Antipatterns on the road to AI

Luca Antiga

Online Tech Conference
- Italian edition -

23-24-25 Marzo, 2021


The view from my windshield


The view from my windshield

Introducing myself

Introducing myself

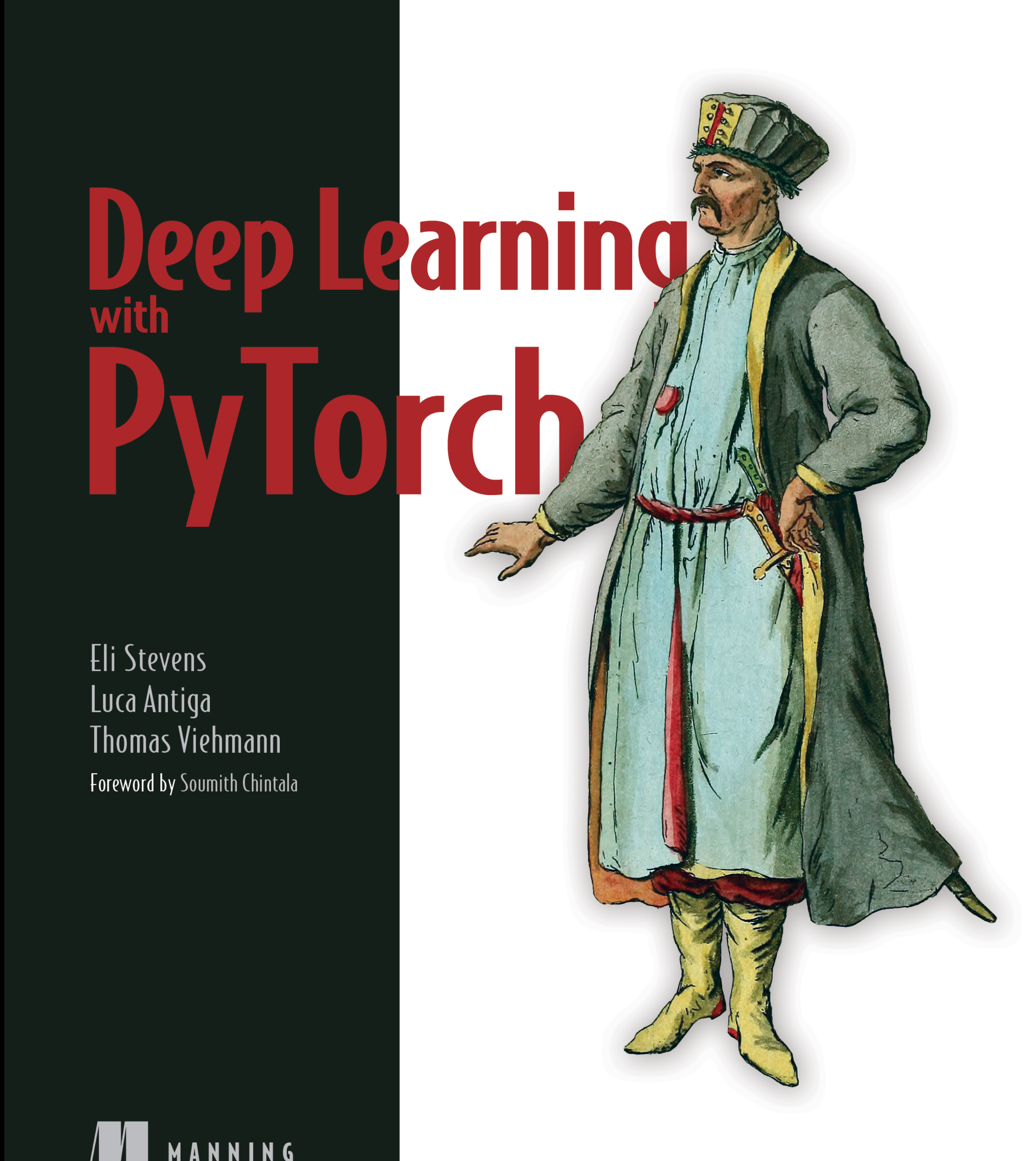
Bioengineer, medical imaging / modeling research 2000-2009

Co-founded Orobix (2009), Tensorwerk (2019), Orobix Life (2020)

PyTorch core contributor 2017-2018

Co-authored Deep Learning with PyTorch, Manning (2020)

