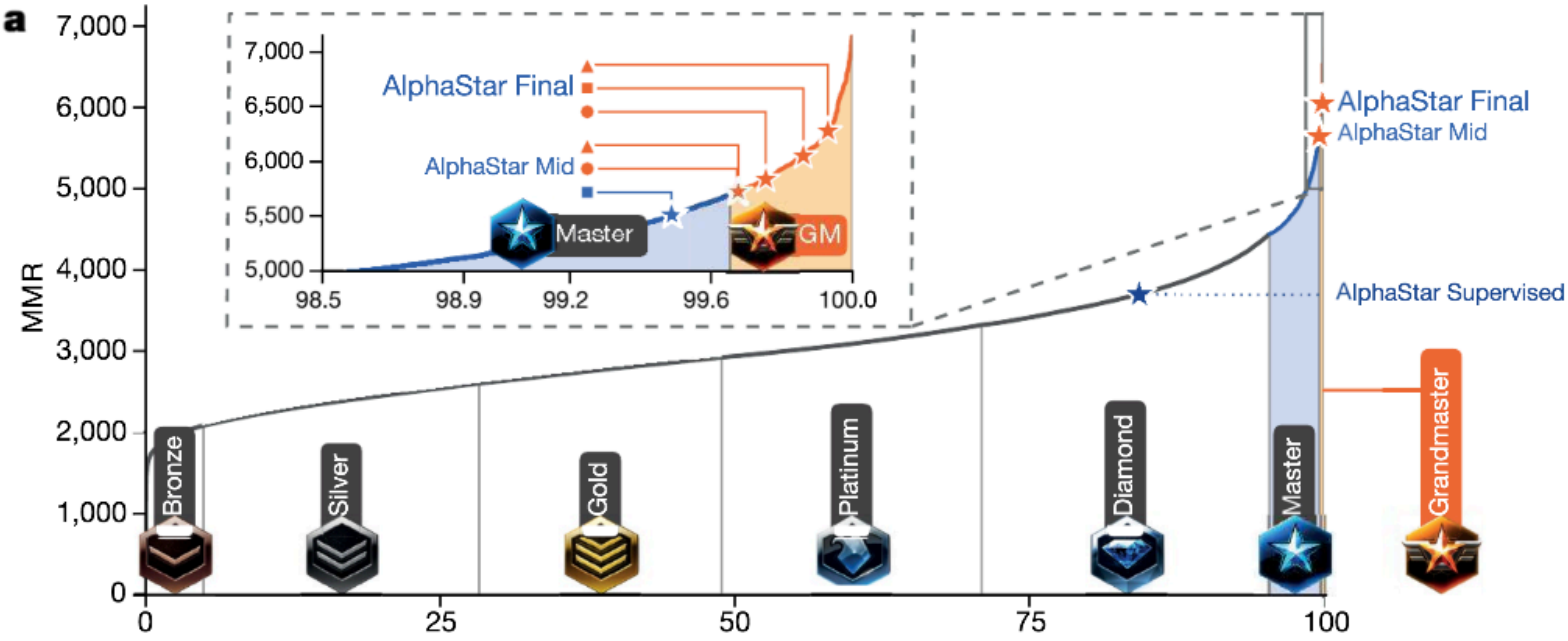




http://matt.might.net/articles/phd-school-in-pictures/



[Picasso]



First imitates humans, then innovates

["AlphaStar", Vinyals et al., Nature 2019]

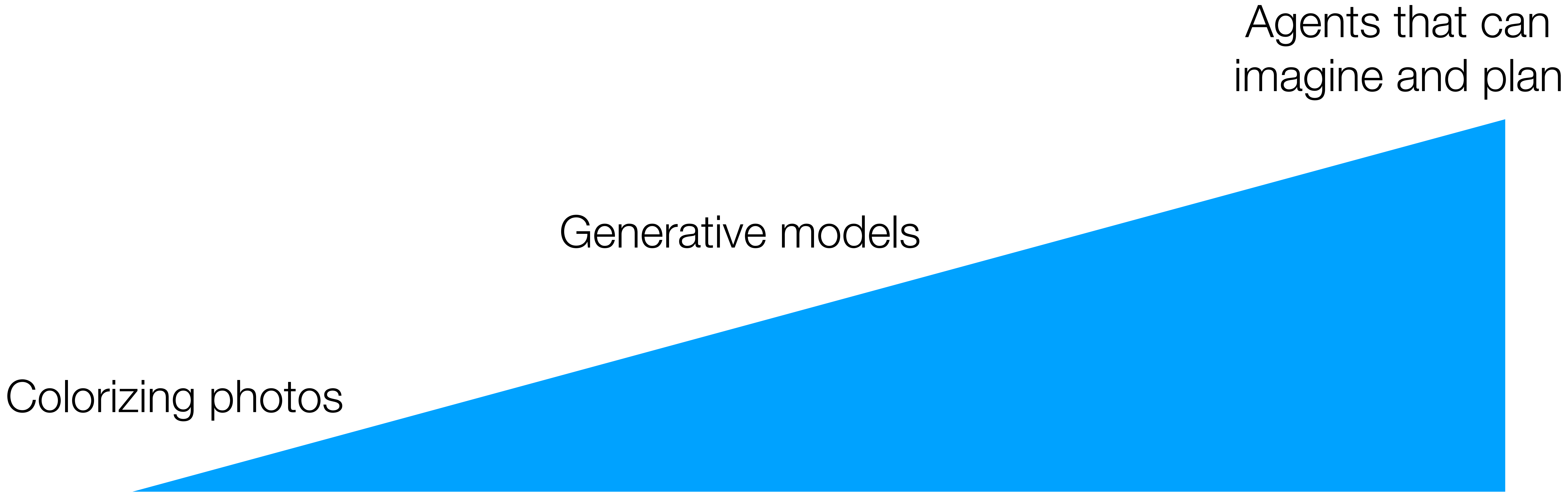
Know something no one else knows

- It’s necessary but not sufficient to master the core knowledge of your field

- Acquire a unique skillset
  - Take difficult or unusual classes
  - Read old papers
  - Take on a complementary hobby
  - Talk to people in other fields

