

What can we learn from a billion agents?

Ayush Chopra
ayushc@mit.edu



AgentTorch



EQUIVALENT TO
9 LARGE EGGS
PER CARTON

108⁵⁰
FT
PER HEN
IN
ROTATED PASTURES

The Vital Farms logo features the word "Vital" in a large, white, cursive script font inside a red heart shape. Below it, the word "Farms" is written in a smaller, white, sans-serif font.

EXTENDED BY HAND & SMALL FAMILY FARMS

PASTURE-RAISED

108^{SD}
FT
PER LINE
100%
PICTURES

The Vital Farms logo is located in the bottom right corner. It features the word "Vital" in a large, white, cursive script font, with "Farms" in a smaller, white, sans-serif font directly below it. Above "Vital", there is a small, semi-transparent circular badge with the words "EST. 2001".

TENDED BY HAND & SMALL FAMILY FARMS
BASILE LURE RAISINS

10.99





Eggs so expensive? Blame bird migration, not inflation...

The H5N1 Crisis



Wild Birds

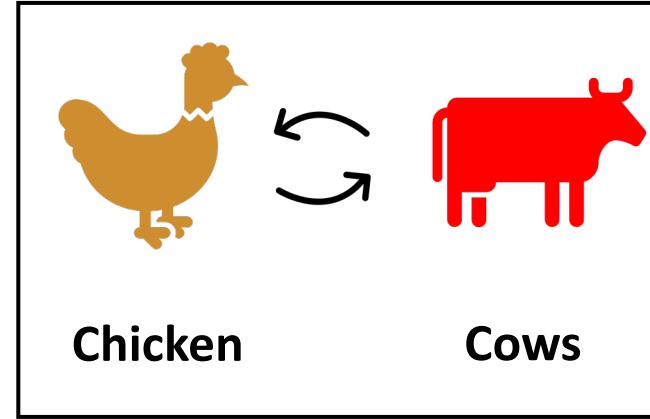
The H5N1 Crisis



Wild Birds



farms

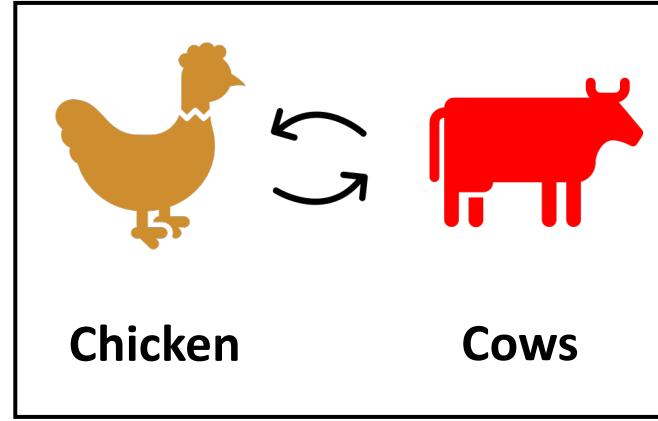


The H5N1 Crisis

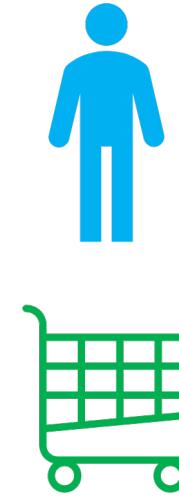


Wild Birds

→
farms



virus
→
food →



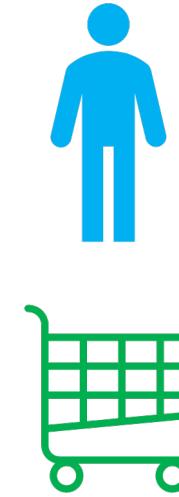
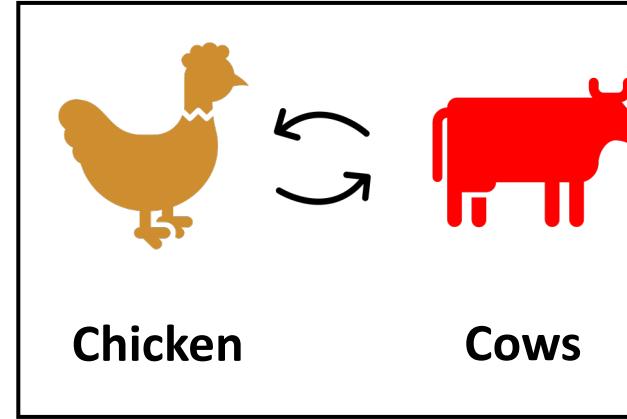
Humans
Grocery
Supply

The H5N1 Crisis



Wild Birds

→
farms



Humans



Grocery
Supply

The H5N1 Crisis

100 million
chicken culled

\$4 billion
revenue loss

Human-human
infections!





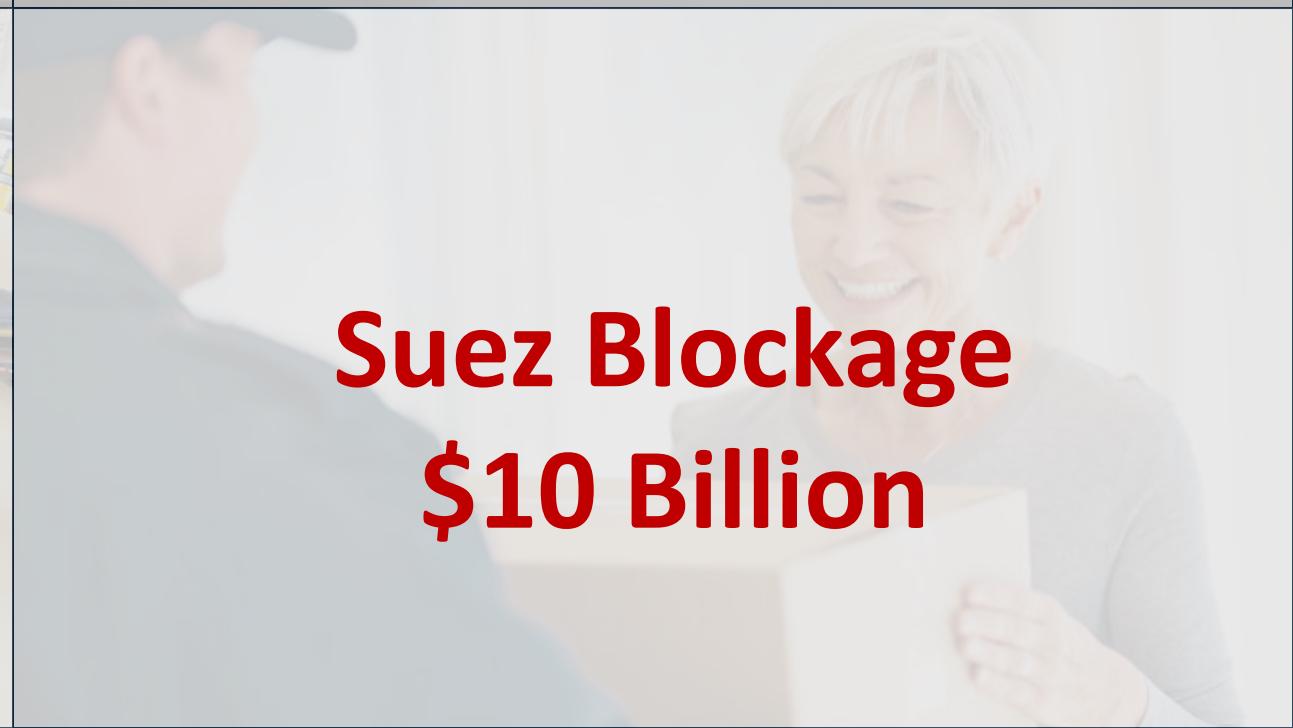
**COVID
\$3.5 Trillion**



**CrowdStrike
\$1.7 Billion**



**Hurricane Ida
\$80 Billion**



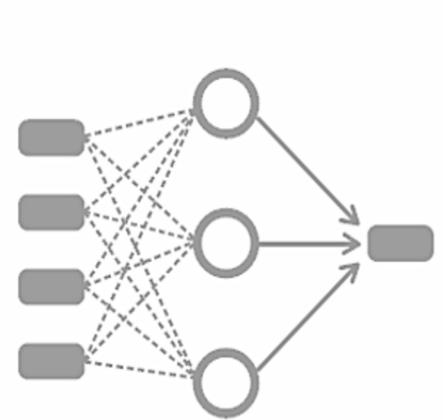
**Suez Blockage
\$10 Billion**

Imagine if... God's Eye View



Capture

Bird Migration
Disease Transmission



Analyze

Farmer Incentives
Supply Disruptions



Act

Stop Spread
Control Prices

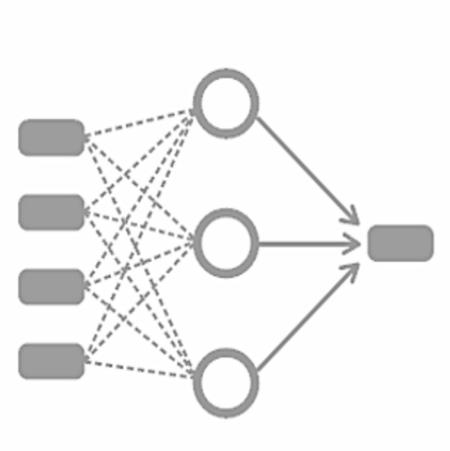
God's Eye View is Not Trivial



Capture

Bird Migration
Disease Transmission

Web of Interconnections



Analyze

Farmer Incentives
Supply Disruptions

Heterogeneous Behaviors



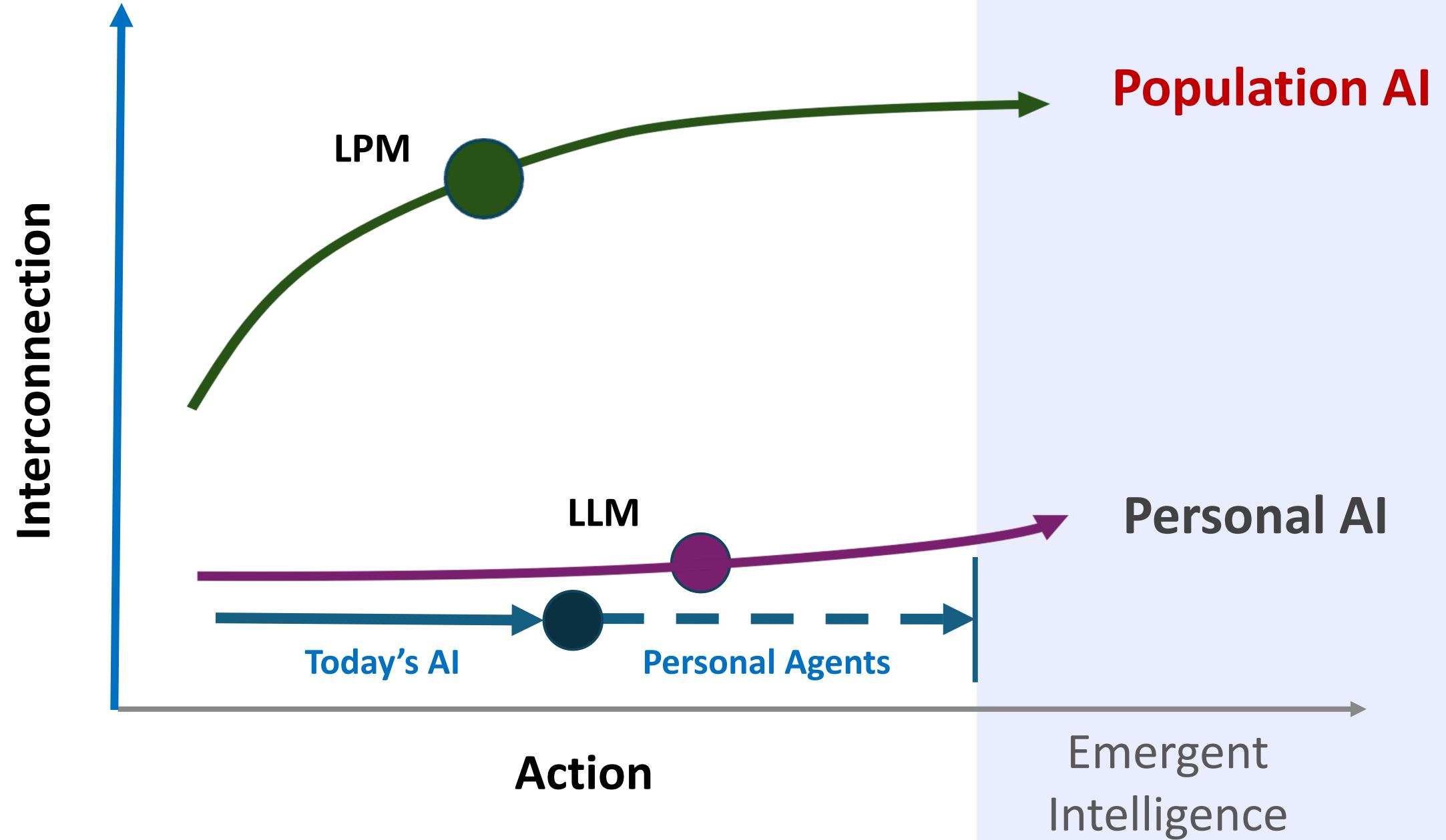
Act

Stop Spread
Control Prices

Multi-scale Decisions

Large Language Population Models

It's not *just* about smarter AI individuals, but grasping our interconnected world





