

On Keeping Things Simple

Angela Bassa
@angebassa, October 2017

Today's Plan



Purposefulness

Complexity

Resilience

Today's Plan



Purposefulness

Complexity

Resilience

Would You Rather?

4

- 100 sensors off by 40%, or 500 off by 10%?
- Only 1% of sensors with truly horrible estimations, or all sensors about 20% off?



Constraints: System

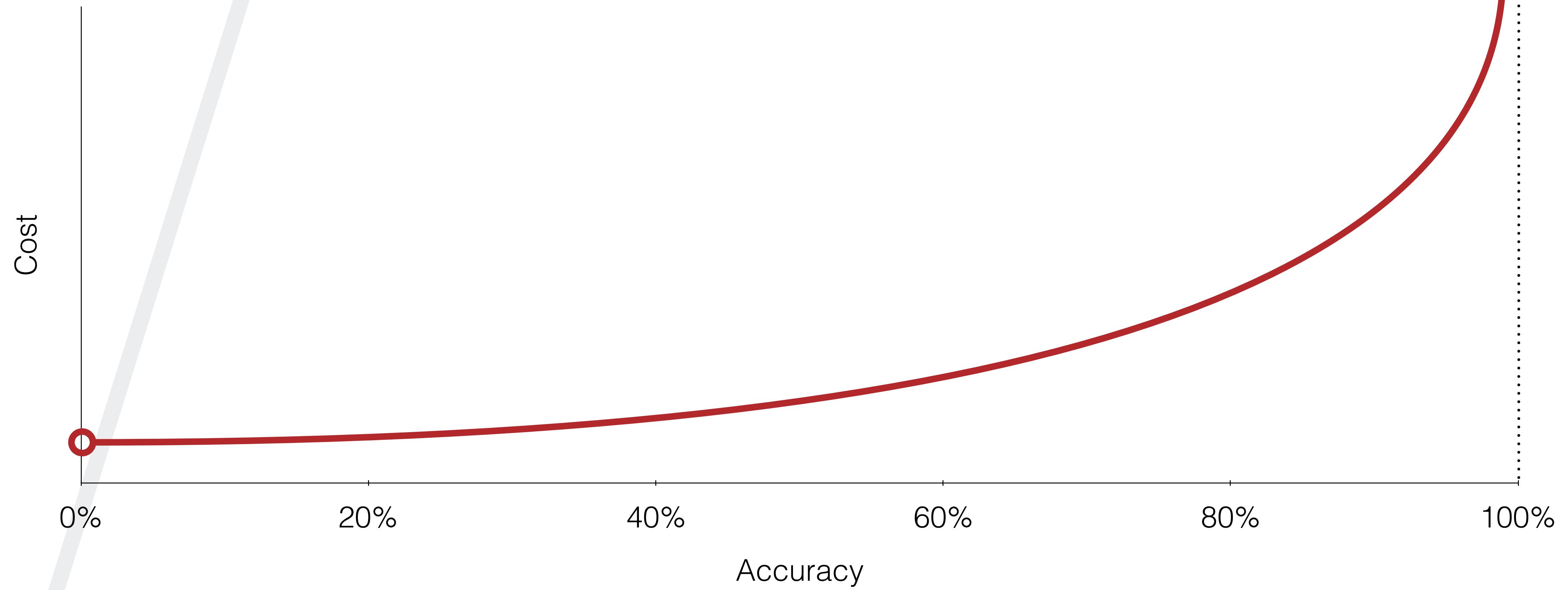
vs.