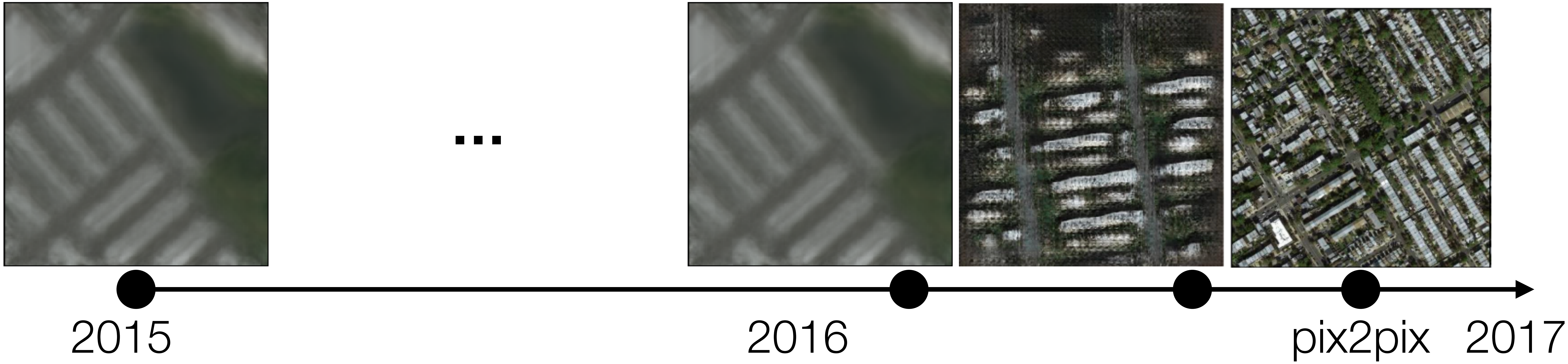
“My answer to "Now What" is "here is a research problem which is unusual, perhaps significant, novel, that I can pose and probably solve because of my background in physics". The situation would not be readily identified as a problem at all by those whose background seems much more relevant than my own.”
— “Now What”, John Hopfield

Build a ramp



“When you are famous it is hard to work on small problems. This is what did Shannon in. After information theory, what do you do for an encore? The great scientists often make this error. They fail to continue to plant the little acorns from which the mighty oak trees grow.”
— Richard Hamming, “You and Your Research”

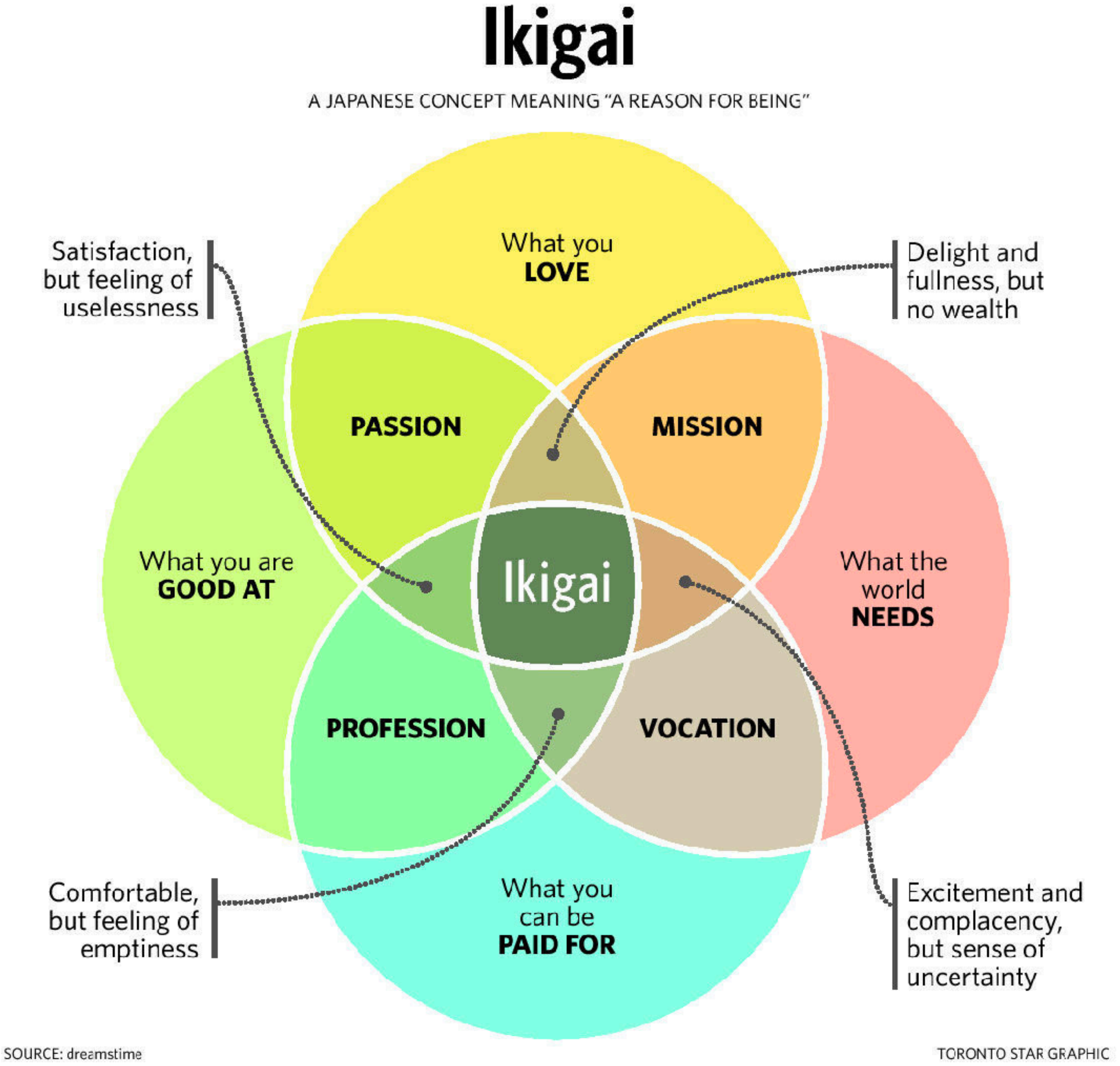
Research takes time



Picking a topic

Don’t worry too much if a topic is unpopular

If it is important, and you do good work on it, you can *make it popular*



Doing good work on that topic

Researcher

Objective

Hypothesis space

Optimizer

Data → Researcher → Discoveries, Theories, Tools

↑

Your brain

Our goal is to produce *knowledge that generalizes*

"One kind of result which probably won't generalize is: "algorithm A works better than algorithm B." Different application areas have their own requirements… The kind of result I believe generalizes to new situations is the na of the tradeoffs between different approaches."