Large Population Models



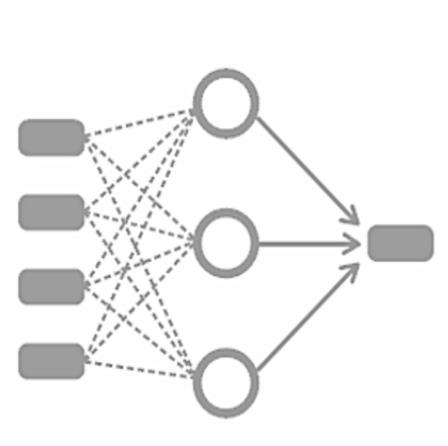
LPMs: Building the God's Eye View



Capture

Bird Migration
Disease Transmission

Web of Interconnections



Analyze

Farmer Incentives
Supply Disruptions

Heterogeneous Data



Act

Stop Spread
Control Prices

Multi-scale Decisions

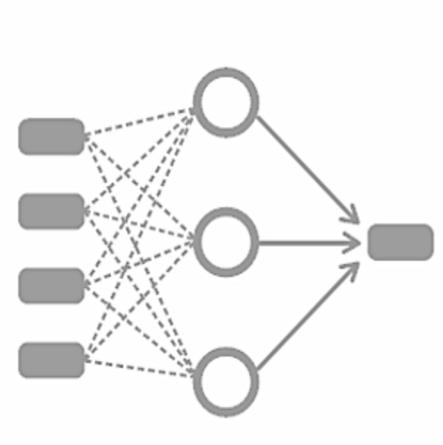
LPM: Technical Pillars



Capture

Multi-scale Dynamics
Stochastic Protocols

Compositional
Simulations



Analyze

Multi-modal data
Adaptive behavior

Differentiable
Learning



Act

Multi-objective Decisions
Real-time Response

Decentralized
Analysis

LPM: Technical Pillars

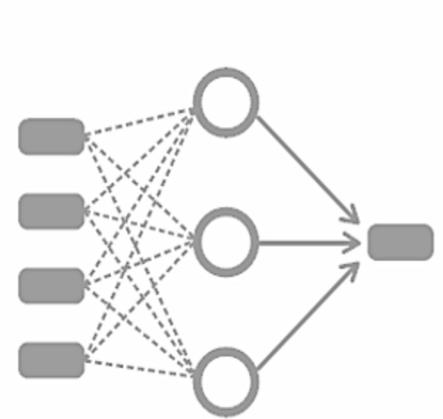


Capture

Multi-scale Dynamics
Stochastic Protocols

Compositional Simulations

AAMAS'23, AAMAS'24,
Nature Medicine'25



Analyze

Multi-modal data
Adaptive behavior

Differentiable Learning

AAMAS'24 (Best Paper Runner-Up),
AAMAS'23



Act

Multi-objective Decisions
Real-time Response

Decentralized Analysis

British Medical Journal'21,
AAMAS'24, ICML-W'22 (Best Paper)

LPM: Simulate a country on your laptop

Execute 300,000 interactions/sec and scale to 60 million agents/GPU

Method	Simulation	Calibration	Analysis
Conventional ABM*	50 hours	100,000 hours	5,000 hours
LPM	5 minutes	20 minutes	10 seconds

+ 600x **+ 3000x** **+ 5000x**



AgentTorch

Large Population Models: Reaching millions around the world

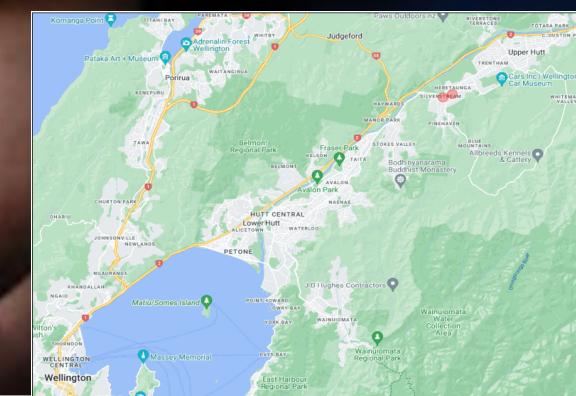


SIMCITY™ BUILD IT!