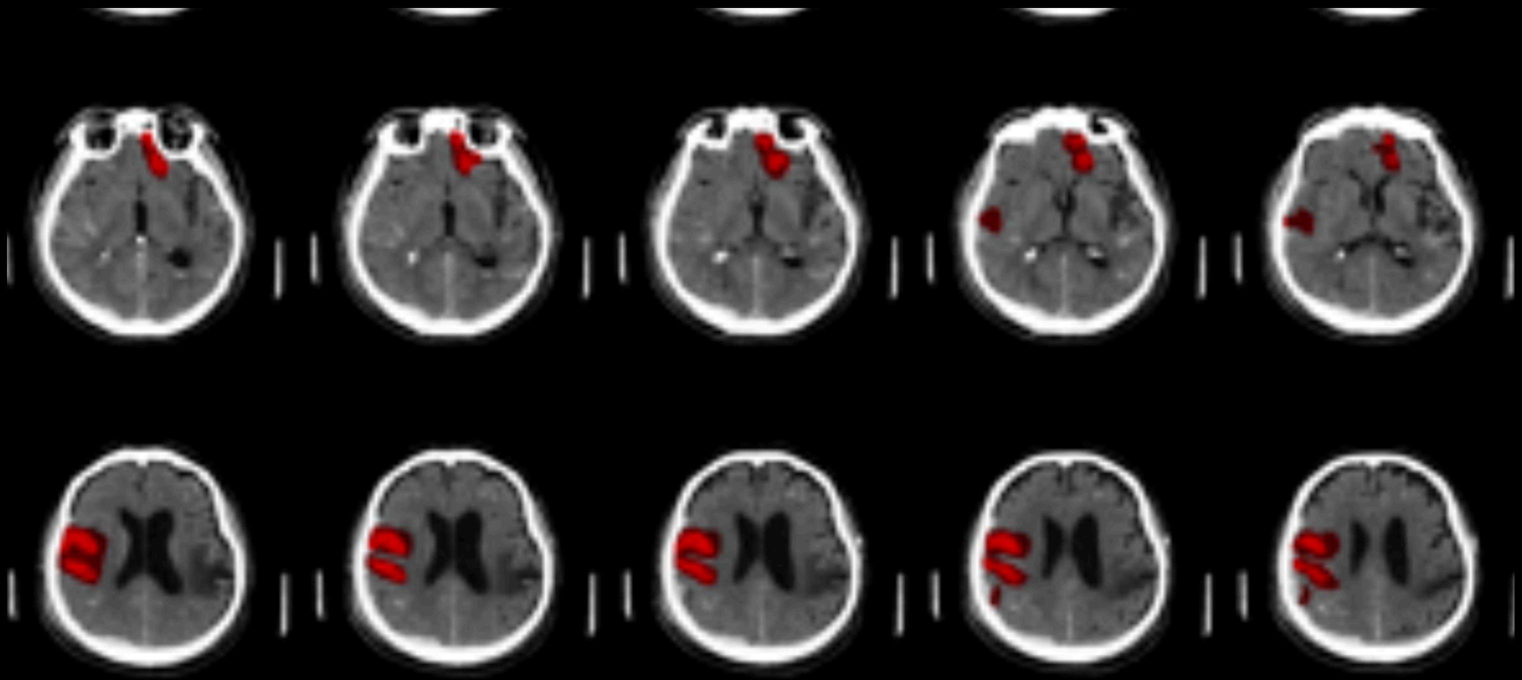
Co-created RedisAI

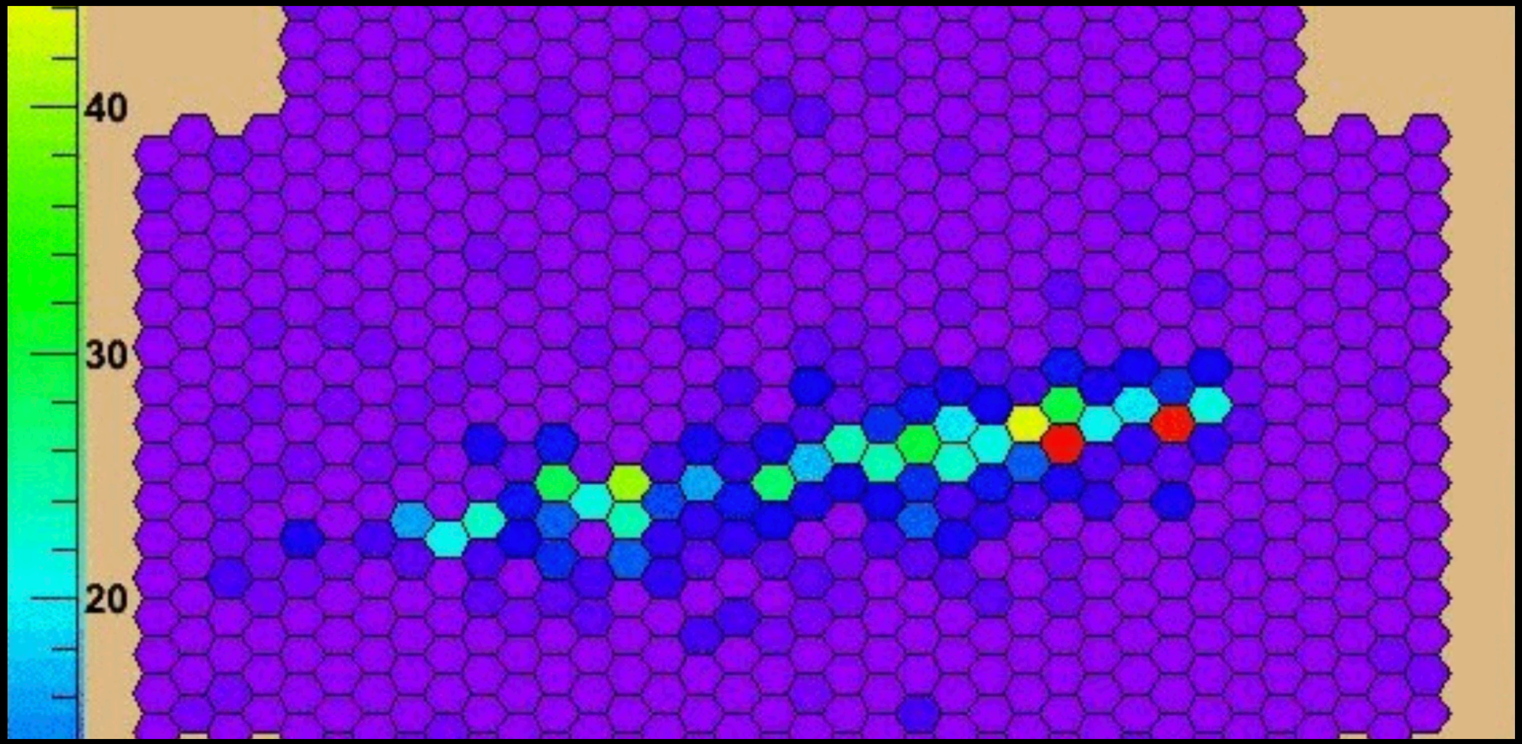


WHO WE ARE

The AI Service Company

At Oròbix, we implement and manage the lifecycle of artificial intelligence (AI) solutions. They can be newly designed or integrated into existing systems, spanning a wide range of industries from healthcare to manufacturing, from gaming to energy.




40
30
20


00:00.526
OAKLEY
19

It doesn't need to hurt

but it will naturally tend to

François Chollet
@fchollet

Deep learning excels at unlocking the creation of impressive early demos of new applications using very little development resources.