— "Which Research Results Will Generalize?" Roger Grosse [https://lips.cs.princeton.edu/which-research-results-will-generalize/]

Fit the data

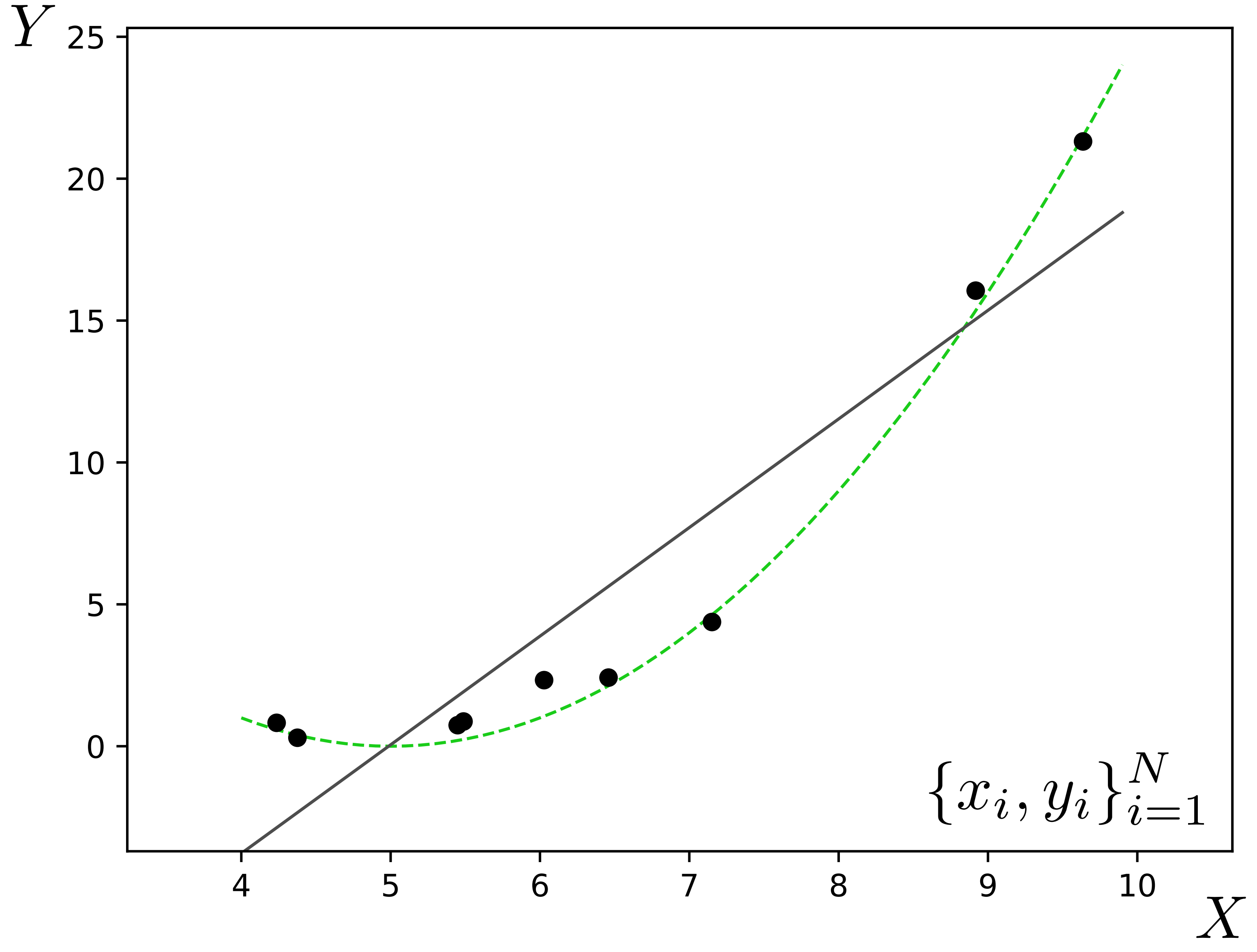
“Now I’m going to discuss how we would look for a new law. In general, we look for a new law by the following process. First, we guess it (audience laughter), no, don’t laugh, that’s the truth. Then we compute the consequences of the guess, to see what, if this is right, if this law we guess is right, to see what it would imply and then we compare the computation results to nature or we say compare to experiment or experience, compare it directly with observations to see if it works.

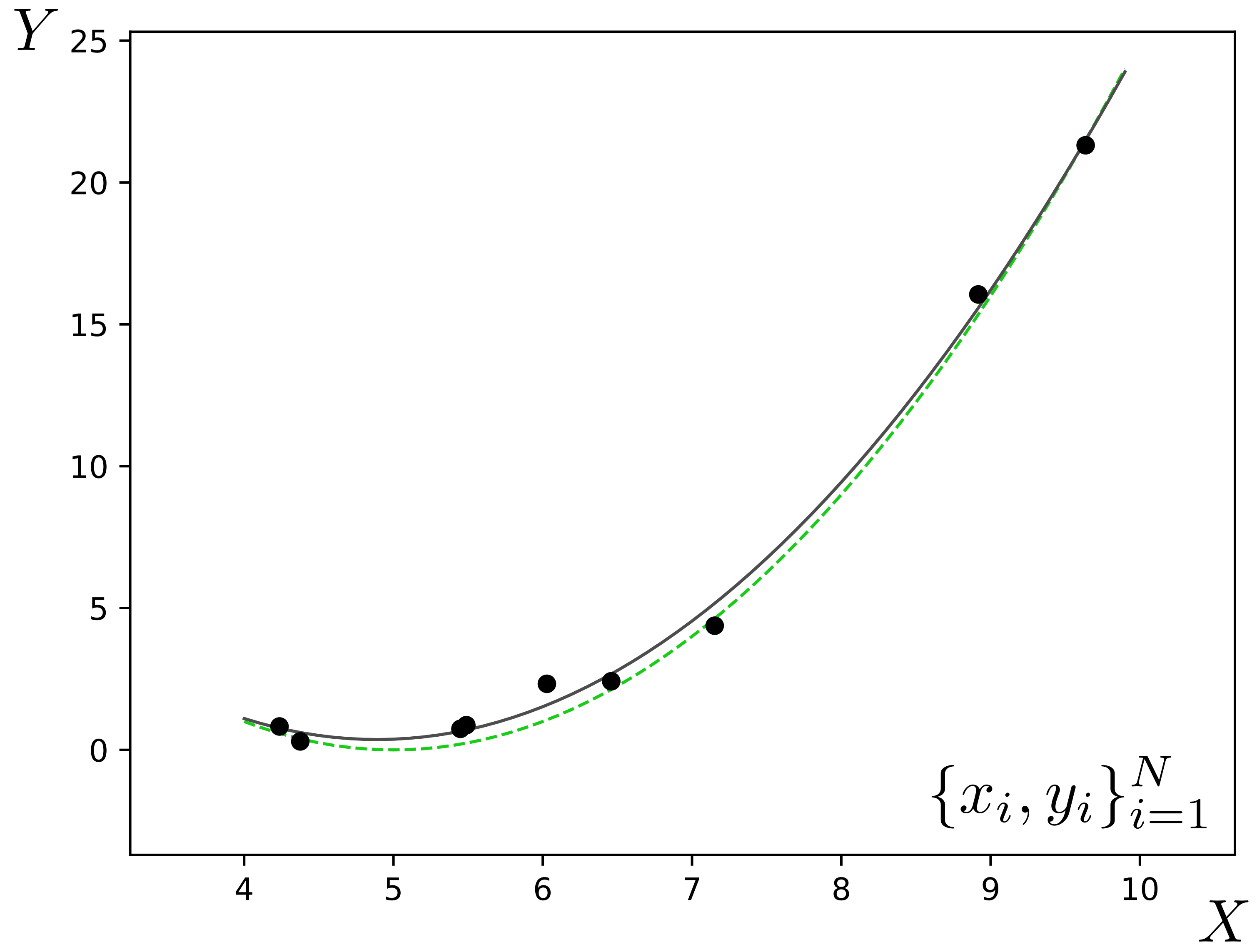
If it disagrees with experiment, it’s wrong. In that simple statement is the key to science. It doesn’t make any difference how beautiful your guess is, it doesn’t matter how smart you are who made the guess, or what his name is … If it disagrees with experiment, it’s wrong. That’s all there is to it.”
— Richard Feynman