Solution

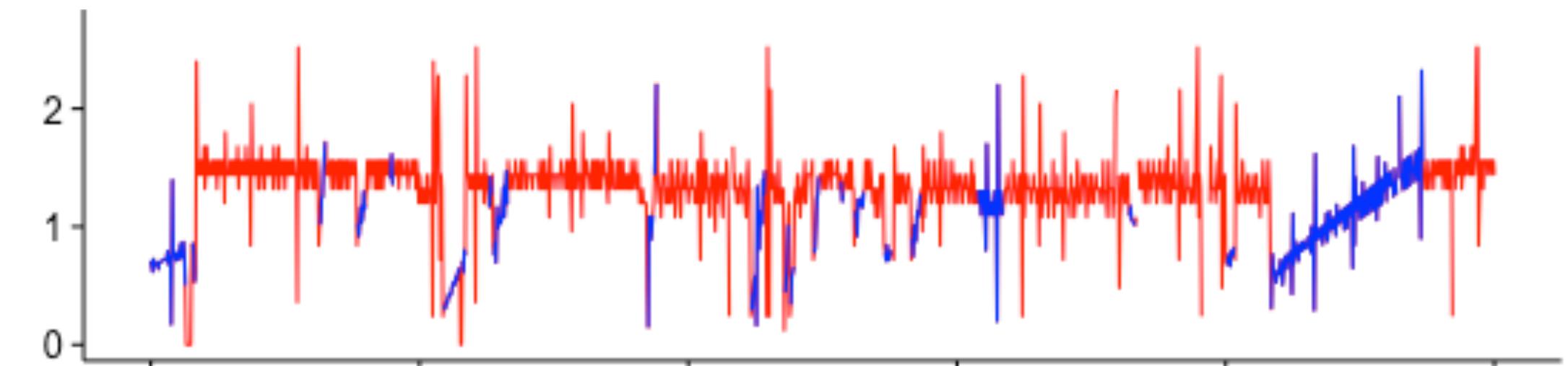
- **Architecture**
(Serverless? Legacy?)
- **Explainability**
(Regulation?)
- **Team structure (SLAs?)**
- **Budget**

- **Accuracy**
- **Speed (<500ms?)**
- **Buildability over time**
- **Supportability**
- **Data handling**

The Cost-Accuracy Tradeoff



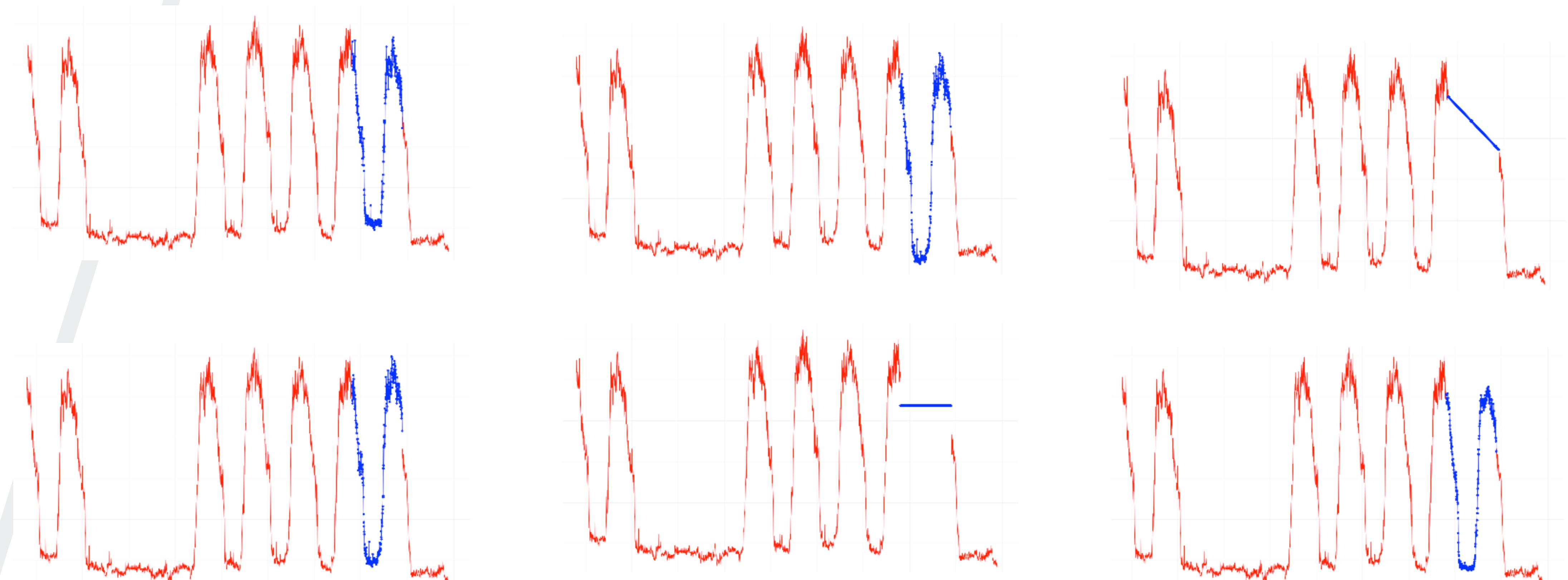
The “Looks Good” Requirement



Mathematically Encode What It Means To “Look Bad:”