The part where it struggles is reaching the level of consistent usefulness and reliability required by production usage.

4:22 AM · Mar 20, 2021 · Twitter for Android

François Chollet
@fchollet

Replying to @fchollet

The reason why is that parametric models trained with gradient descent make it easy to automate something, but have little ability to deviate from the patterns they've learned. Meanwhile, the real world is full of surprises, and handling it requires the ability to adapt.

4:37 AM · Mar 20, 2021 · Twitter for Android

Hidden Technical Debt in Machine Learning Systems

D. Sculley, Gary Holt, Daniel Golovin, Eugene Davydov, Todd Phillips
{dsculley,gholt,dgg,edavydov,toddphillips}@google.com
Google, Inc.

Dietmar Ebner, Vinay Chaudhary, Michael Young, Jean-François Crespo, Dan Dennison
{ebner,vchaudhary,mwyoung,jfcrespo,dennison}@google.com
Google, Inc.

NIPS 2015

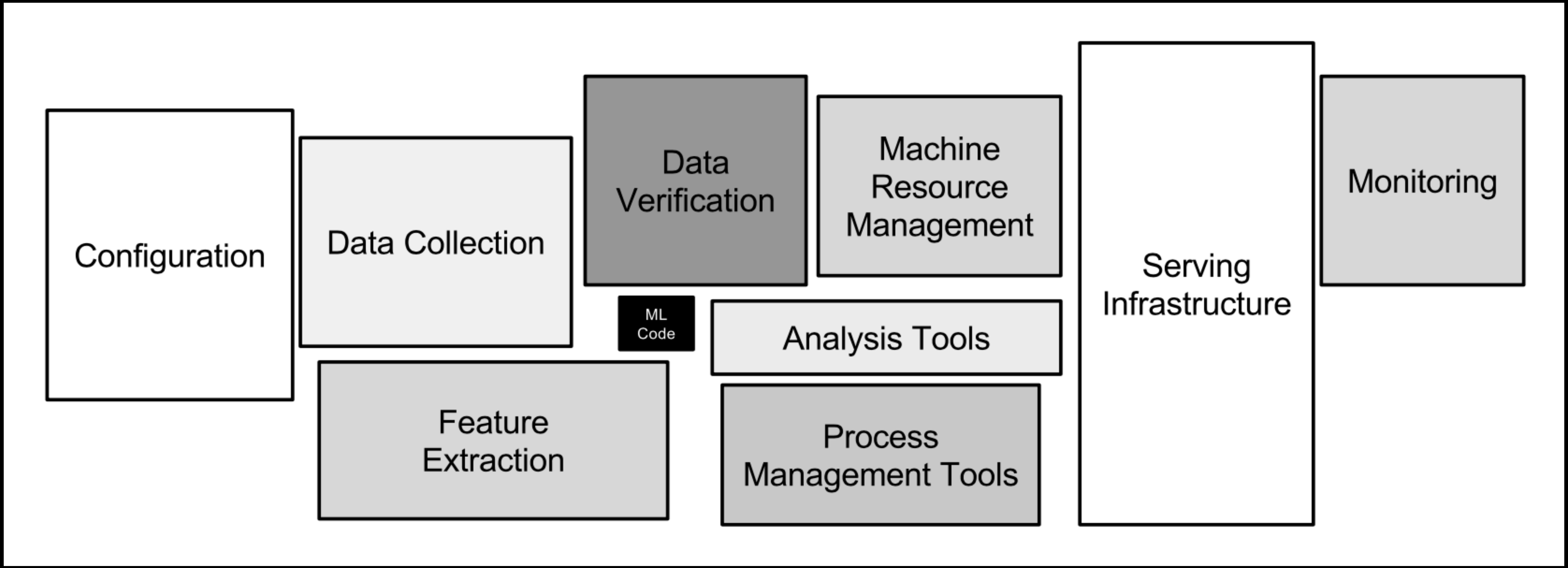
"... ML is required in [...] those cases when the desired behavior cannot be effectively expressed in software logic without dependency on external data."

"The real world does not fit into tidy encapsulation."

— D. Sculley et al. Hidden Technical Debt in Machine Learning Systems

Challenges

- Complex models erode boundaries
- Changing Anything Changes Everything (CACE)
- Data dependencies cost more than code dependencies
- Debt from glue code, pipeline jungles, dead experimental codepaths



"Only a small fraction of real-world ML systems is composed of ML code, as shown by the small black box in the middle. The required surrounding infrastructure is vast and complex."

ML is about process

Many antipatterns are about process

(1/6) The in(de)finite project trap

1. We'll start with the data we have
2. We'll look for something interesting in the data, find new correlations
3. We'll get back to the client in 2 months with the new insights

(1/6) The in(de)finite project trap

Antidotes

- Keep goals tight, identify a workable problem subspace
- Don't promise to come up with new correlations
- Keep stakeholders engaged at all times