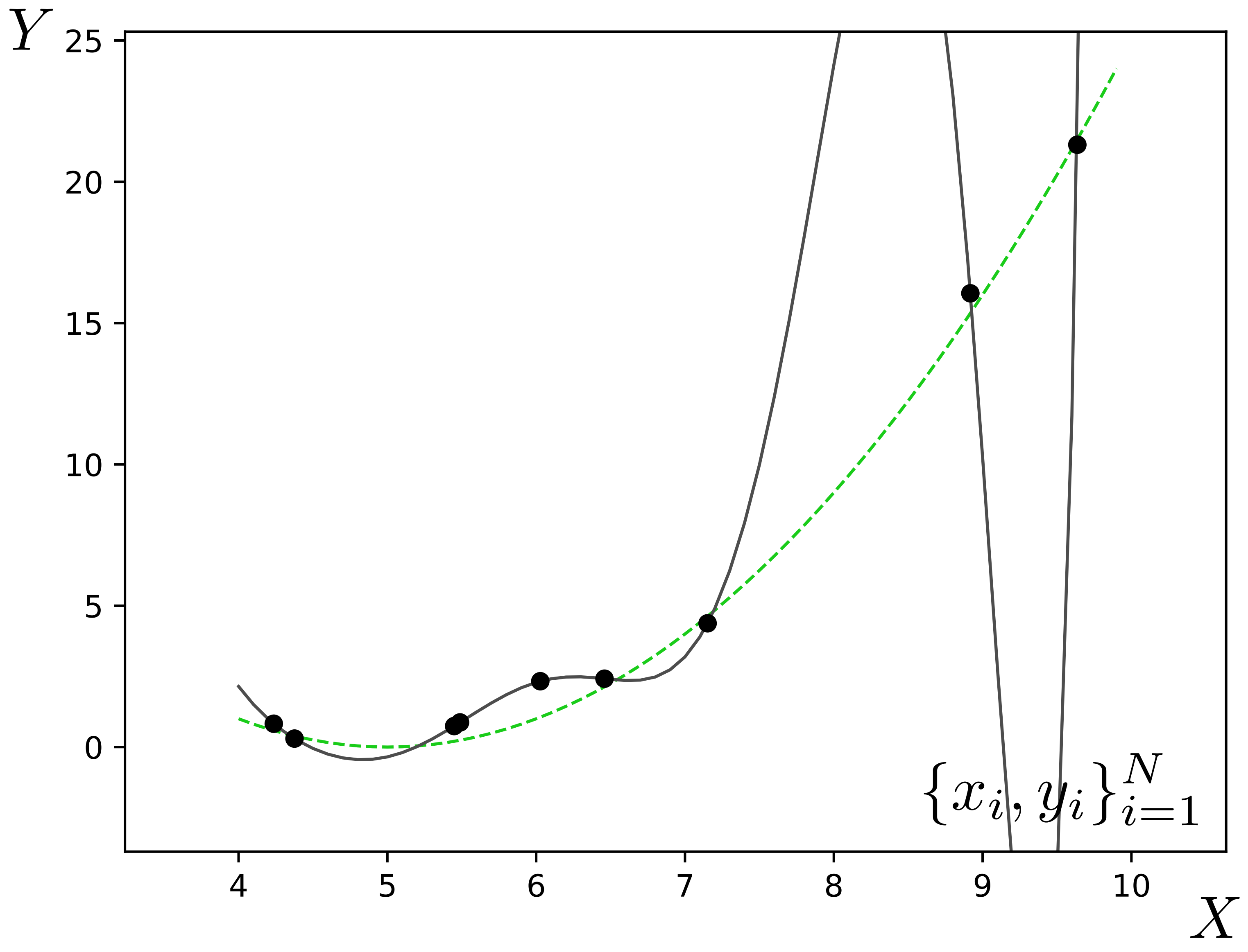
Omit needless bits

"Perfection is finally attained not when there is no longer anything to add, but when there is no longer anything to take away"
— Antoine de Saint Exupéry

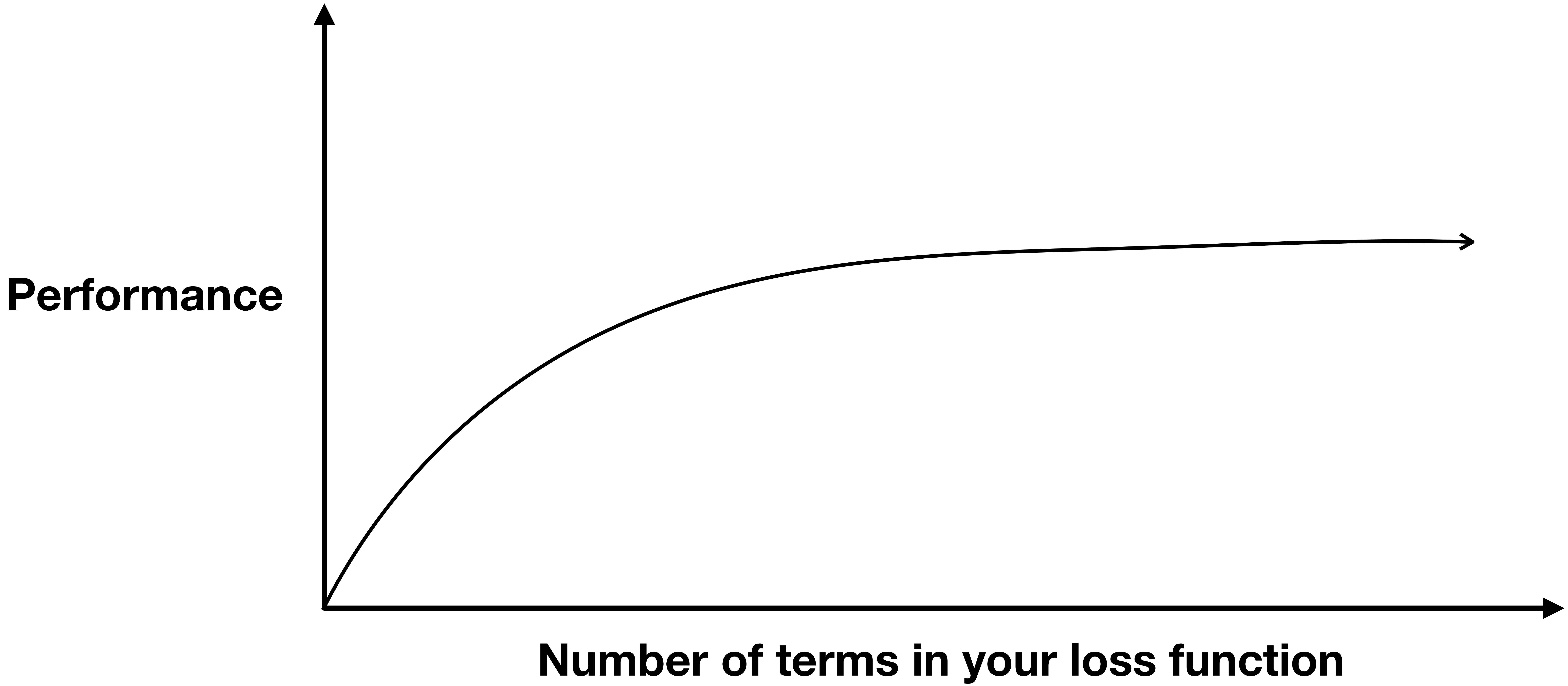
Underfitting	Appropriate model	Overfitting
$\boxed{K = 1}$	$\boxed{K = 2}$	$\boxed{K = 10}$