- **Edges Don’t Match Up**
- **Values Are Not Typical**
- **Slope Is Not Typical**
- **Smoothness Is Different From Surroundings**

Tradeoffs



Designing A *Functional* Objective Function

- When the prediction gets worse, the cost metric should increase
- The cost metric should not bias the prediction to overestimate (or underestimate)
- The cost metric should not be dominated by a few data points— unless you’re doing this on purpose



10

**There Can
Be Only
One**

“*Meta*” Objective Function

$$= f \left(\text{Model}, \text{Business} \right)$$

$$\begin{aligned} & \cdot \text{Bias/Variance trade-off} & & \cdot \text{Computation costs} \\ & \cdot \text{Computational complexity} & & \cdot \text{Model explainability} \\ & \cdot \text{Cross validation} & & \cdot \text{Product requirements} \\ & \cdot \text{Design optimization} & & \cdot \text{Data availability} \\ & & & \cdot \text{Architectural complexity} \end{aligned}$$

"Your intuition will only
be accurate if your
cumulative experience
is a representative
sample of reality."

Adam Piore

Today's Plan



Purposefulness

Complexity

Resilience

Garbage In Garbage Out

At Scale

14





15

**It's Not That
No One
Cares**

**It's That
Non-Data
Folks Are
Unprepared
To Care**

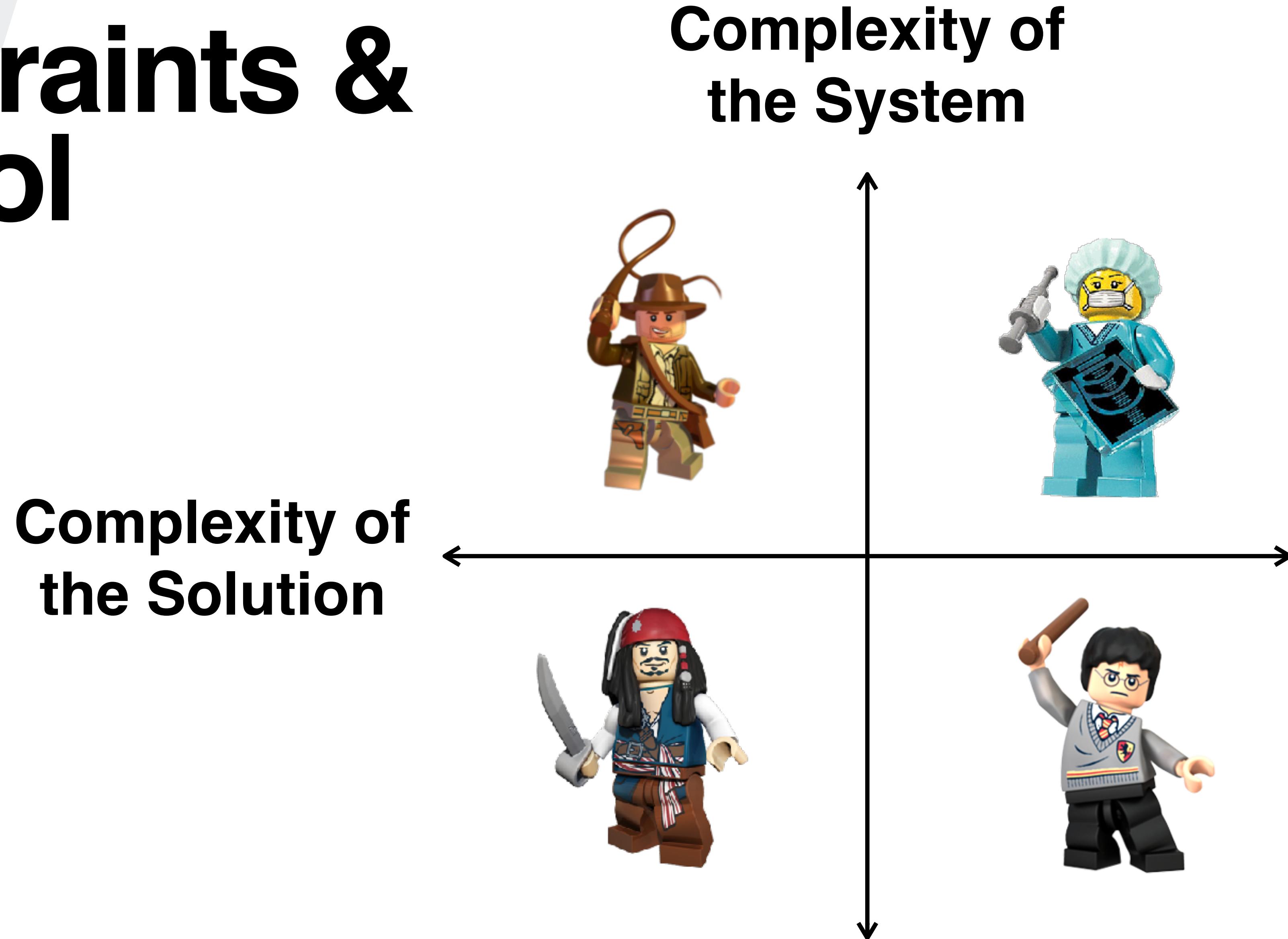
@angebassa

How Do We Measure Simplicity?

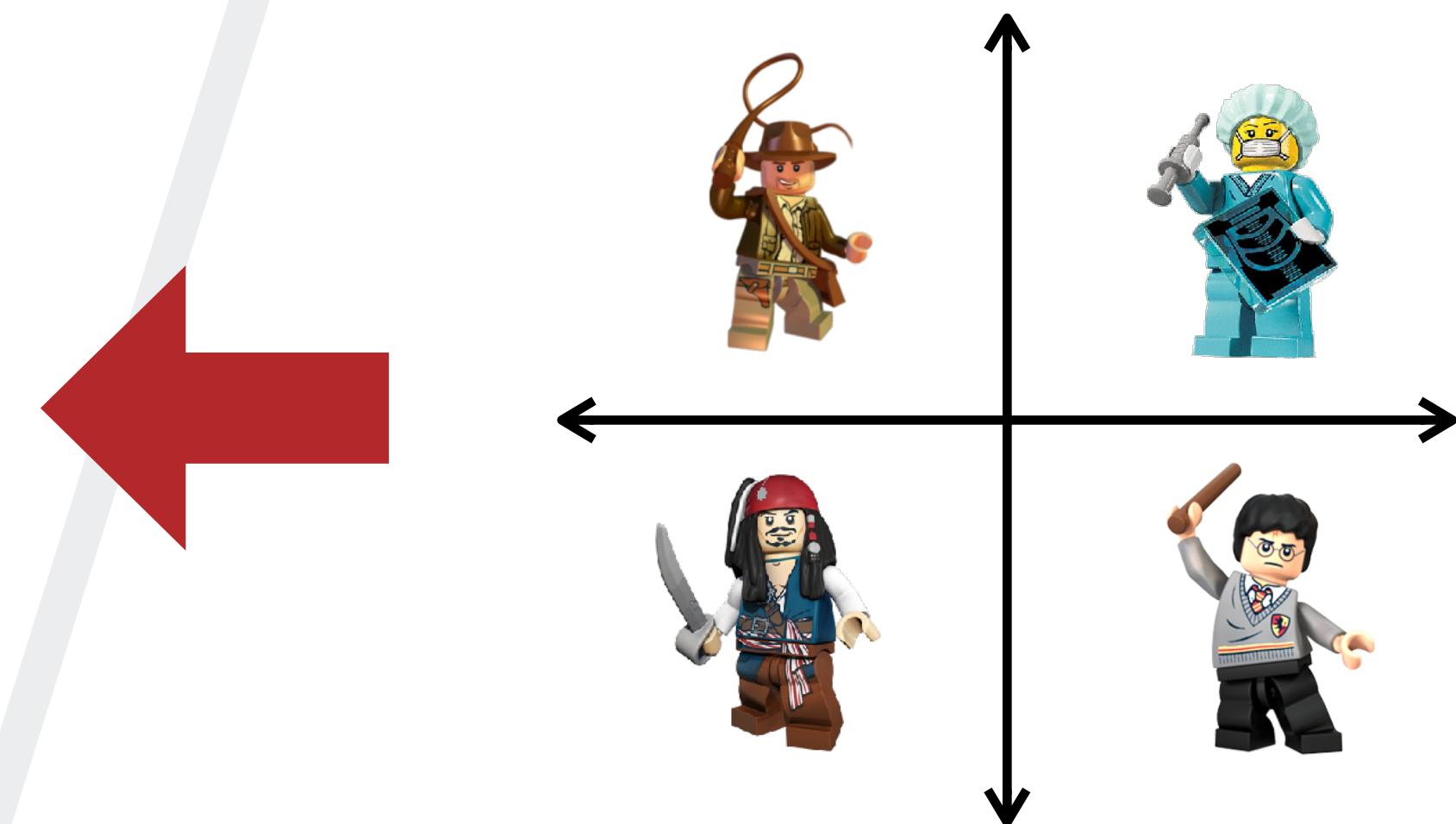
- Number of moving parts
- Number of seconds to complete
- Type safety
- How many team members can explain? Debug? Implement?
- Off the shelf parts vs. custom code
- Cost to run, to hire, to maintain, to store