(2/6) The expectation trap

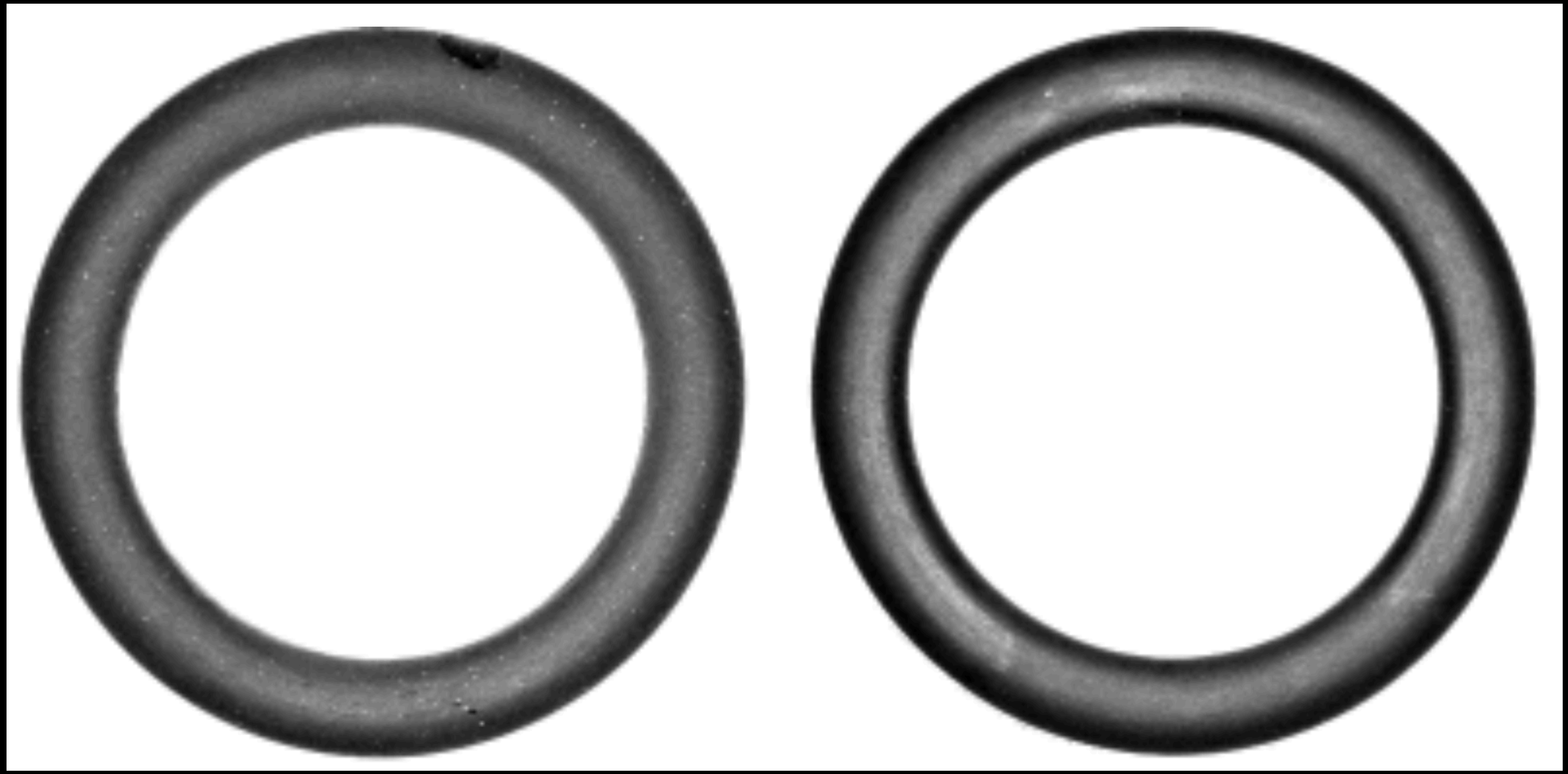
1. We were given lots of historical data
2. We promised 99% performance for this process
3. We'll get back to the client when we are close

(2/6) The expectation trap

Antidotes

- Historical data probably won't allow you to get there
- Need to sample the distribution in production
- Responsibility for performance is shared