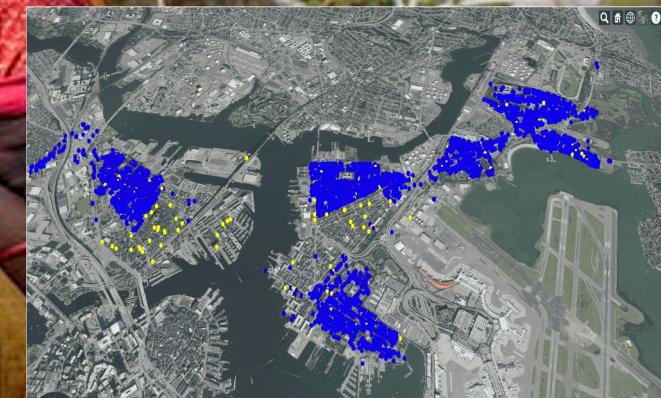




5 million citizens

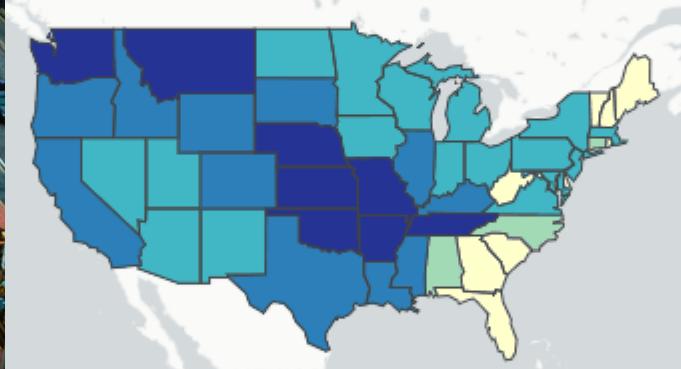


30 million households



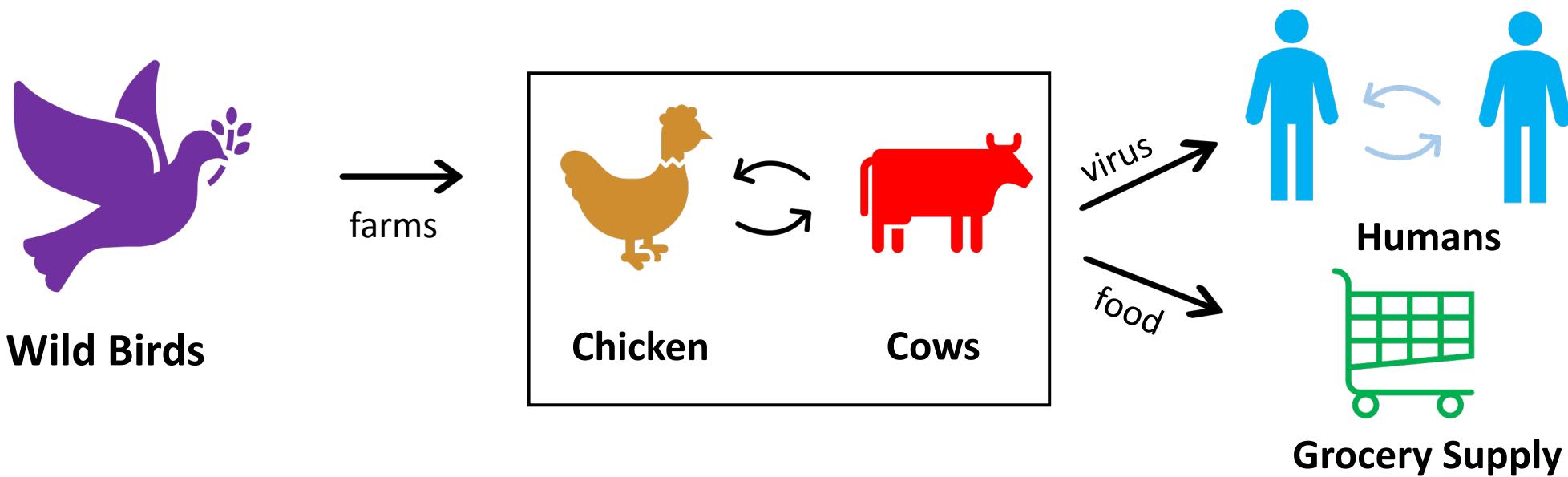


\$ billion supply chains

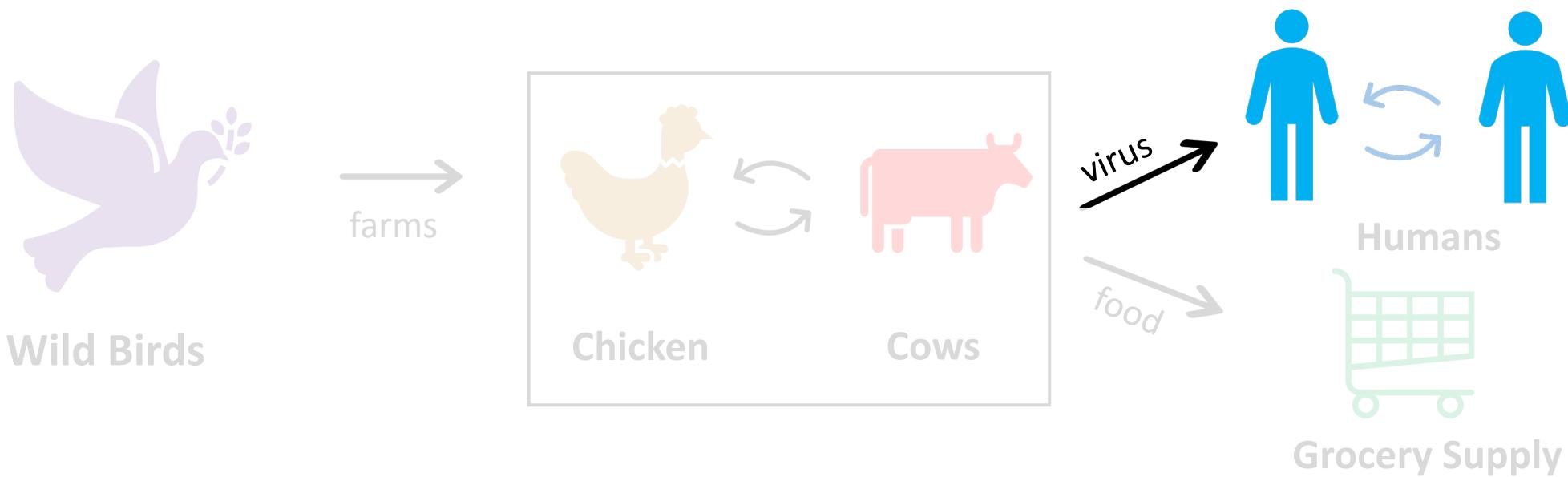




The H5N1 Crisis

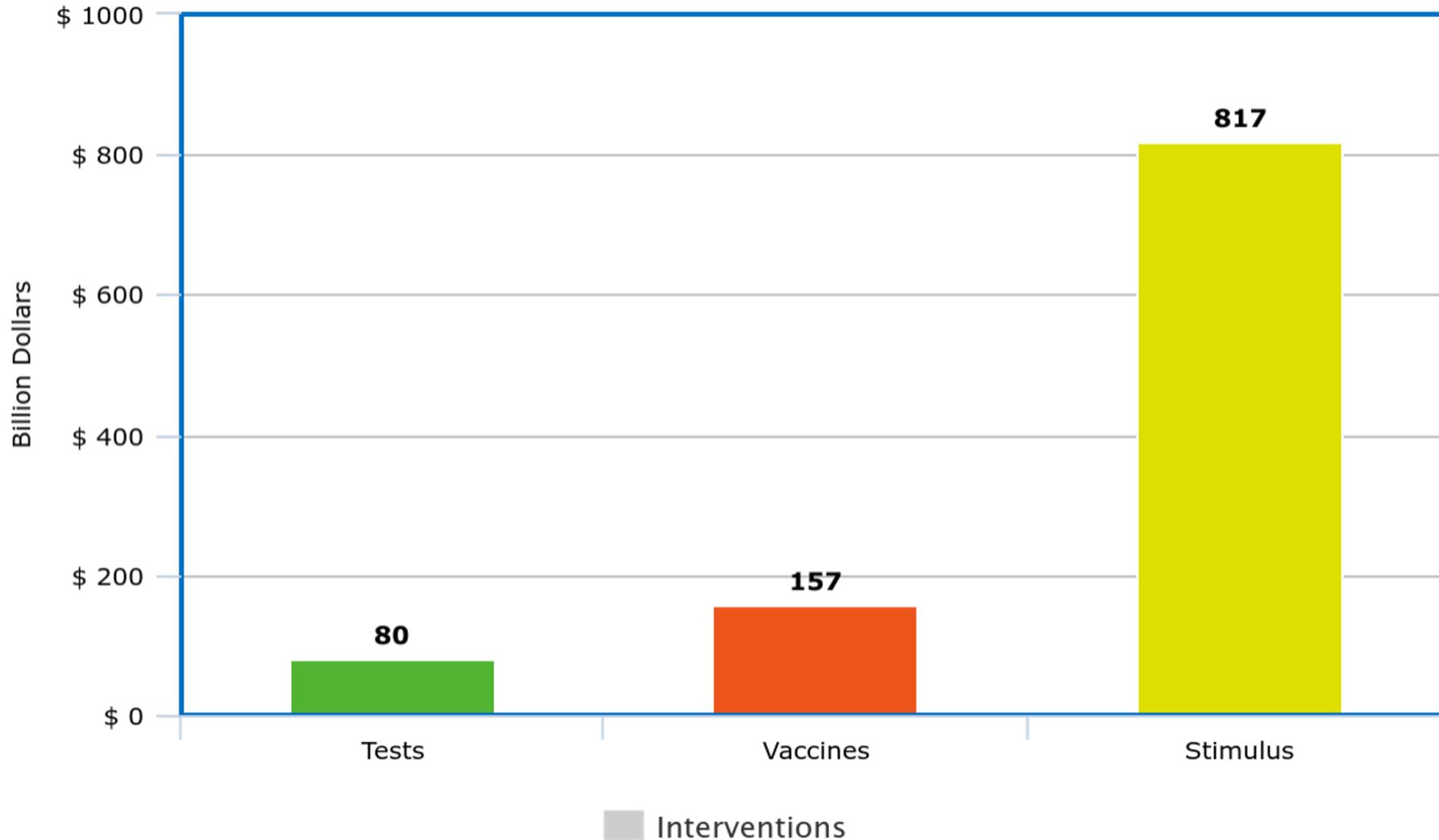


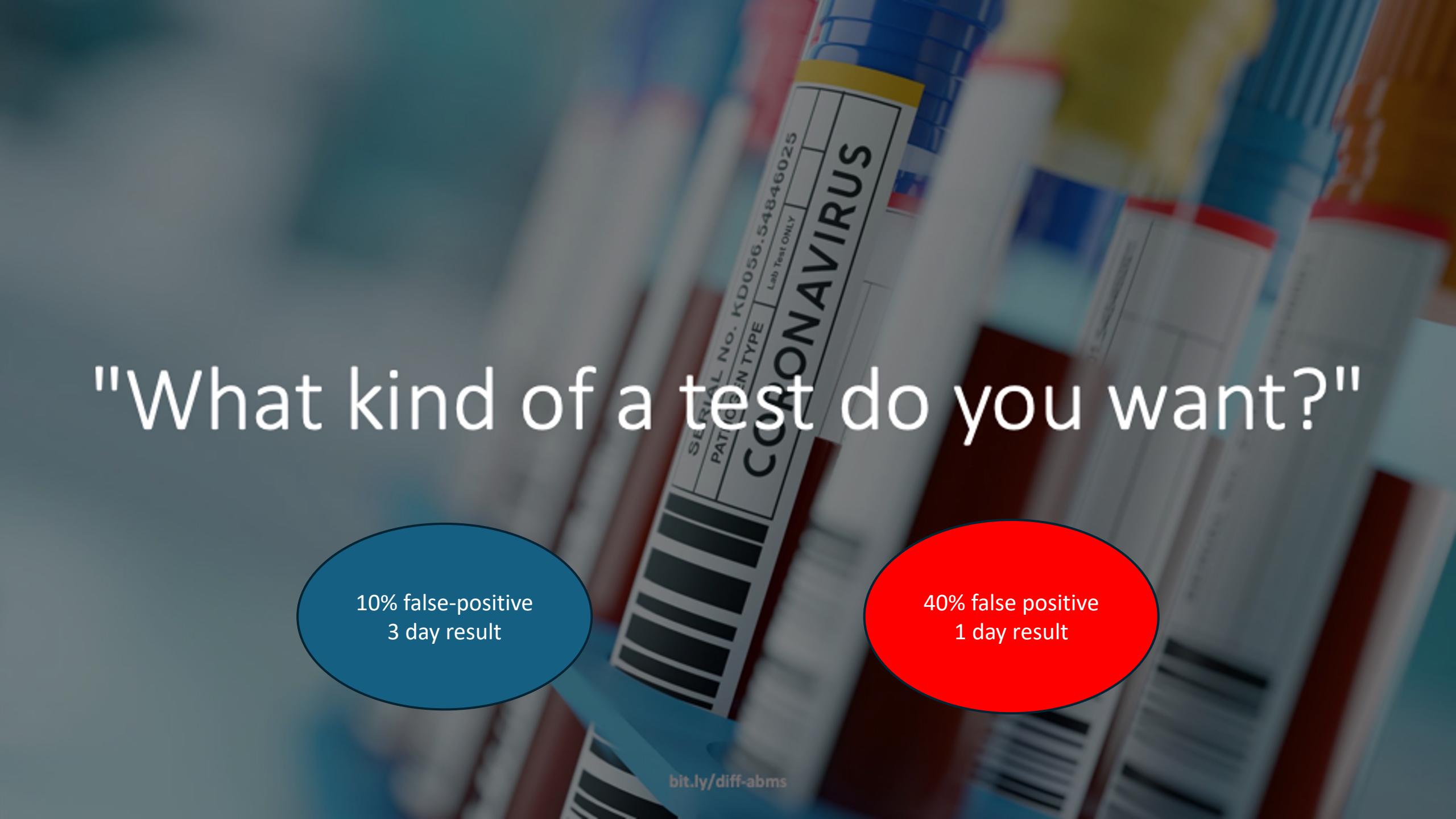
Human Disease Transmission?





Trillion-Dollar Pandemic





"What kind of a test do you want?"