Constraints & Control

17



How To Stay



- **Breadth: Lots of shallow solutions**
- **Evaluation criteria: include accuracy *and* simplicity**
- **Definitions: What is good enough?**
- **Engender serendipity**
- **Don't forget to get clever**

“I wouldn’t give a fig for the simplicity on the near side of complexity, but I would give my right arm for the simplicity, on the far side of complexity.”

Oliver Wendell Holmes, Sr.

Today's Plan



Purposefulness

Complexity

Resilience

Solution Reliability

System
Reliability



The Thing About Reliability

Golden Data Sets

Input	Output
a_0	a_0'
b_0	b_0'
c_0	c_0'
...	...

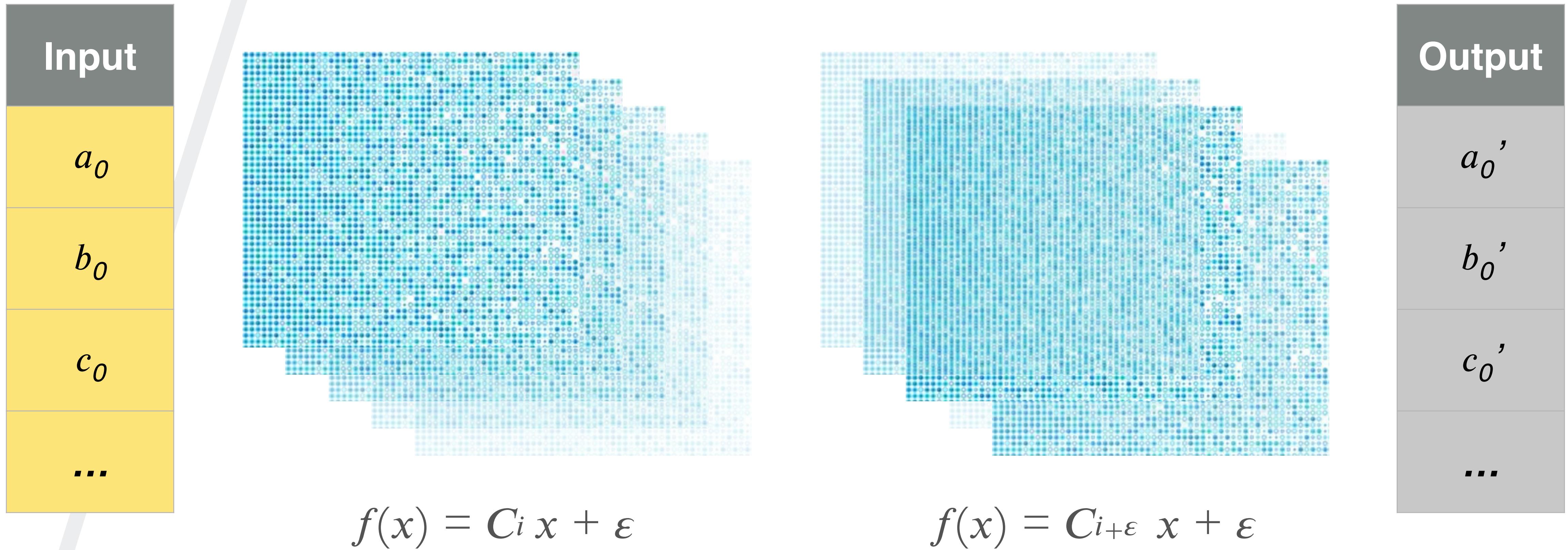
Input
a_0
b_0
c_0
...



$$f(x) = Cx + \varepsilon$$

Output
a_0'
b_0'
c_0'
...