Be careful where you sample from



Bias hides in the process

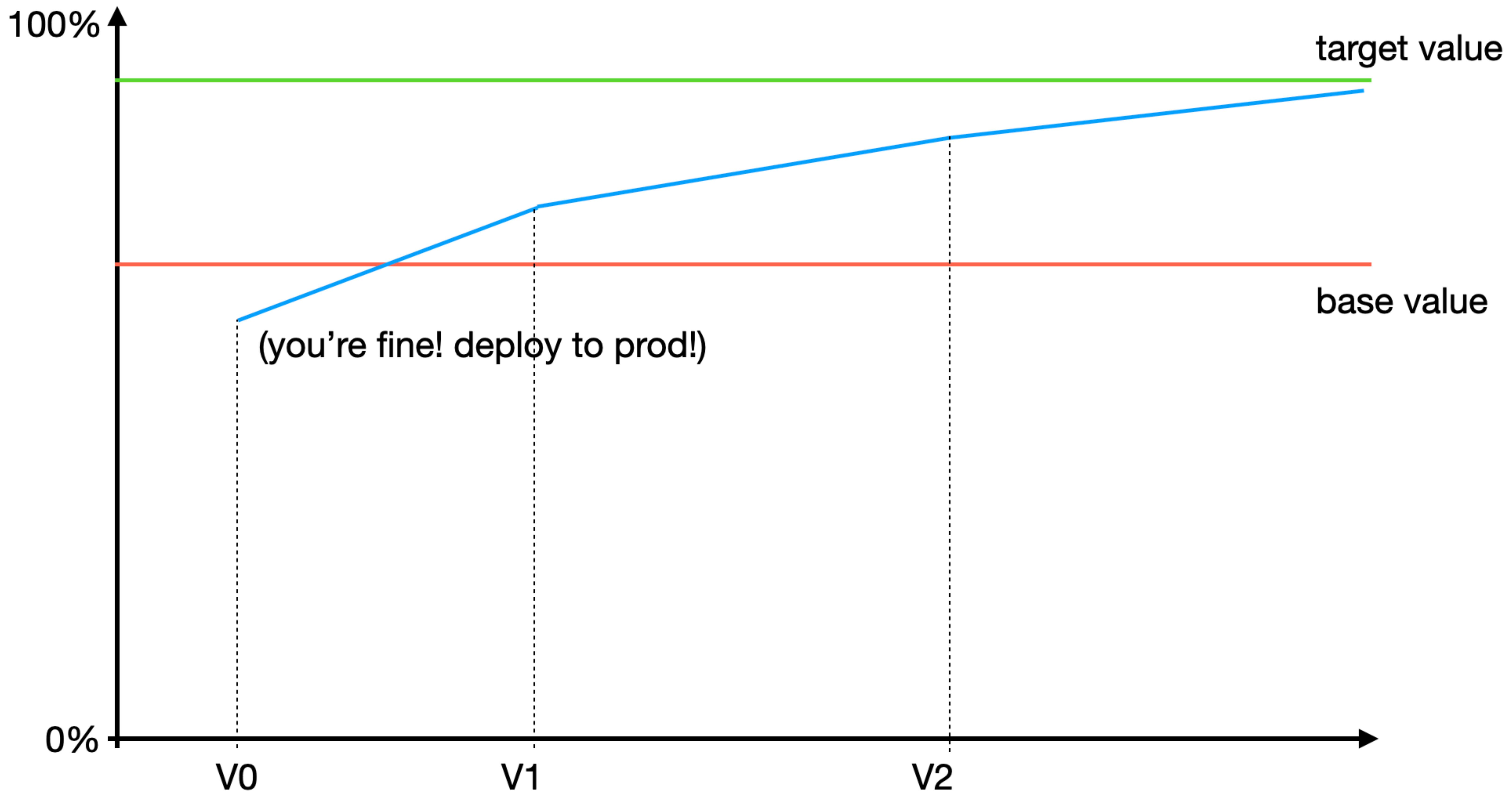
(3/6) Validation procastination

1. Performance is ok, but we're going to try with this other model to see if we can squeeze in some more accuracy
2. We'll push the test with the client by one month to see if we can reach the performance we promised

(3/6) Validation procastination

Antidotes

- Delaying measuring is a form of escapism
- By measuring partial success you can demonstrate improvement
- Validation data is the one you should spend more time curating


100%
target value
base value
(you're fine! deploy to prod!)
0%
V0
V1
V2

(4/6) Delayed deployment

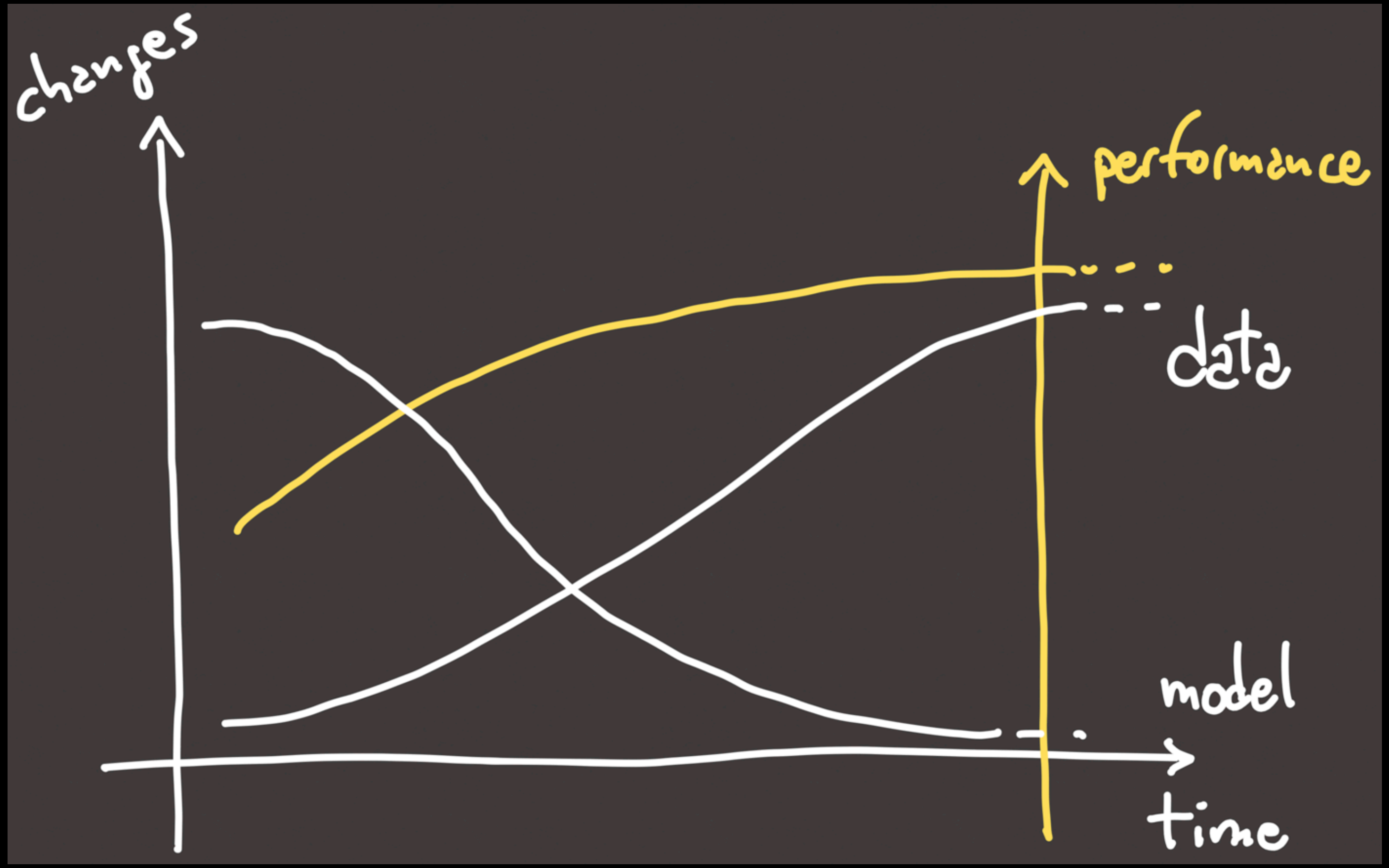
1. We were given lots of historical data
2. We promised 99% performance for this process
3. When the system reaches 99% on validation, we'll deploy to production

(4/6) Delayed deployment

Antidotes

- Integrate into production pipelines as early as possible
- Real challenge starts when system is integrated, no point in delaying
- Favor tools that allow to go to production sooner, eliminate friction