10% false-positive
3 day result

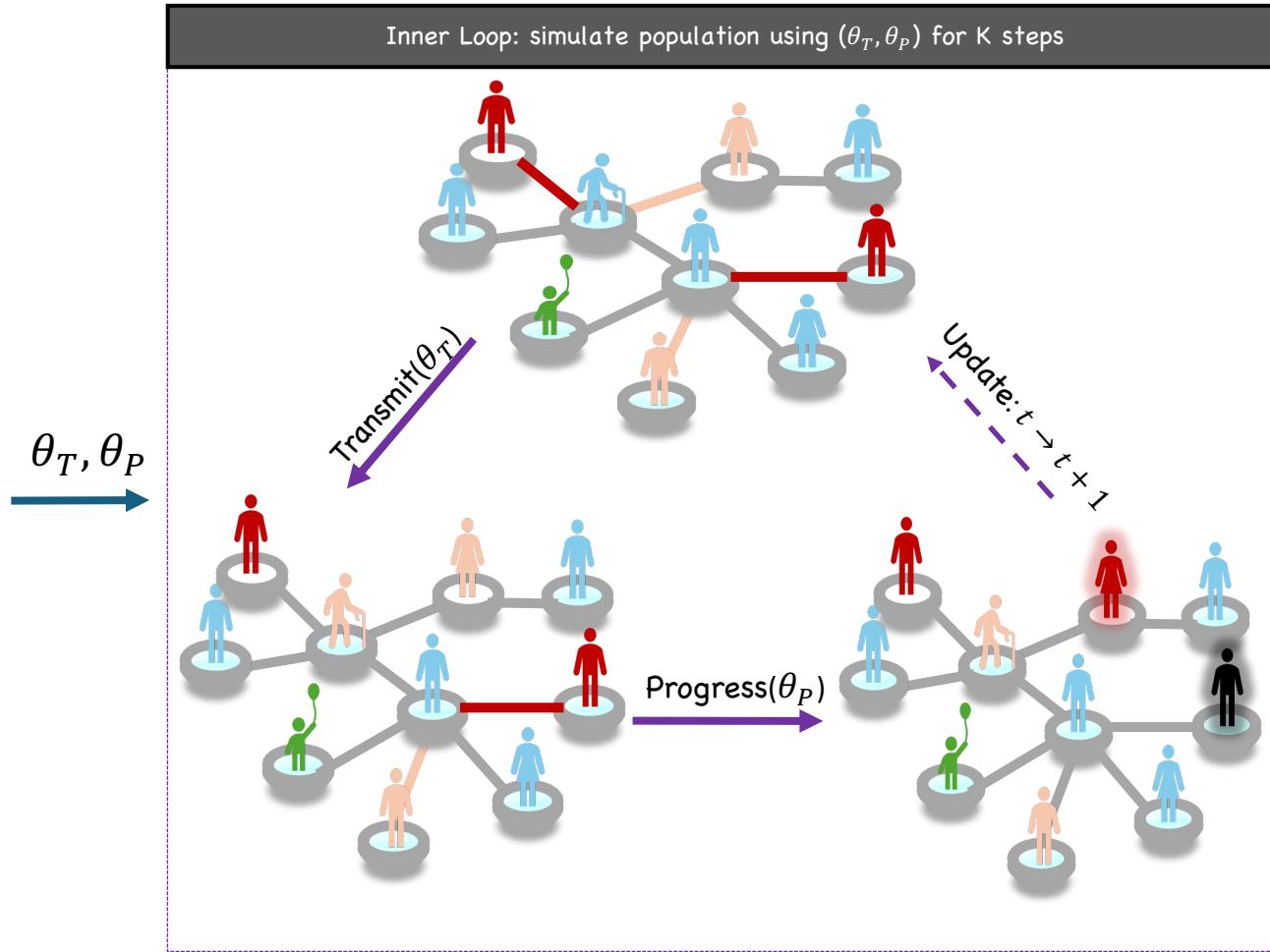
40% false positive
1 day result

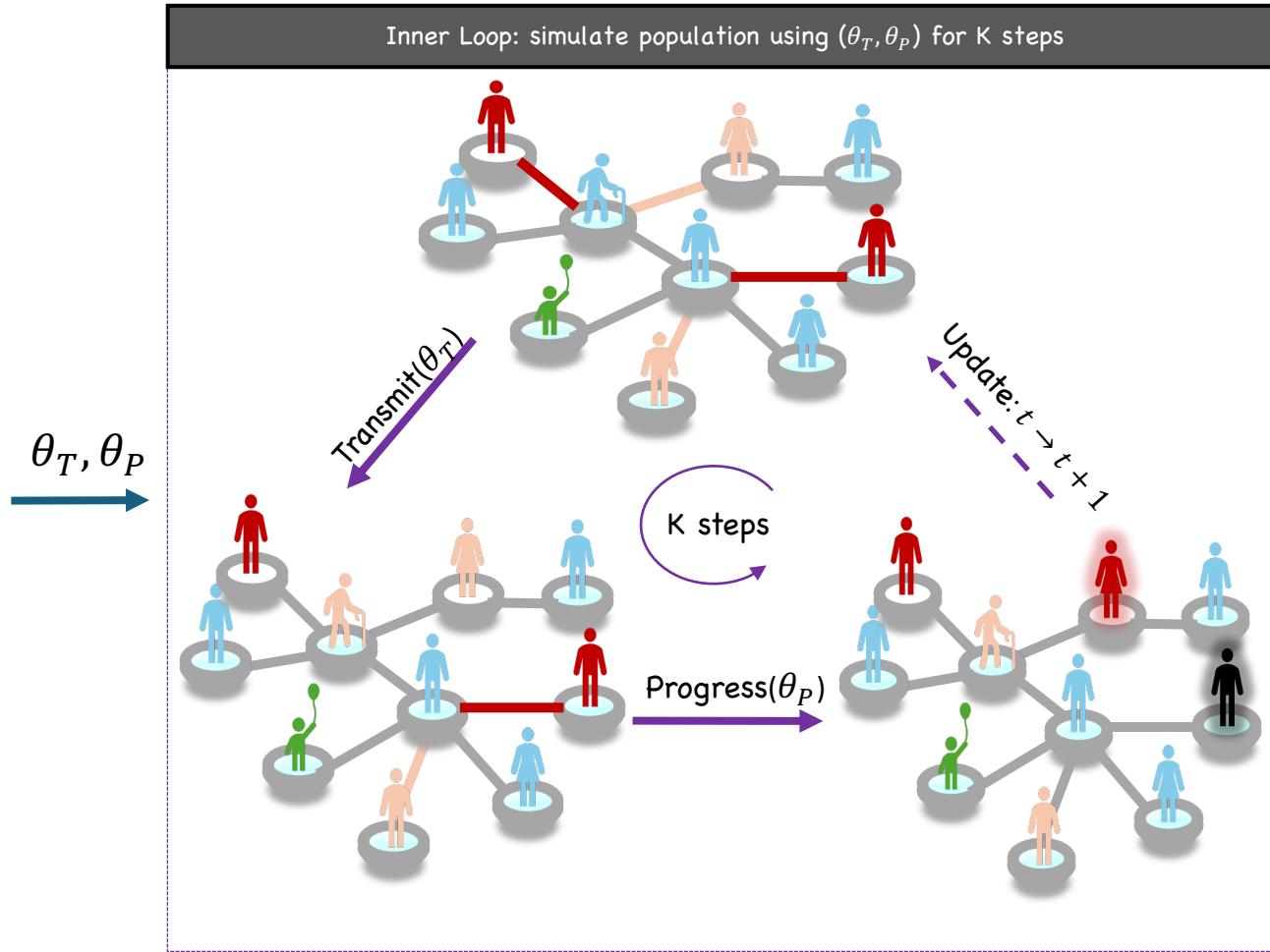
New York City

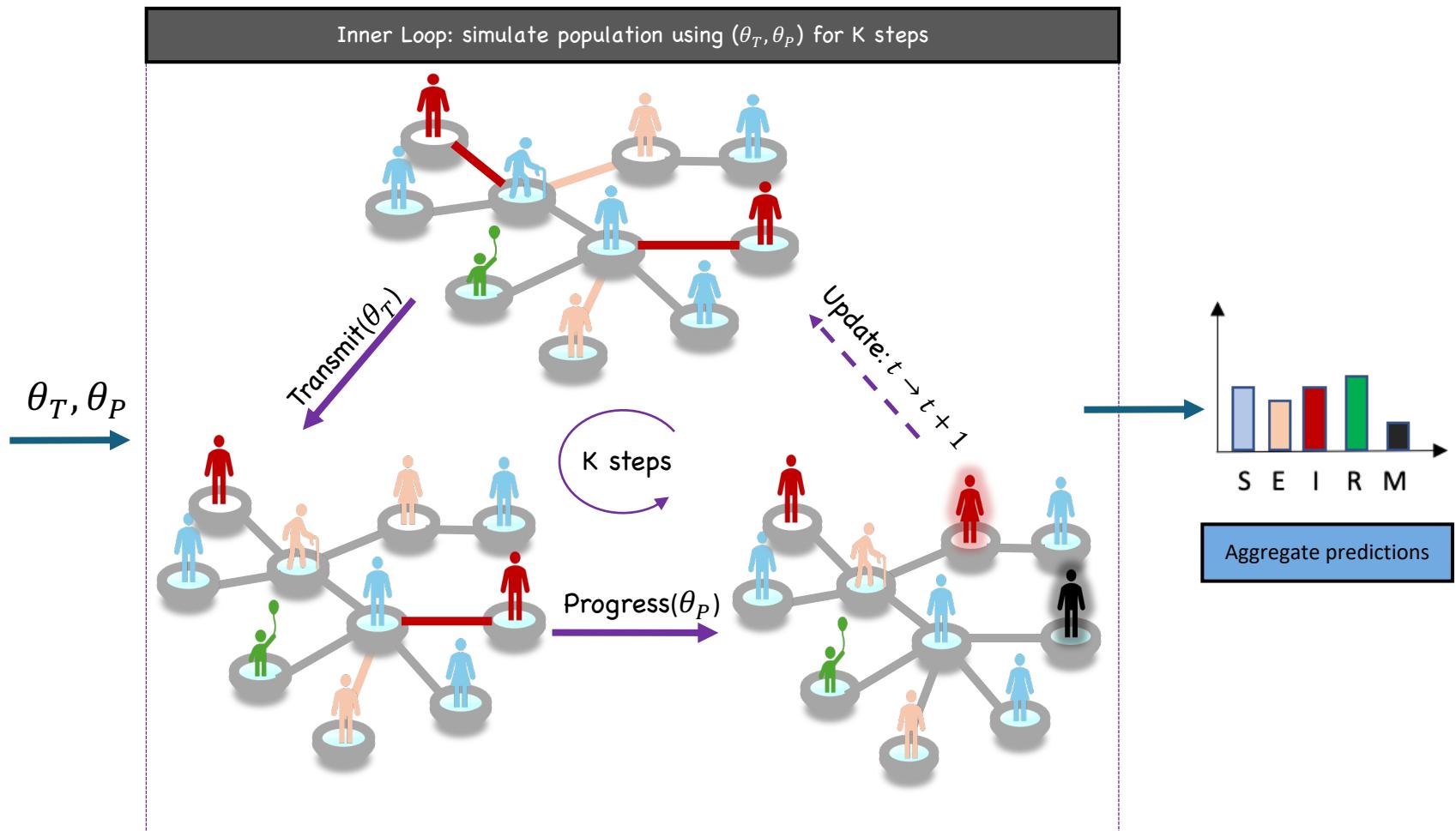
- 5 boroughs
- 8.4 million people
- 3.3 million households
- 200,000 small businesses



2020-2022







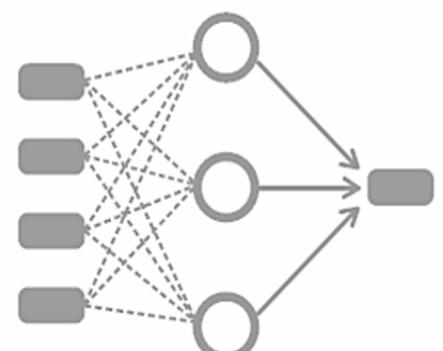
LPMs for Public Health



Capture

Disease Transmission
Viral Evolution

Simulate disease spread
across 8 million agents



Analyze

Wastewater Signal
Compliance Behavior

Calibrate to disease and
intervention data



Act

Test: Speed vs Accuracy?
\$1M: how Vaccinate?

Optimize policies while
preserving privacy

LPMs for Public Health



Capture

Disease Transmission
Viral Evolution

Simulate disease spread
across 8 million agents



Analyze

Wastewater Signal
Compliance Behavior

Calibrate to disease and
intervention data



Act

Test: Speed vs Accuracy?
\$1M: how Vaccinate?

Optimize policies will
preserving privacy

Get Started in 3 lines of code

Custom Templates, High-res visualizations



Capture

Disease Transmission
Viral Evolution

Simulate disease spread
across 8 million agents

```
● ● ●  
from agent_torch.core import envs  
from agent_torch.core import models, populations  
  
# create  
simulation = env.create_from_template(  
    model = models.epidemiology,  
    region=populations.NYC)  
  
# execute  
simulation.execute()  
  
# visualize  
simulation.visualize(  
    geo_index='household/positions' )
```

Compose: LLMs

Scale LLMs to prompt millions of agents



Capture

Disease Transmission
Viral Evolution

Simulate disease spread
across 8 million agents

```
● ● ●  
from agent_torch.core import Archetype, Behavior  
from agent_torch.populations import NYC  
  
# Create an object of Archetype class  
# n_arch estimates a predictive posterior over outcomes  
archetype = Archetype(n_arch=7)  
  
# Create an object of Behavior class  
work_behavior = Behavior(archetype=archetype.llm(prompt),  
                         region=NYC)  
  
will_work = work_behavior.sample()
```

Compose: ODEs

Simulate across multiple time scales!



Capture

Disease Transmission
Viral Evolution

Simulate disease spread
across 8 million agents

```
from chirho.dynamical.handlers.solver import TorchDiffEq
from chirho.dynamical.ops import simulate

simulation = envs.create(
    model=epidemiology,
    populations=NYC,
    archetype={'immune_dynamics':
                archetype.ode(chiro_ode, eval_times)
    })
simulation.execute()
```



Capture

Disease Transmission
Viral Evolution

Simulate disease spread
across 8 million agents

Implement Stochastic and Conditional Policies

Ensure Differentiability via custom operators



```
from torch.distributions import Bernoulli
from agent_torch.distributions import Bernoulli as AT_Bernoulli

p = torch.tensor([0.7], requires_grad=True)

out1 = Bernoulli.sample(p)
out1.backward()
p.grad # [None.]

out2 = AT_Bernoulli.sample(p)
out2.backward()
p.grad # [0.7]
```

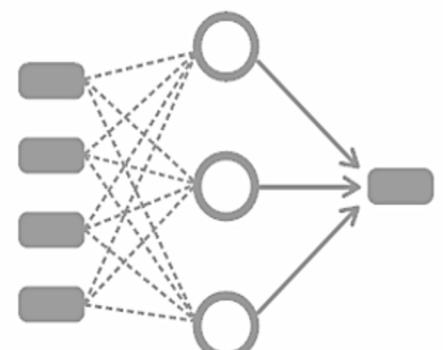
LPMs for Public Health



Capture

Disease Transmission
Viral Evolution

Simulate disease spread
across 8 million agents



Analyze

Wastewater Signal
Compliance Behavior

Calibrate to disease and
intervention data



Act

Test: Speed vs Accuracy?
\$1M: how Vaccinate?