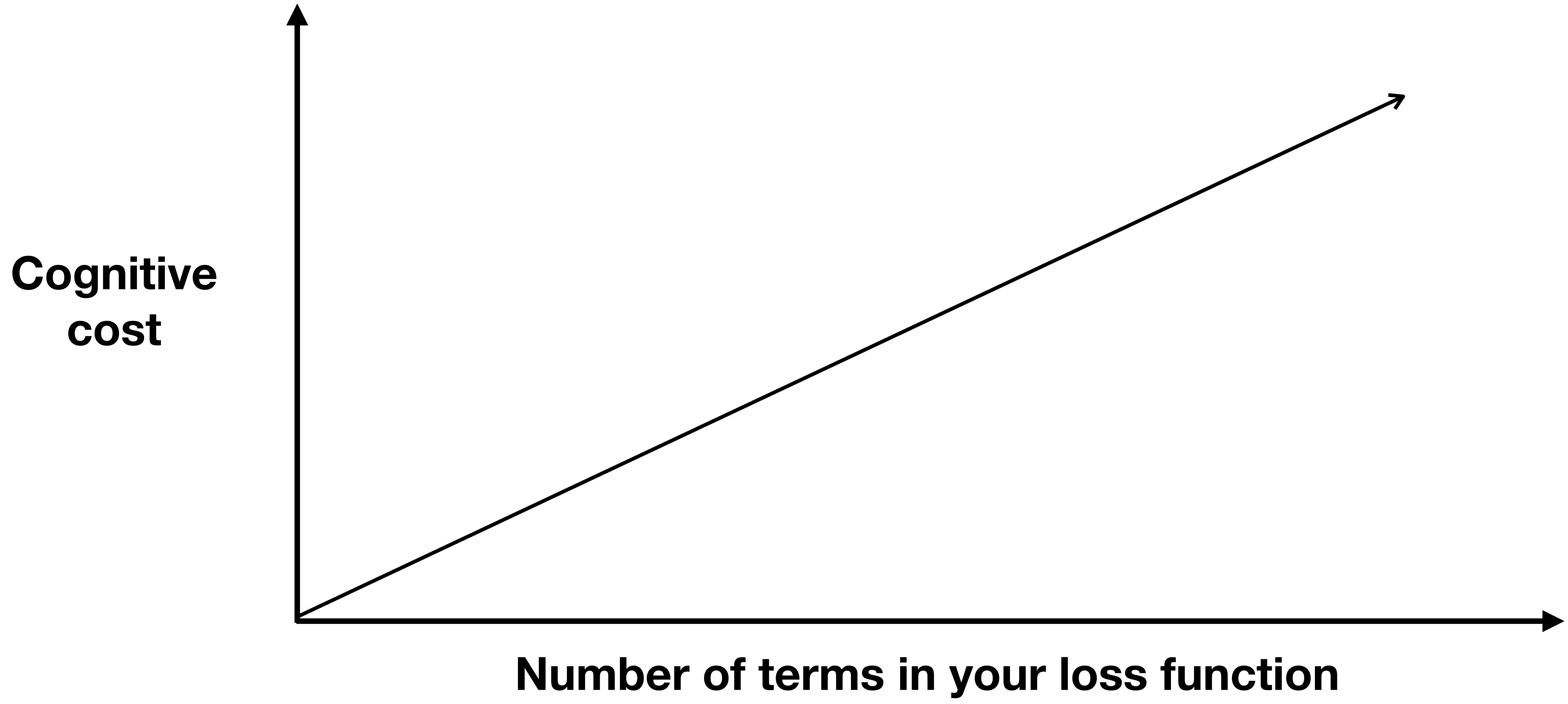




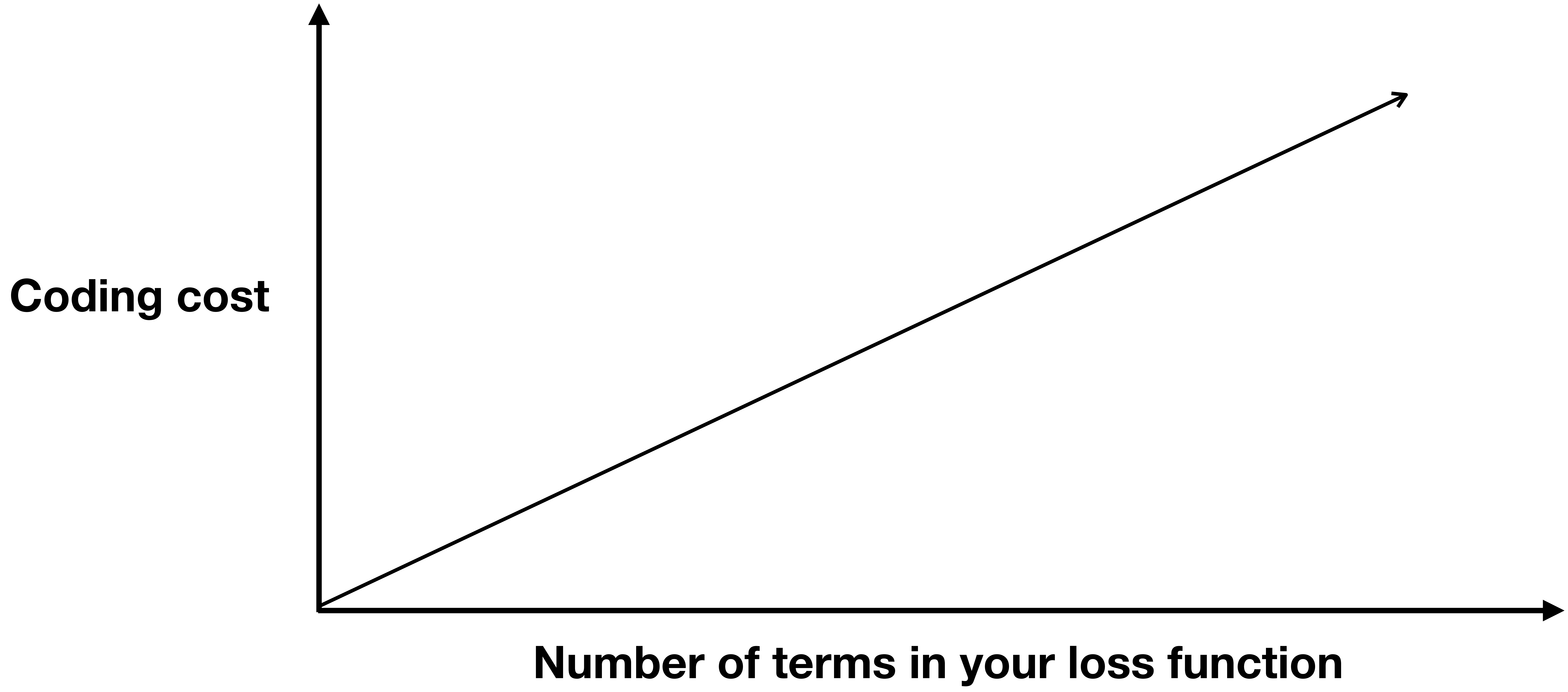
Information creep



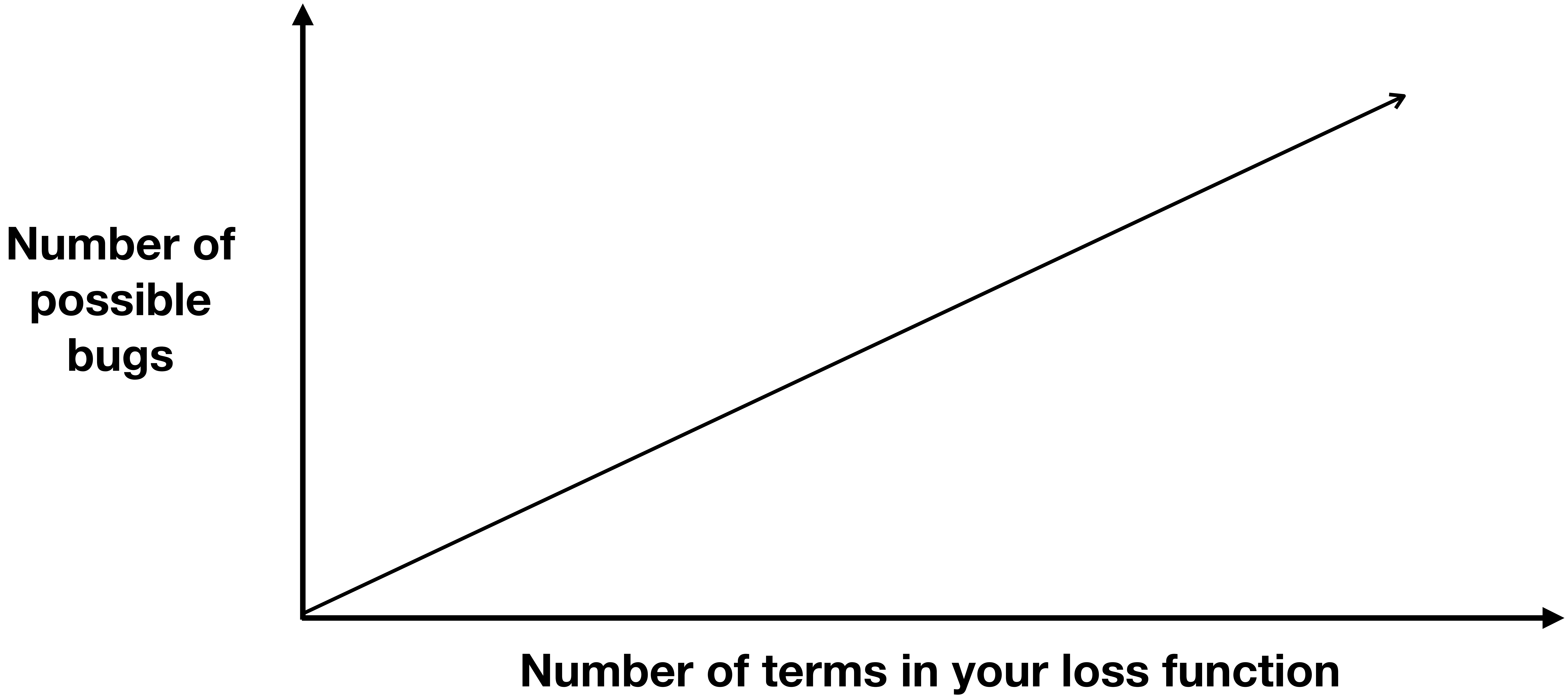
Information creep



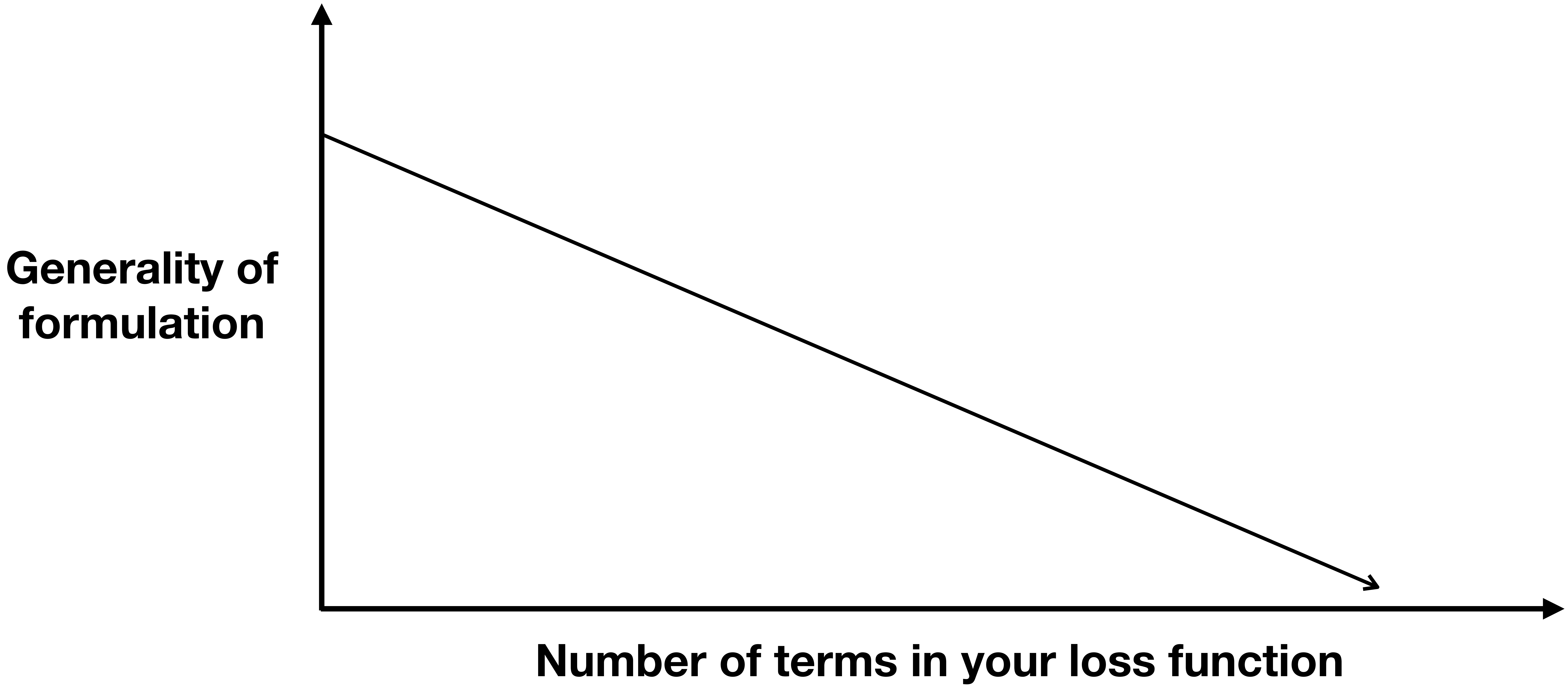
Information creep



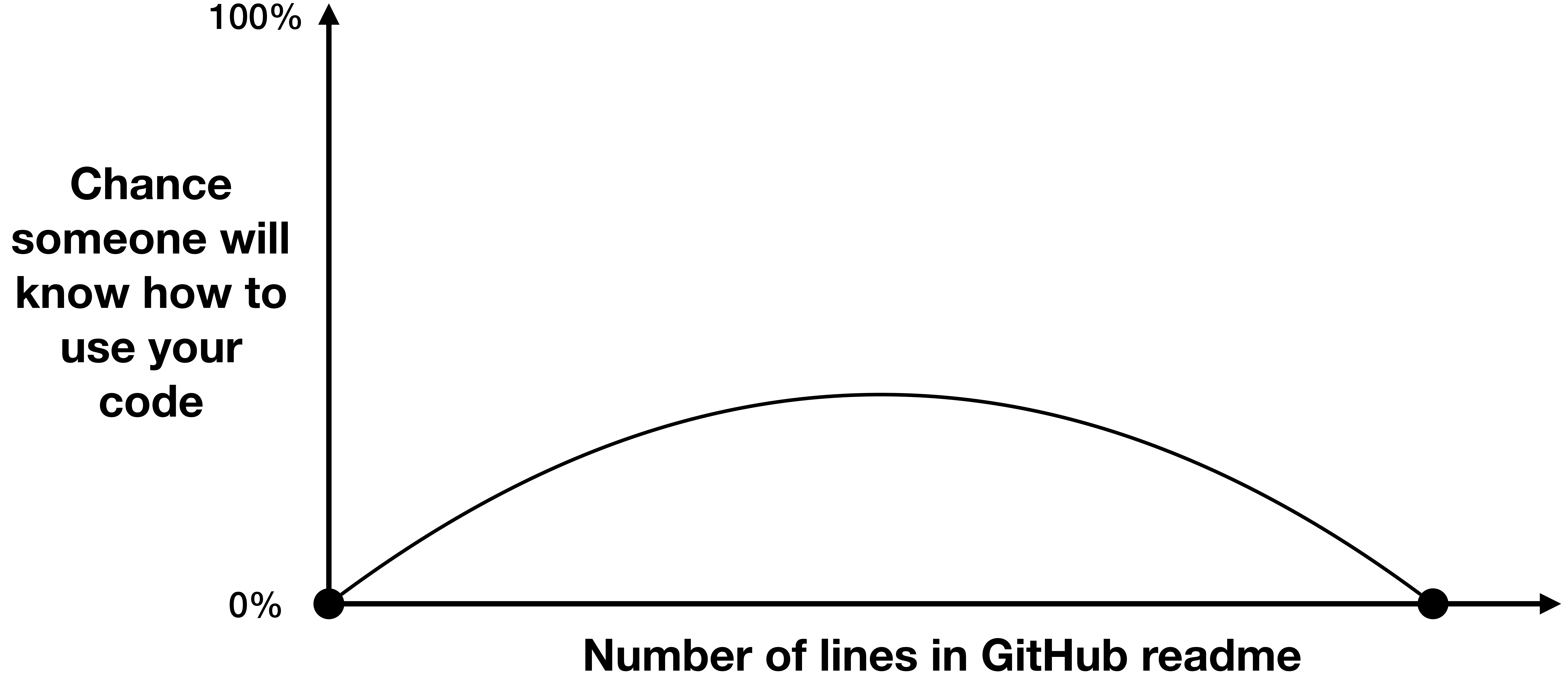
Information creep



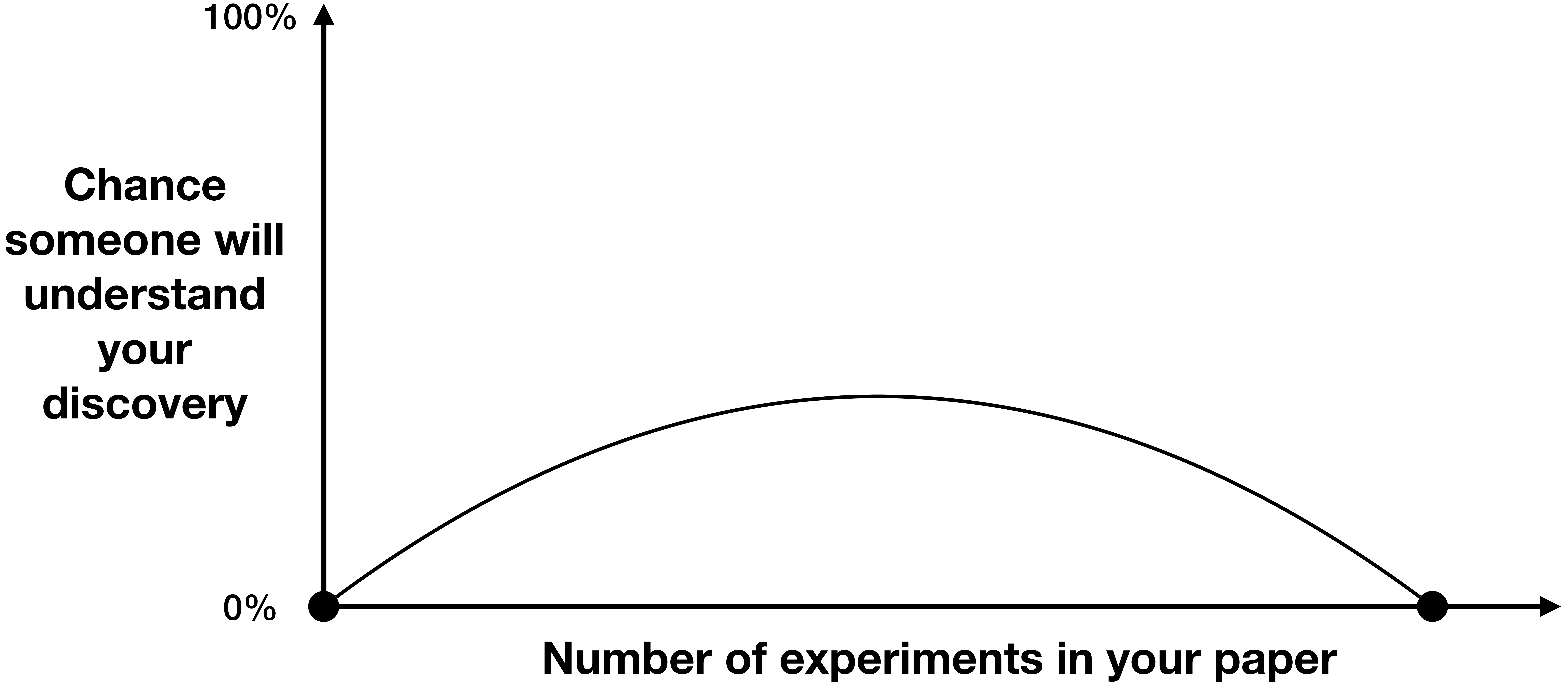
Information creep



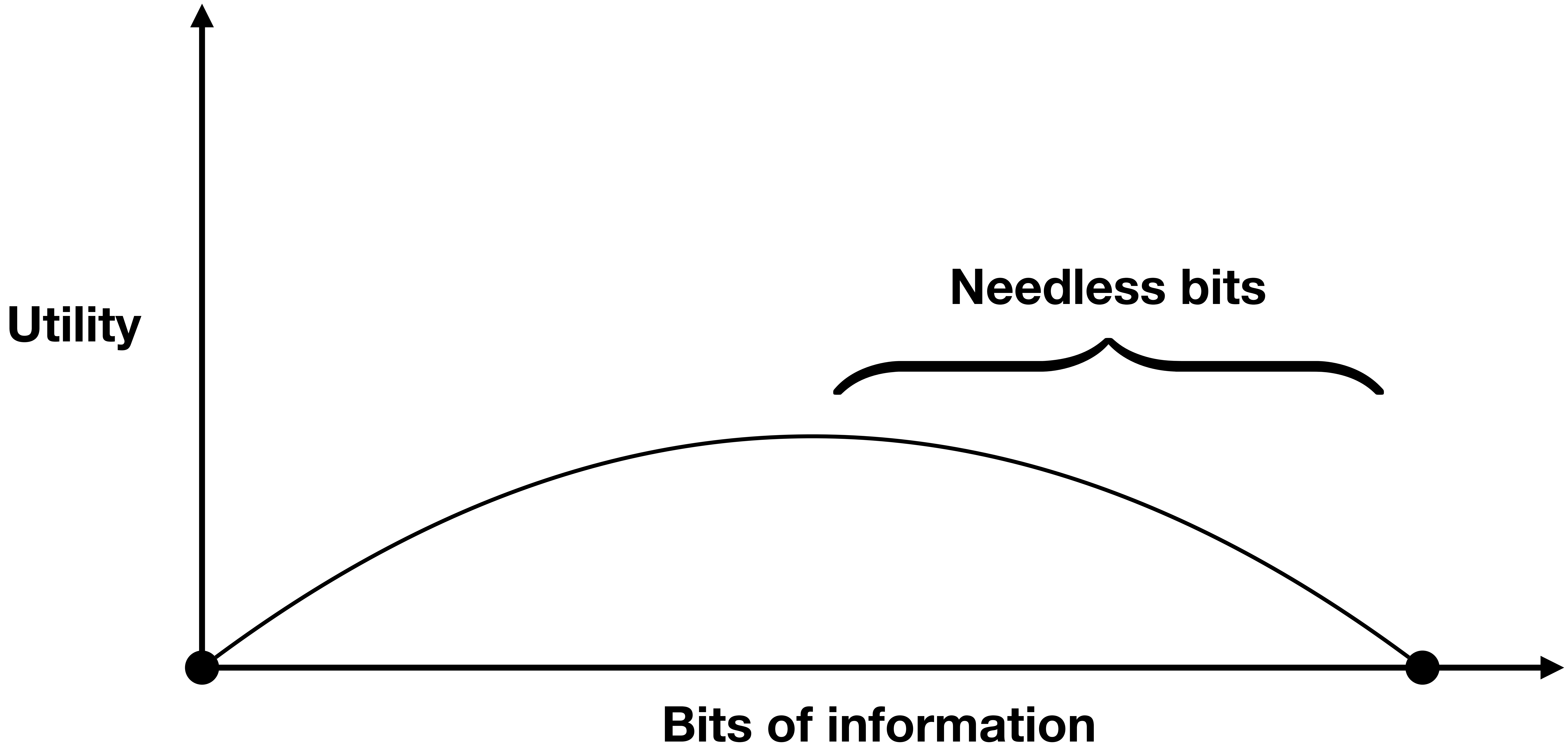
Information creep



Information creep



Information creep



Do the most, with the least

$$\frac{\text{Products: Discoveries, Explanations, Results, Tools, ...}}{\text{Costs: Words, Equations, Concepts, Lines of code, GPUs, People, Time, Money, ...}} = \text{Metric for research}$$

Products

Discoveries
Explanations
Results
Tools
...

Costs

Words
Equations
Concepts
Lines of code
GPUs
People, Time, Money
...

= Metric for research

Avoid fallacies of conspicuous consumption

— This paper is really great, it used thousands of GPUs!

— The equations in this paper are so hard to decipher, I bet it is really powerful stuff

— “We present a simple but effective approach”

— The human brain is so fantastically complex we cannot hope to match it with today’s basic algorithms

Alpern’s razor (via Ted Adelson):
“Among competing hypotheses, the most boring is the most likely to be true.”

The regularizing force of human fallibility

- The vast majority of information at any given academic conference is forgotten — that's a good thing
- We tend to forgot all but the simplest and starkest discoveries
- Complex and subtle discoveries are usually either overfit or unimportant

[https://www.youtube.com/watch?v=mrw4KlP5en0]

Reviews

Review of “Generative Adversarial Networks”, Goodfellow et al. 2014:

“The theoretical work is primitive, and the experiments are pretty basic.”