Many real-world applications are data-variant



Plan for change

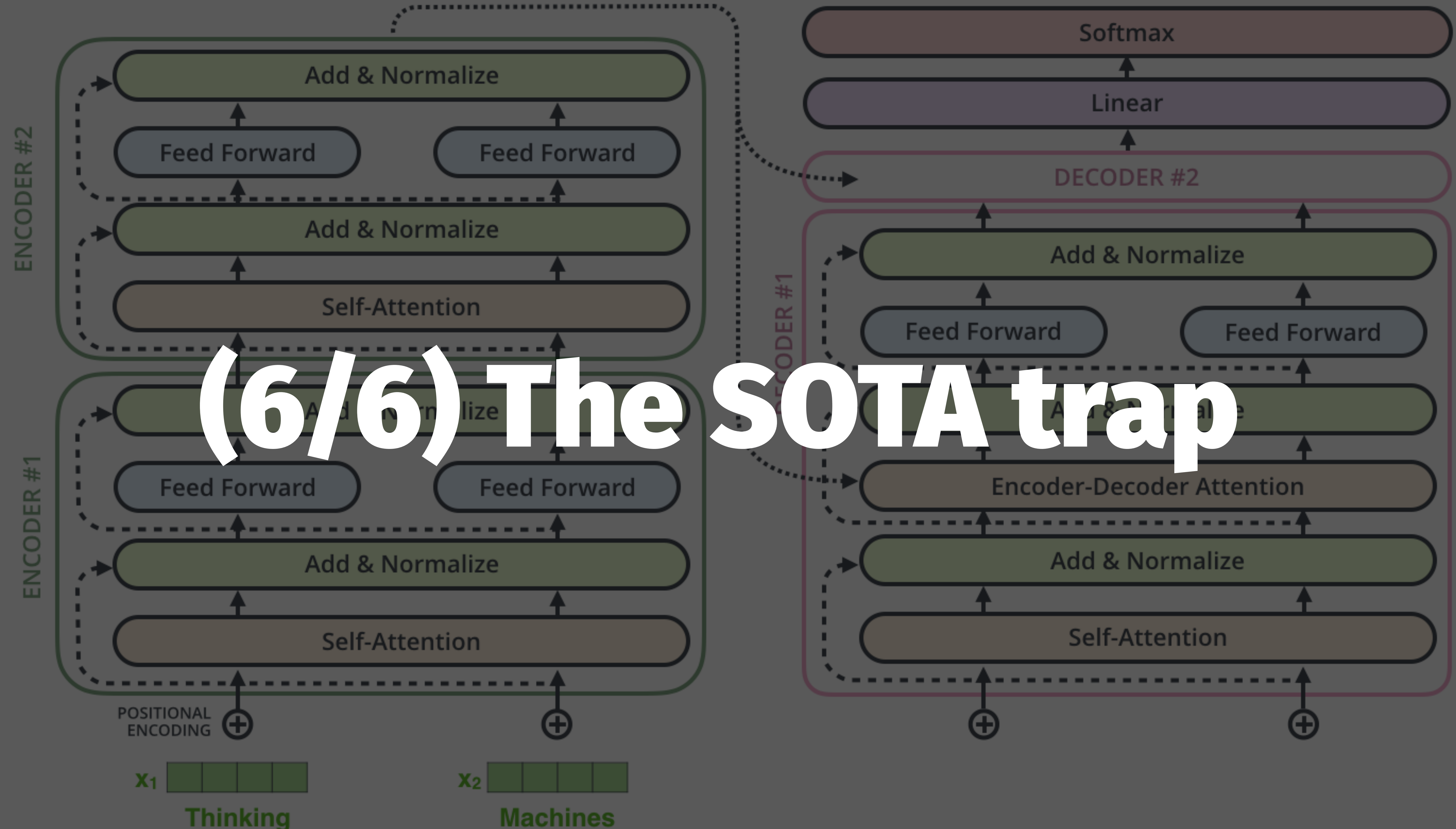
(5/6) The tooling trap

(5/6) The tooling trap

Antidotes

- Easy to get enthusiastic and cargo cult
- Tooling helps with process if process is clear
- Establish minimal process over minimal tooling
- **A great tool is a tool that disappears**

(6/6) The SOTA trap



(6/6) The SOTA trap

Antidotes

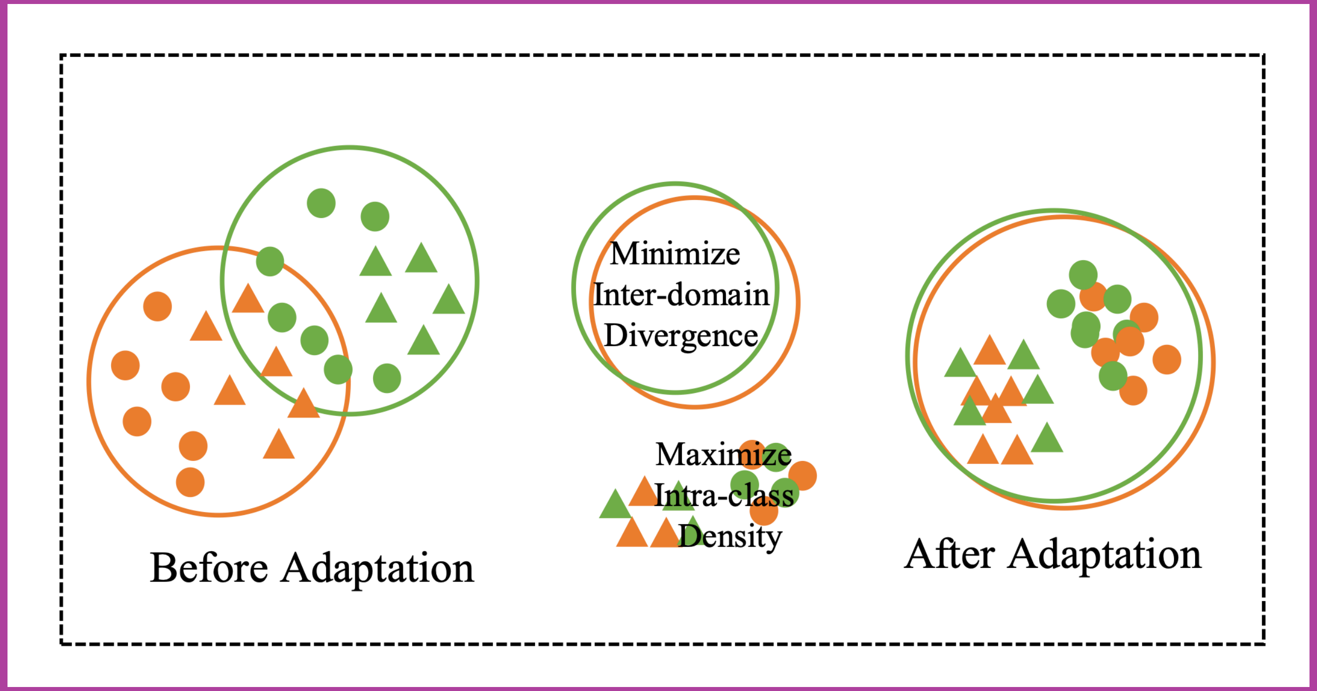
- Your goal is solving the problem, not pleasing your peers
- It's first about the data (and the goal), then the model
- Therefore it's probably about the process