Optimize policies while
preserving privacy

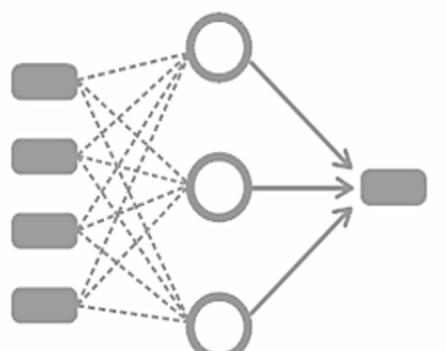
LPMs for Public Health



Capture

Disease Transmission
Viral Evolution

Simulate disease spread
across 8 million agents



Analyze

Wastewater Signal
Compliance Behavior

Calibrate to disease and
intervention data



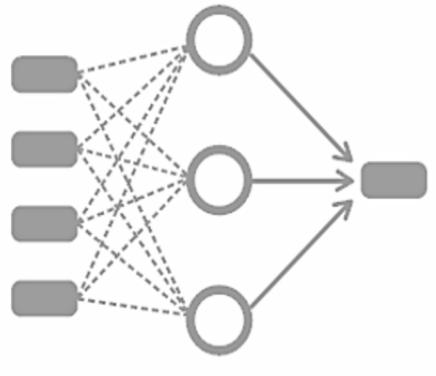
Act

Test: Speed vs Accuracy?
\$1M: how Vaccinate?

Optimize policies will
preserving privacy

Compose end-to-end with NNs

Calibrate the simulator instead of a surrogate!



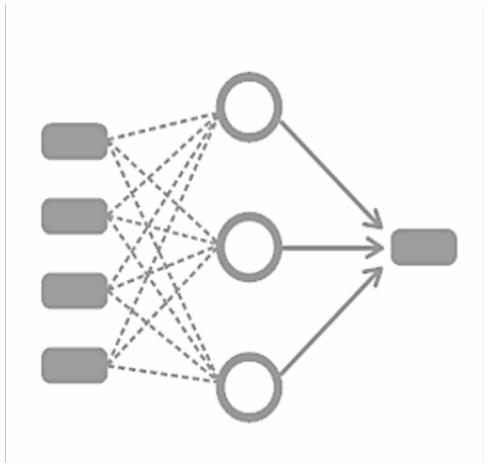
Analyze

Wastewater Signal
Compliance Behavior

Calibrate to disease and
intervention data

```
● ● ●  
# agent_"torch" works seamlessly with the pytorch API  
from torch.optim import SGD  
  
nn = compose_nn()  
  
for i in range(n_epochs):  
    parameters = compose_nn.forward()  
  
    simulation.step(params=parameters)  
    cases, employment = simulation.predict()  
  
    simulation.optimize(SGD)  
    simulation.reset()
```

AgentTorch simulators can be calibrated directly, without a surrogate



Analyze

Wastewater Signal
Compliance Behavior

Calibrate to disease and
intervention data

Method	Calibration	Analysis
Calibrate a Surrogate	100,000 hours	5,000 hours
Calibrate the Simulator	20 minutes	10 seconds

+ 3000x

+ 5000x

- Calibration: Learn from multi-modal historical data + generate sparse simulation data!
- Analysis: Compute sensitivity without executing ANY simulations!

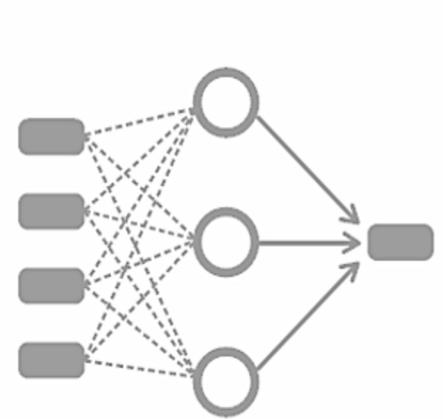
LPMs for Public Health



Capture

Disease Transmission
Viral Evolution

Simulate disease spread
across 8 million agents



Analyze

Wastewater Signal
Compliance Behavior

Calibrate to disease and
intervention data



Act

Test: Speed vs Accuracy?
\$1M: how Vaccinate?

Optimize policies while
preserving privacy

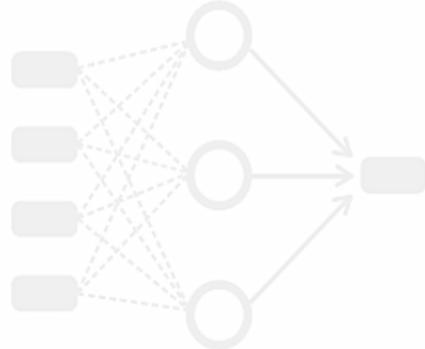
LPMs for Public Health



Capture

Disease Transmission
Viral Evolution

Simulate disease spread
across 8 million agents



Analyze

Wastewater Signal
Compliance Behavior

Calibrate to disease and
intervention data



Act

Test: Speed vs Accuracy?
\$1M: how Vaccinate?

Optimize policies while
preserving privacy

Optimize Policies in Natural Language

Custom analyzers for AgentTorch API



Act

Test: Speed vs Accuracy?
\$1M: how Vaccinate?

Optimize policies while
preserving privacy

```
from agent_torch.analyzer import SimulationAnalysisAgent

analyzer = SimulationAnalysisAgent(
    openai_api_key=OPENAI_API_KEY,
    simulation=simulation, # The simulation executor
    document_retriever=retriever
)

# exploratory
response = analyzer.query("Which age group has lowest median income, how much is it?")

# counterfactual
response = analyzer.query("How would deaths change if the R0 had been 4.5?")

# prospective
analyzer.query("What would happen if isolation time is increased and virus become more transmissible?")
```

Deploy decision-making to the edge



Act

Test: Speed vs Accuracy?
\$1M: how Vaccinate?

Optimize policies while
preserving privacy

Execute AgentTorch simulations over decentralized protocols



```
from agent_torch.core import decentralize

# map simulation to real protocol
real_sim = decentralize(
    simulation,
    protocol="ble_contact_tracing")

real_sim.deploy()

real_sim.sync()
```



Act

Test: Speed vs Accuracy?
\$1M: how Vaccinate?

Optimize policies while
preserving privacy

Execute over Contact Tracing Protocols

Official Exposure Notification app in 5 US states / territories



Guam

Guam Covid Alert



Cyprus

COVTracer EN



Hawaii

AlohaSafe Alert



Minnesota

COVIDaware MN



- Congressional Testimony
- World's largest open-source project for COVID-19

Close Sim2Real Gap: Contact Intelligence

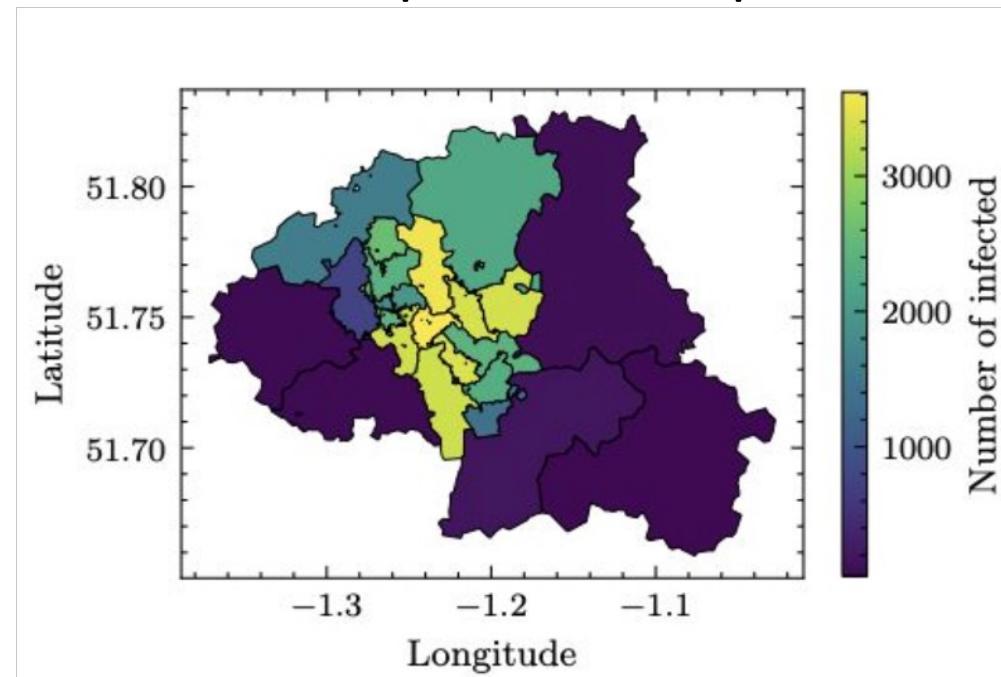


Act

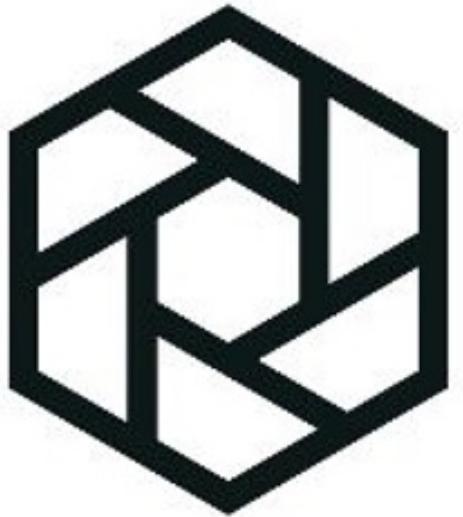
Test: Speed vs Accuracy?
\$1M: how Vaccinate?

Optimize policies while
preserving privacy

Synchronize simulation and protocol for
real-time and private computation



Track disease without leaking individual demographic,
disease status or geolocation