Testing Harnesses



System Reliability

Solution Reliability



26

**Math Is Not
Negligent.**

**Math Don't
Care.**

Common Blind Spots

Software Engineers
Deep Learners
PhDs
Statisticians
Visualization Experts
Managers



Sampling
Rigorous Stats
Ship-It-ness
Unit Tests
Overfitting
P-hacking

Dynasties And Intellectual Inbreeding

Multiple-Objective Decision Analysis

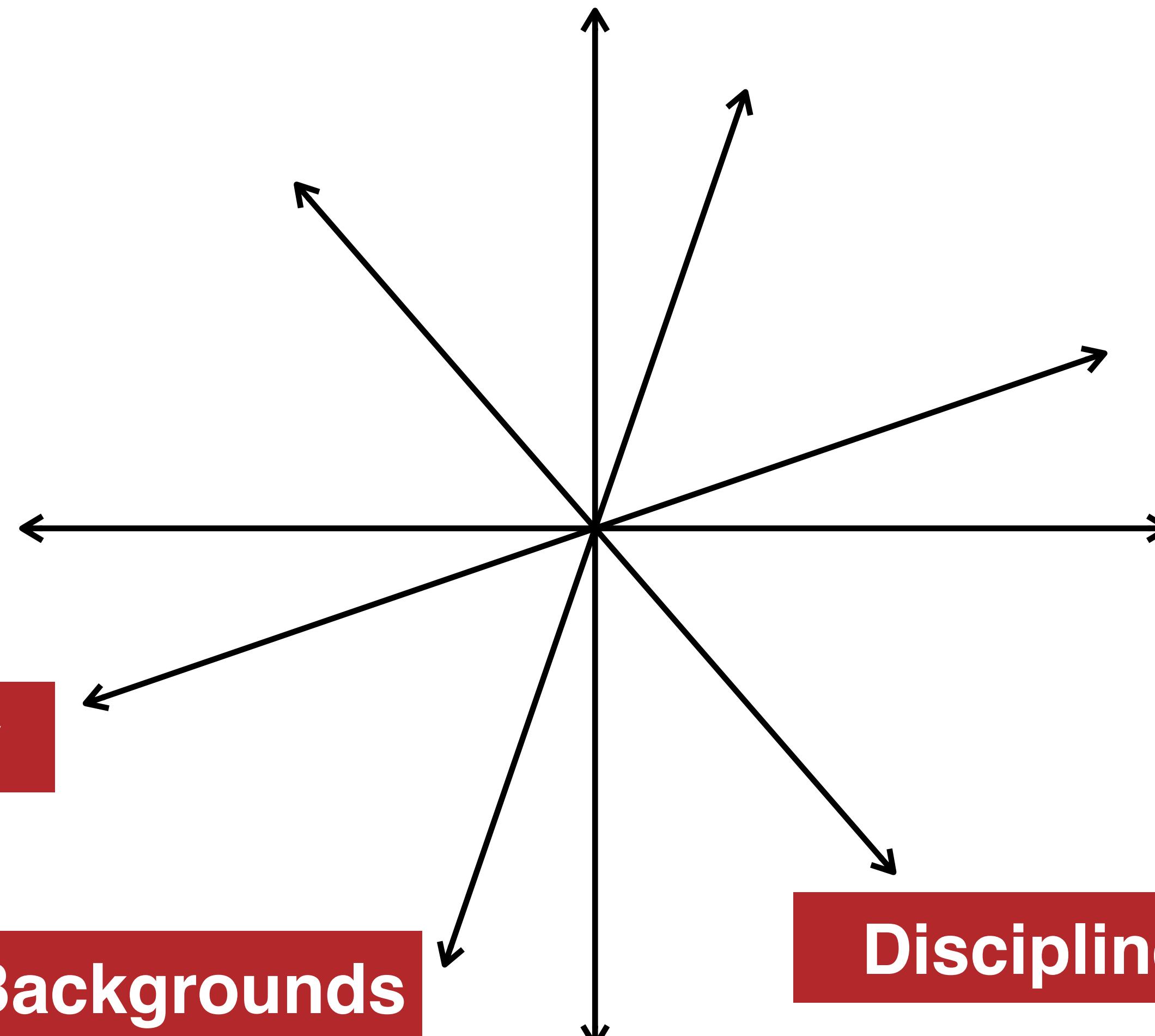
Perspectives

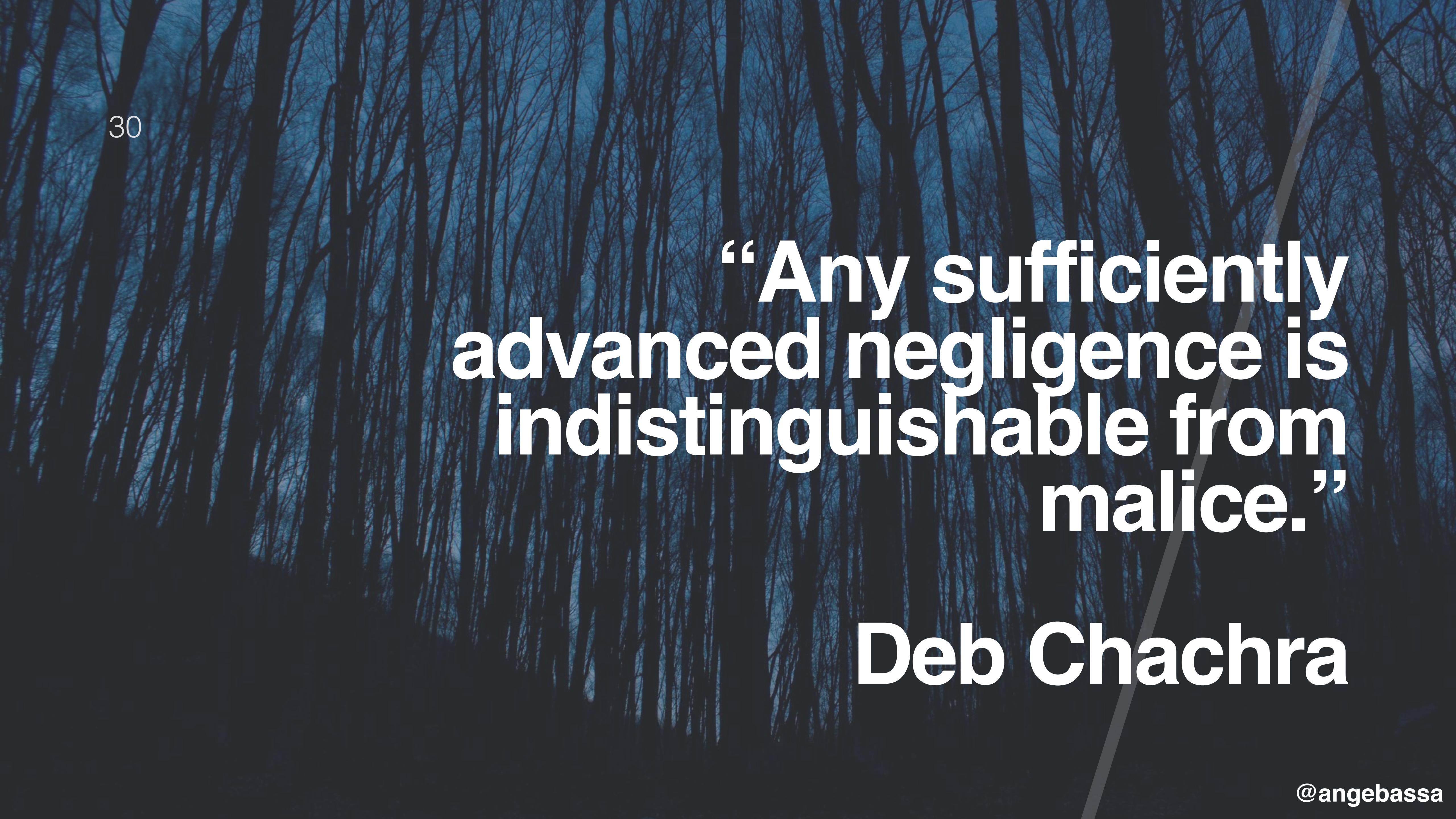
Seniority

Backgrounds

Languages

Disciplines





“Any sufficiently advanced negligence is indistinguishable from malice.”