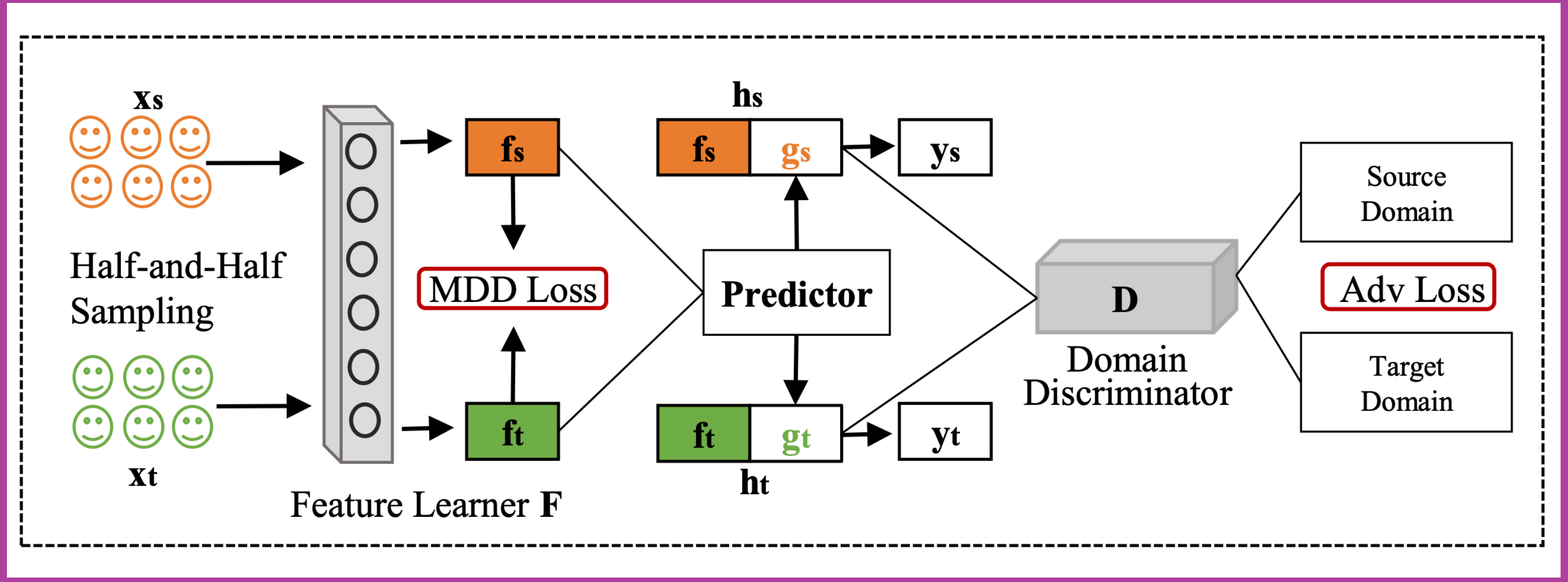
Current strategies @ Orobix

- We have a production strategy from the start
- We emphasize monitoring and governance
- We write models predictably (PyTorch Lightning)
- We focus on a small family of models at a time