AgentTorch

Imagine... your enterprise a living lab

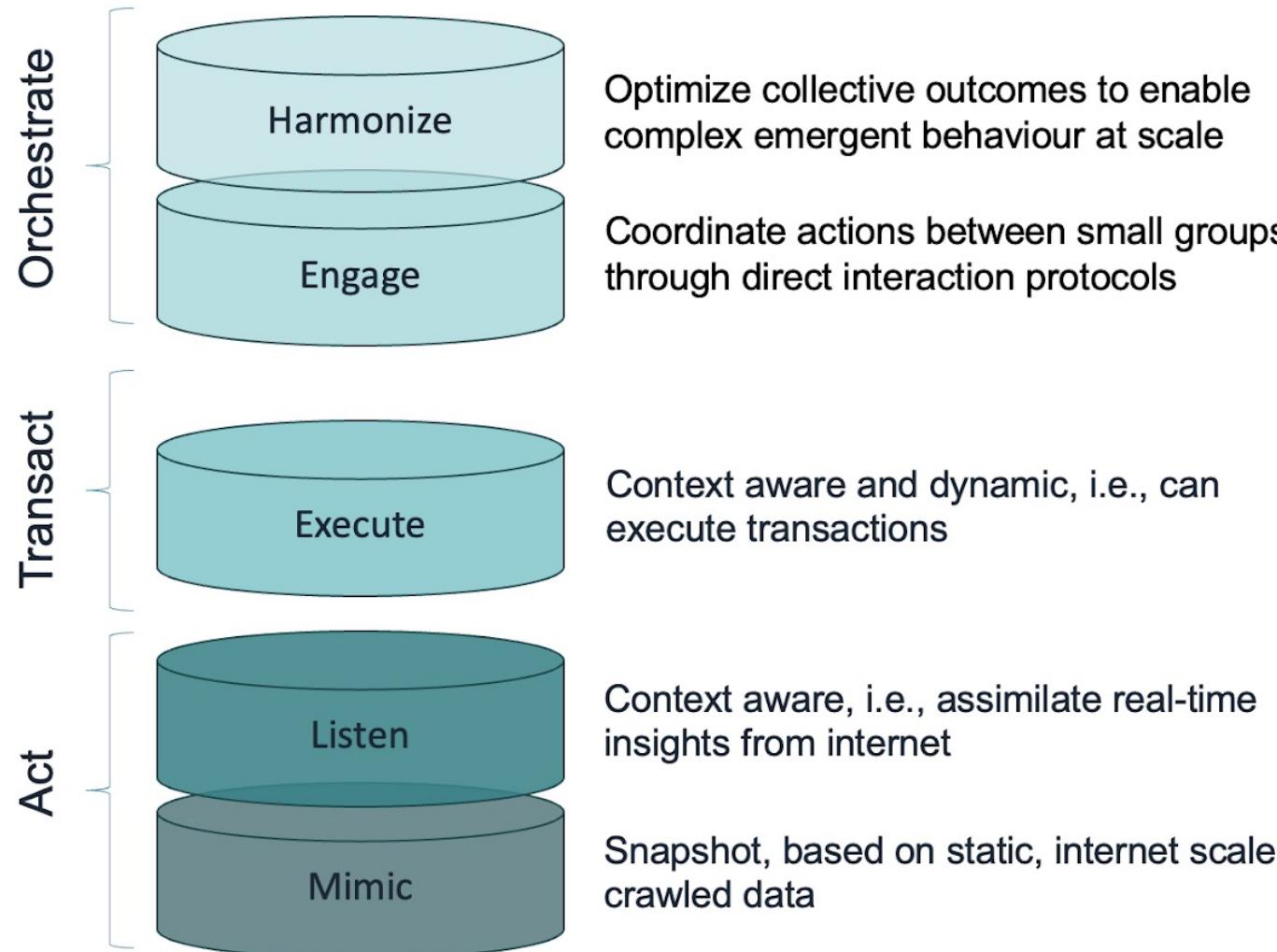


Large Population Models



Appendix

Levels of Agentic Systems

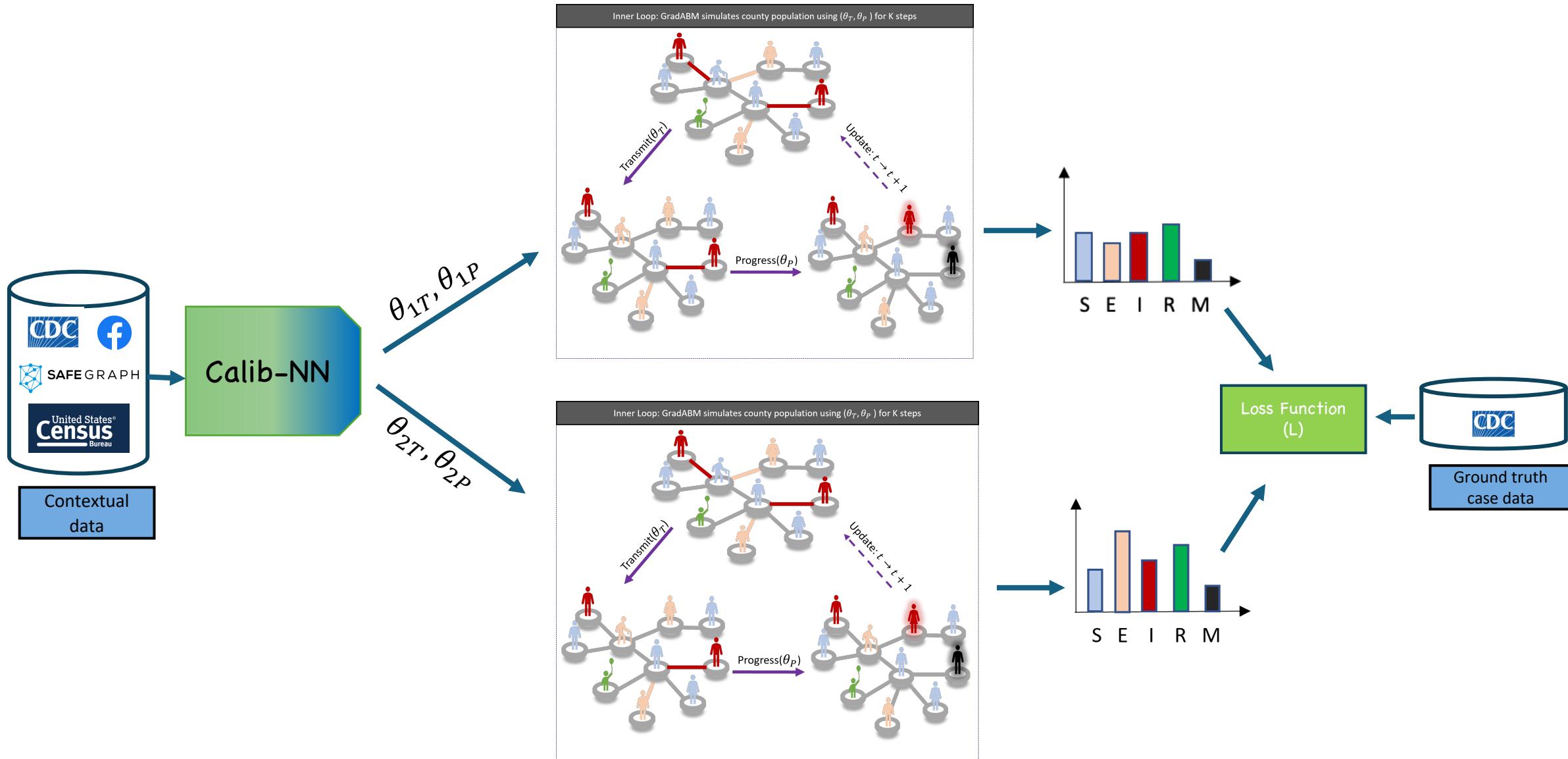


Comparison with other frameworks

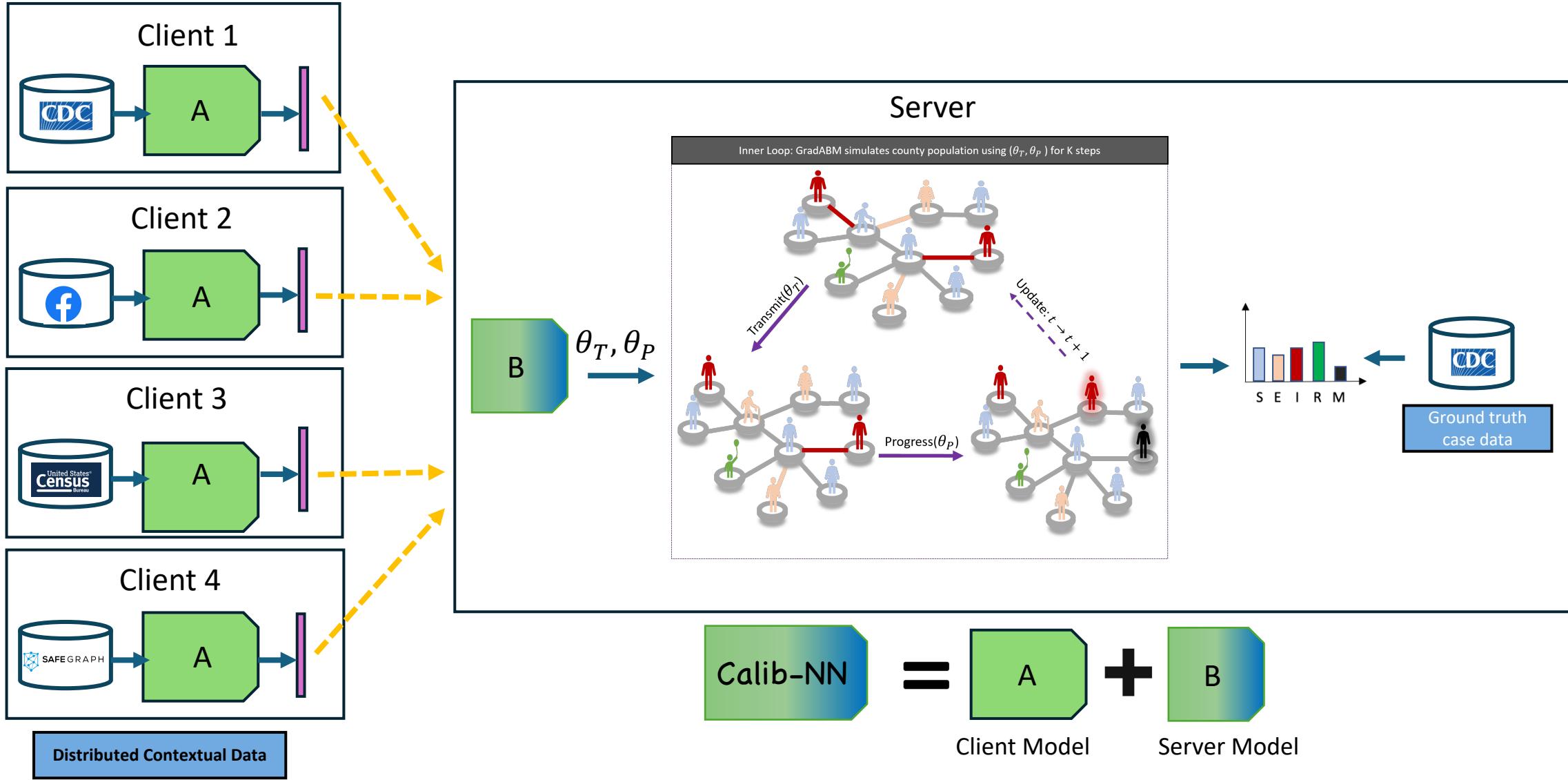
Feature	AgentTorch	Concordia	Flame	Mesa
GPU Execution	✓	✓	✓	✗
Million-agent Populations	✓	✗	✓	⚠
Differentiable Environments	✓	✗	✗	✗
Mechanistic Environments	✓	✗	✓	✓
LLM Integration	✓	✓	✗	✗
Neural Composition	✓	⚠	✗	✗

Data for Differentiable Learning

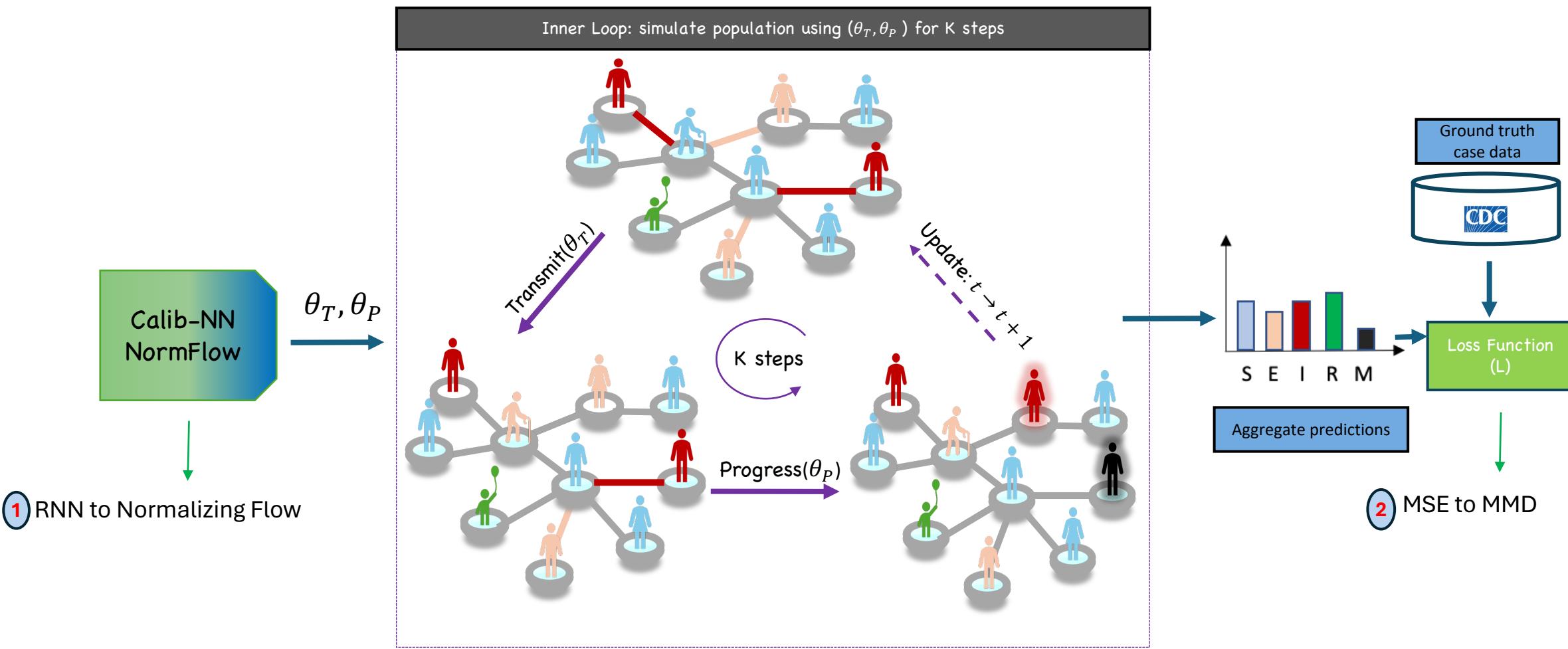
Multi-task learning over simulation environments, reduce overfitting