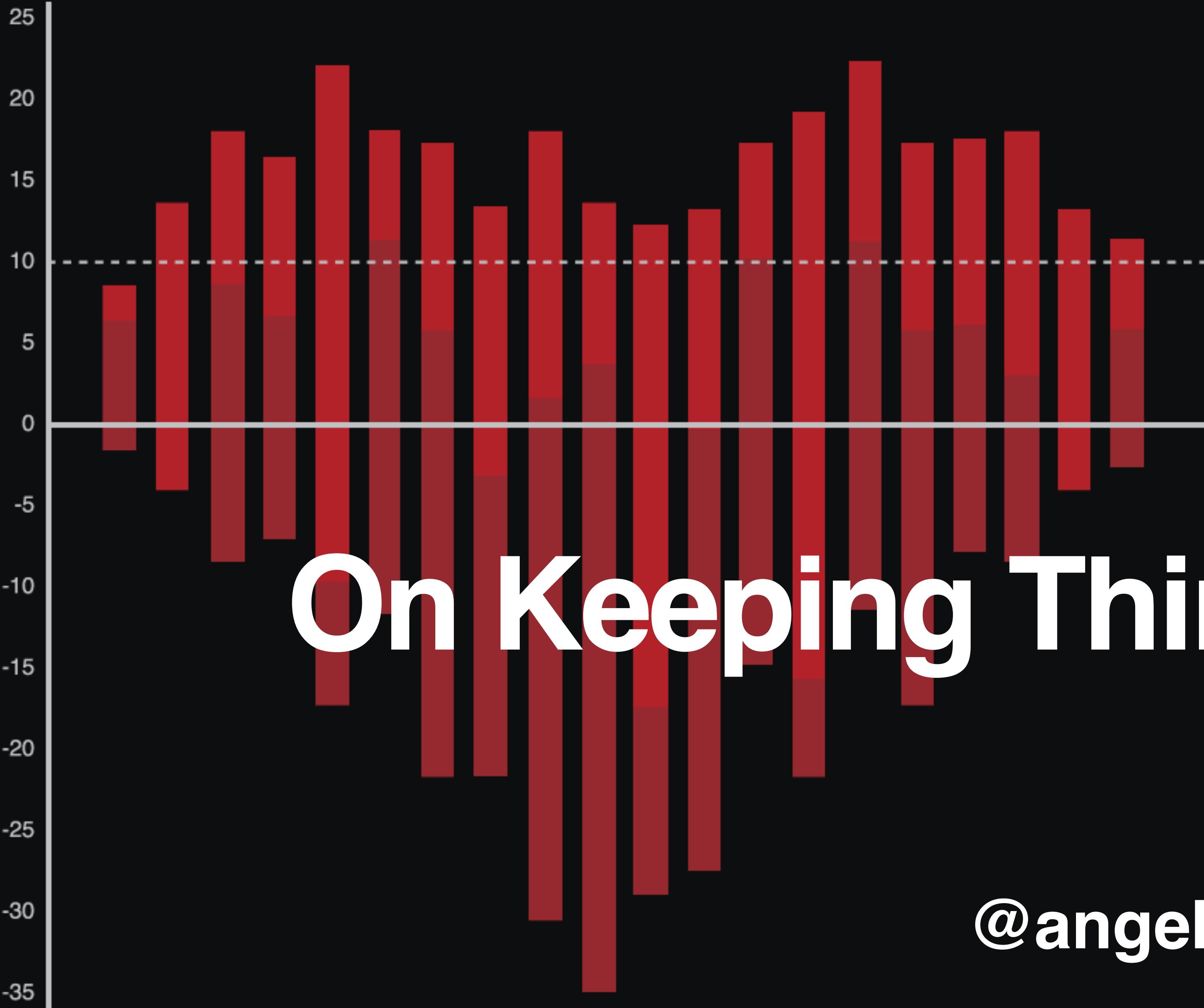
Deb Chachra

Questions?



©Dog Training Nation

@angebassa



On Keeping Things Simple

Angela Bassa
@angebassa, October 2017