Current R&D focus @ Orobix

Model drift, domain adaptation, self-supervised learning

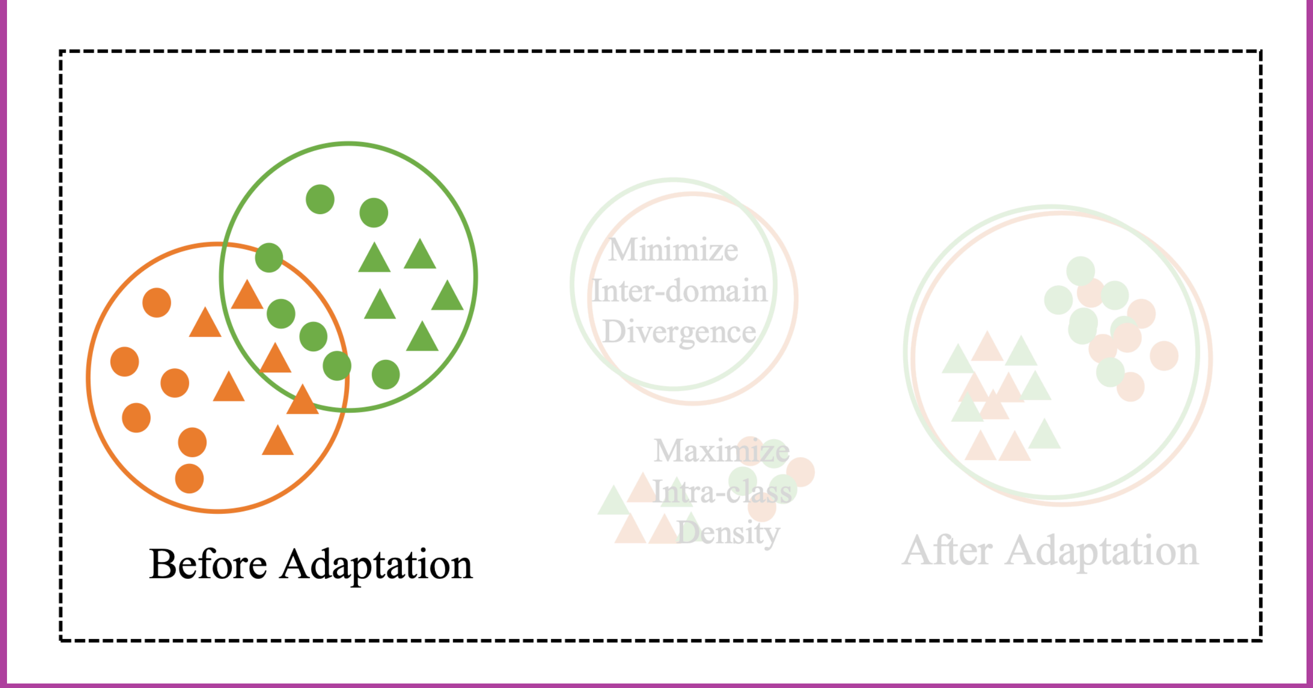




https://github.com/orobix/mdd-domain-adaptation

Model drift

New library for data and concept drift for Pytorch




TorchDrift

TorchDrift: drift detection for PyTorch ¶

TorchDrift is a data and concept drift library for PyTorch. It lets you monitor your PyTorch models to see if they operate within spec.

We focus on practical application and strive to seamlessly integrate with PyTorch.

Get started:

Installation

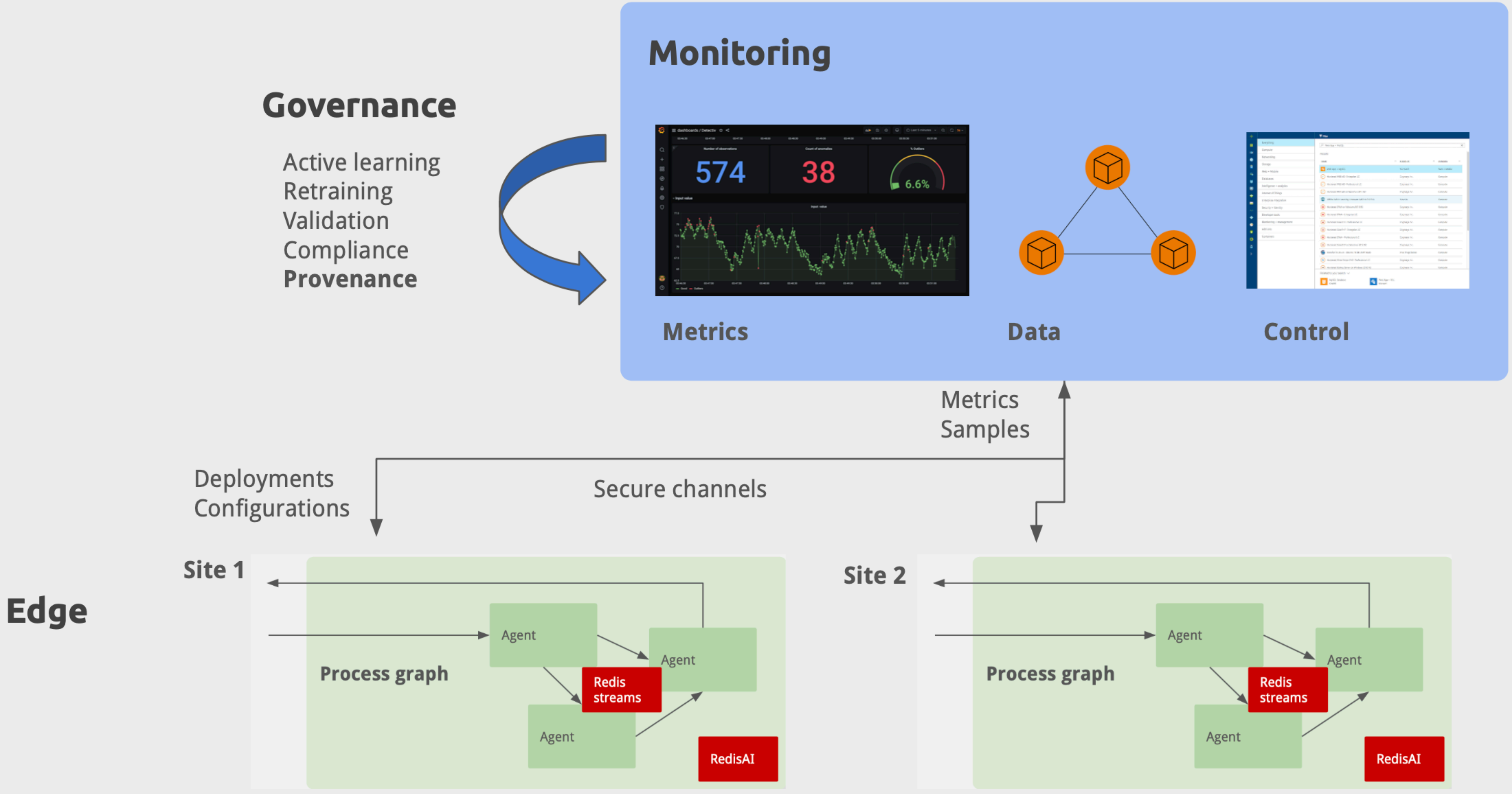
Examples:

Drift detection on image classifiers
- Load data
- Build a model
- Simulating drifted data

https://torchdrift.org

Orobix + MathInf GmbH

invariant.ai®



Thank you

@lantiga