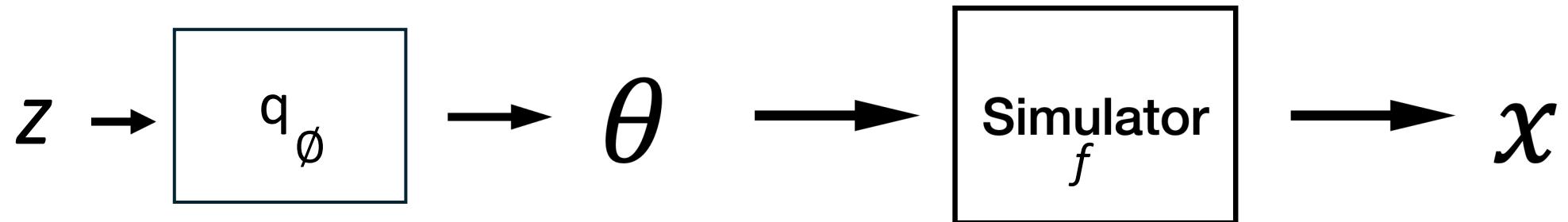
Federated calibration of simulation parameters

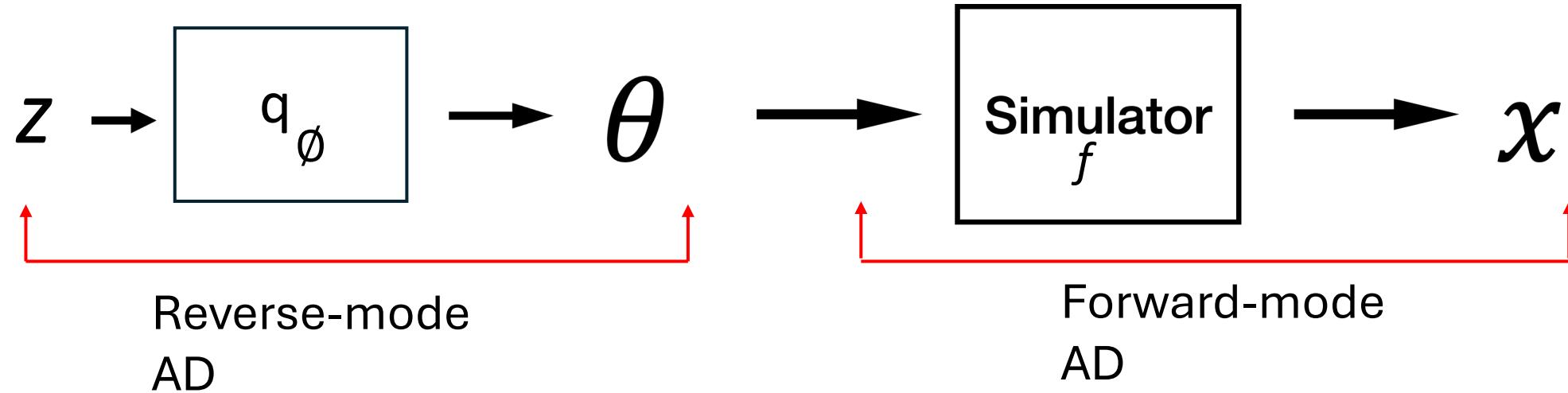


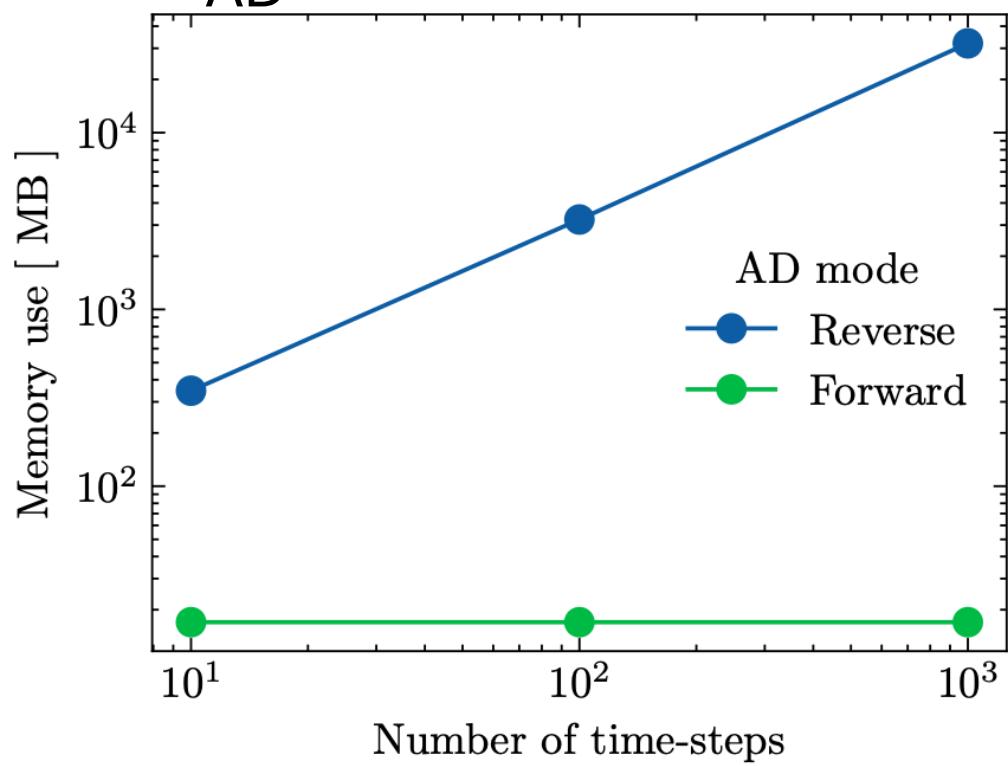
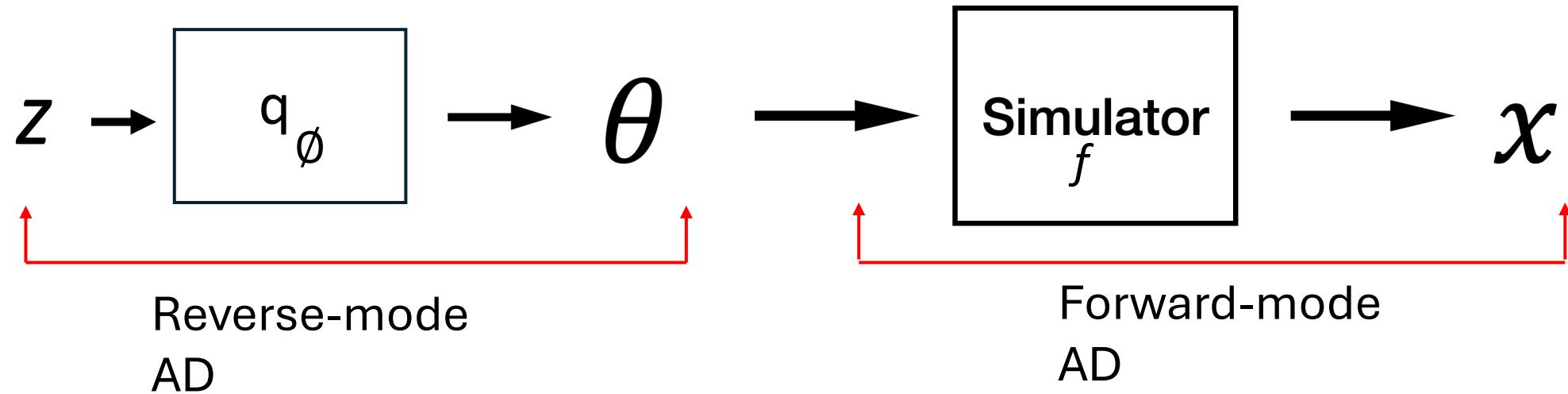
Estimate posteriors over simulation parameters

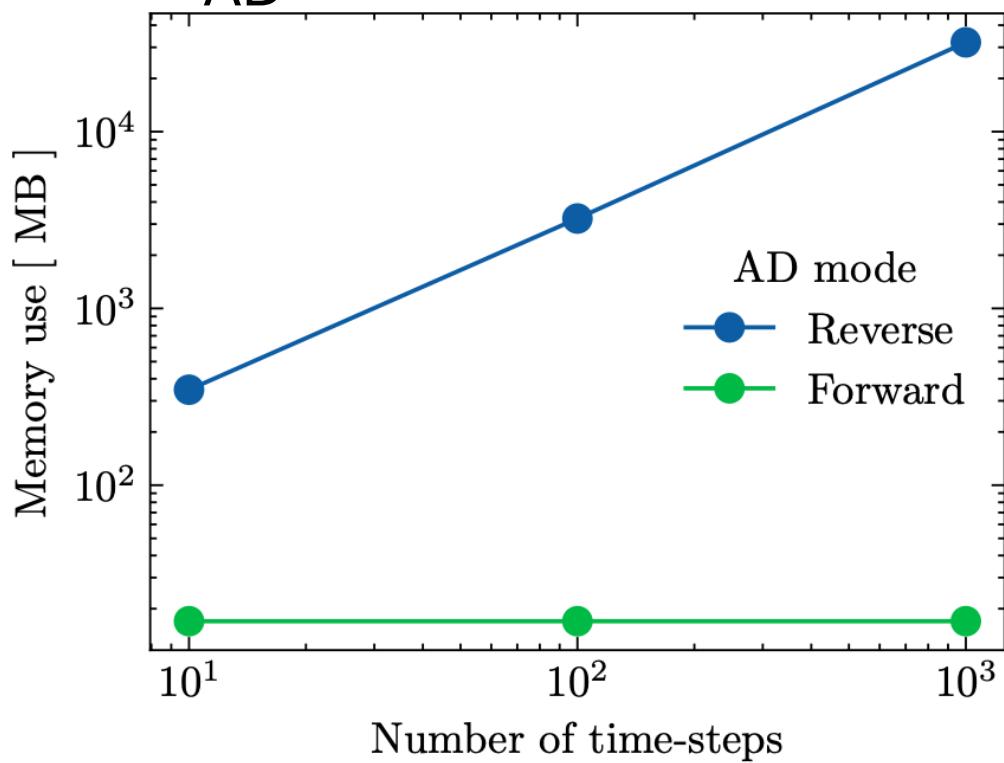
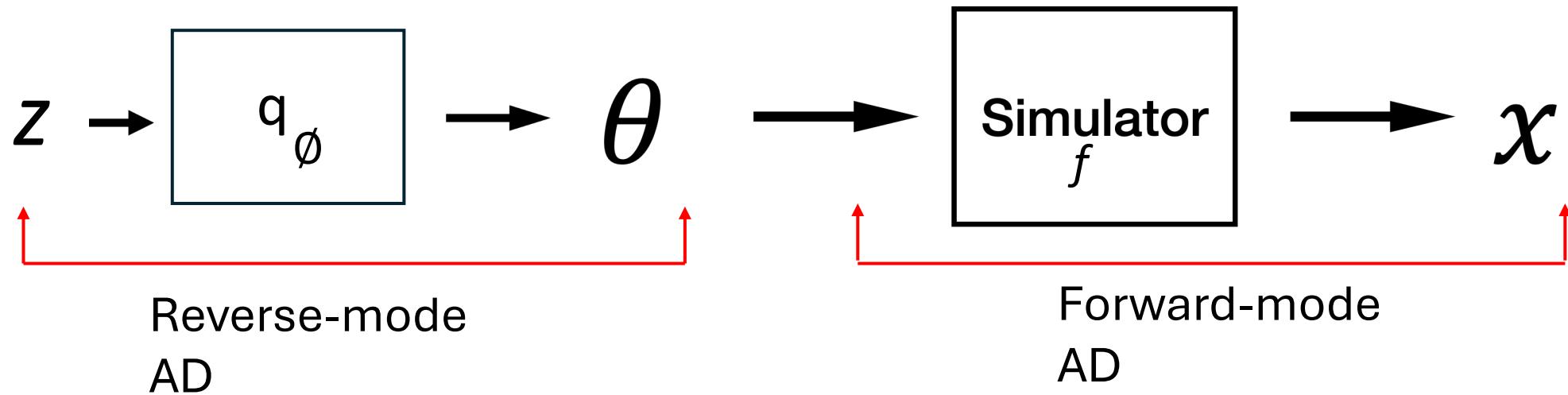


Advanced Differentiable Learning: Memory Optimization









$$\frac{\partial x}{\partial \phi} = J_f \cdot g_q$$

$$J_f = \frac{\partial f(\theta)}{\partial \theta}$$

$$g_q = \frac{\partial q_\phi(z)}{\partial \phi}$$

Advanced Differentiable Learning: Zero-shot Sensitivity Analysis

Sensitivity Analysis is critical for validation

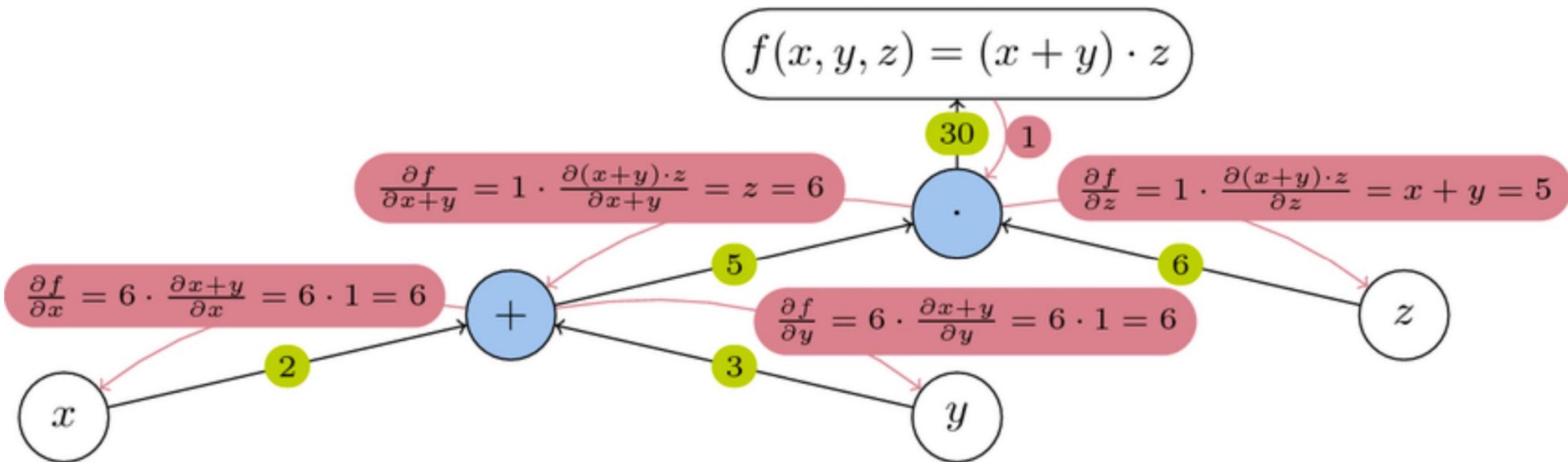
The impact of uncertainty on predictions of the CovidSim epidemiological code

Wouter Edeling¹, Hamid Arabnejad¹ , Robbie Sinclair³, Diana Suleimenova², Krishnakumar Gopalakrishnan¹ , Bartosz Bosak⁴, Derek Groen², Imran Mahmood², Daan Crommelin^{1,5} and Peter V. Coveney^{1,3,6} 

Epidemiological modelling has assisted in identifying interventions that reduce the impact of COVID-19. The UK government relied, in part, on the CovidSim model to guide its policy to contain the rapid spread of the COVID-19 pandemic during March and April 2020; however, CovidSim contains several sources of uncertainty that affect the quality of its predictions: parametric uncertainty, model structure uncertainty and scenario uncertainty. Here we report on parametric sensitivity analysis and uncertainty quantification of the code. From the 940 parameters used as input into CovidSim, we find a subset of 19 to which the code output is most sensitive—imperfect knowledge of these inputs is magnified in the outputs by up to 300%. The model displays substantial bias with respect to observed data, failing to describe validation data well. Quantifying parametric input uncertainty is therefore not sufficient: the effect of model structure and scenario uncertainty must also be properly understood.

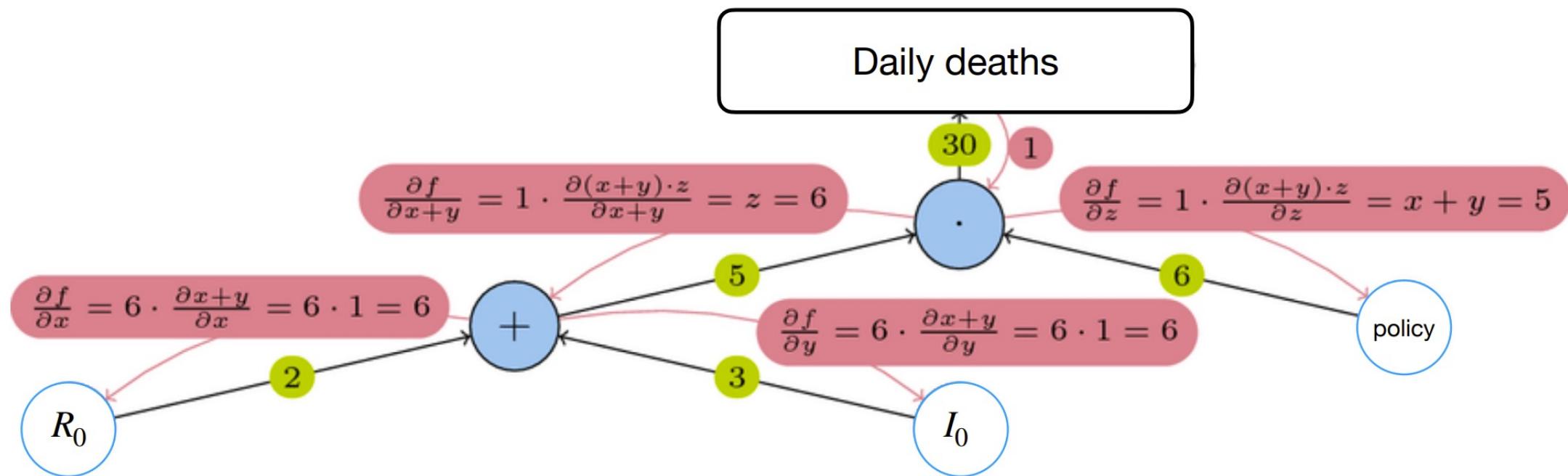
Ensemble execution. Consequently, through the use of adaptive methods we make the uncertainty analysis of CovidSim tractable, but our analysis nevertheless required us to perform thousands of runs, each with its own unique set of input parameters. Specifically, we used the Eagle supercomputer at the Posnan

Recap: Reverse-mode automatic differentiation



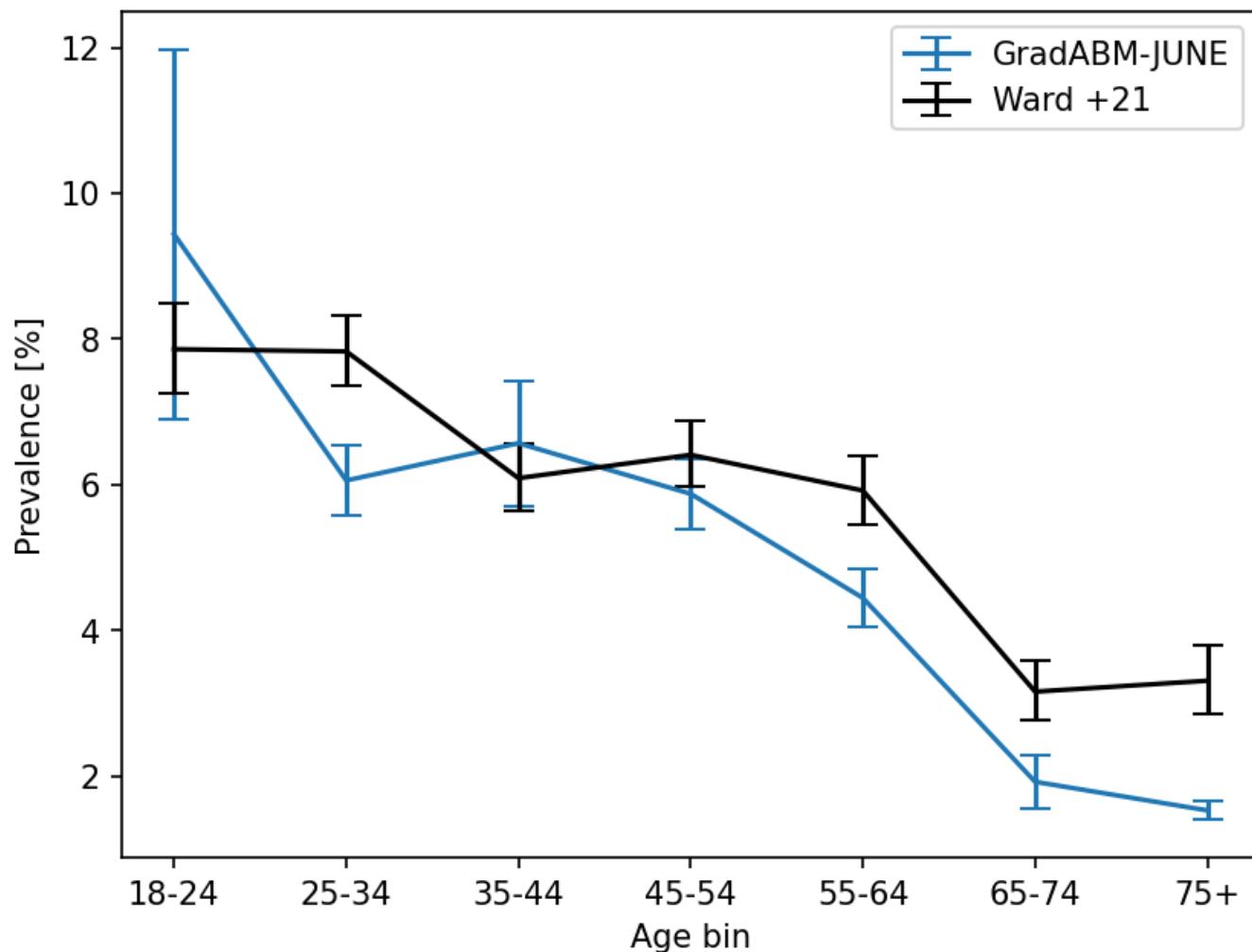
Sensitivity analysis via reverse-mode automatic differentiation

Reverse-mode automatic differentiation is independent of the number of parameters!!



How effective *really* were lockdown policies?

Analyze retrospective decisions by reproducing seroprevalence studies in-silico



What could we have done differently?

Design counterfactual lockdown policies with multiple constraints in-silico!

