What can we learn from a **billion** agents?

Ayush Chopra
ayushc@mit.edu

AgentTorch

mit media lab

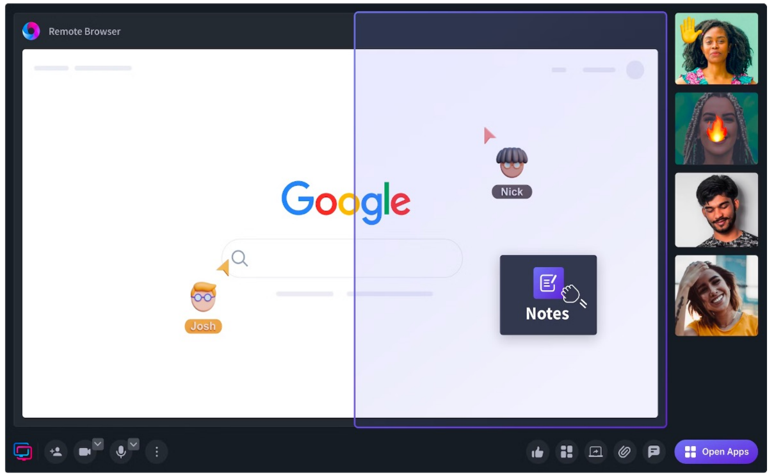
Massachusetts Institute of Technology

Ayush Chopra

MIT (PhD’25, MS’22)
Adobe, JP Morgan, RemoteHQ (acq)

Ayush Chopra

MIT (PhD’25, MS’22)
Adobe, JP Morgan, RemoteHQ (acq)
25 Patents in AI and Simulations



RemoteHQ

Remote Work, *someday*?
Share **browsers** not screens!

#1 on ProductHunt,
acquired by Presence

Adobe

Retail in the Browser?
Distributed AI, Virtual Try-on, Visual Search

Outstanding Young Engineer Award,
Started Adobe Fashion: Research to Product

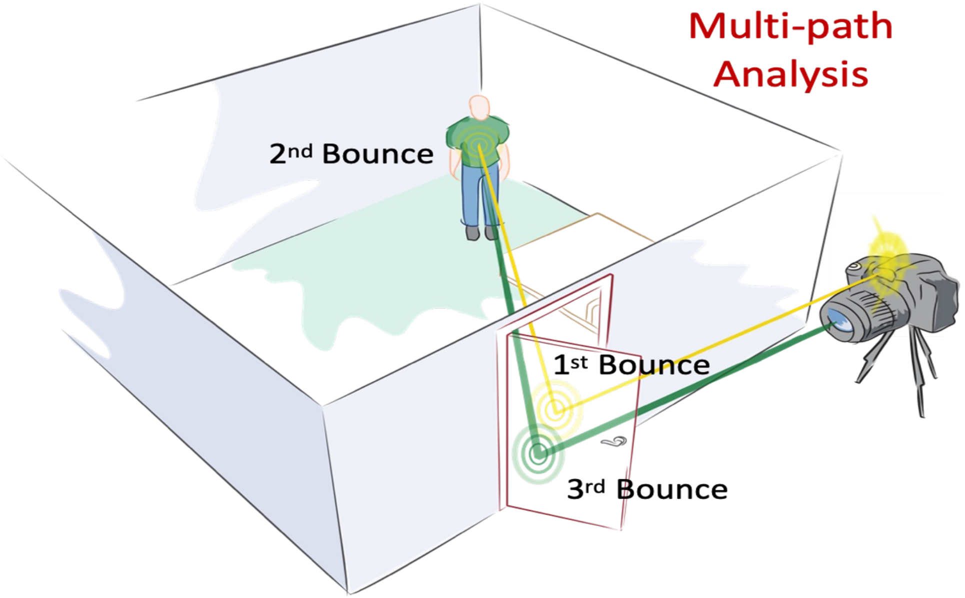
MIT Media Lab

Seeing the Invisible?
Large Population Models

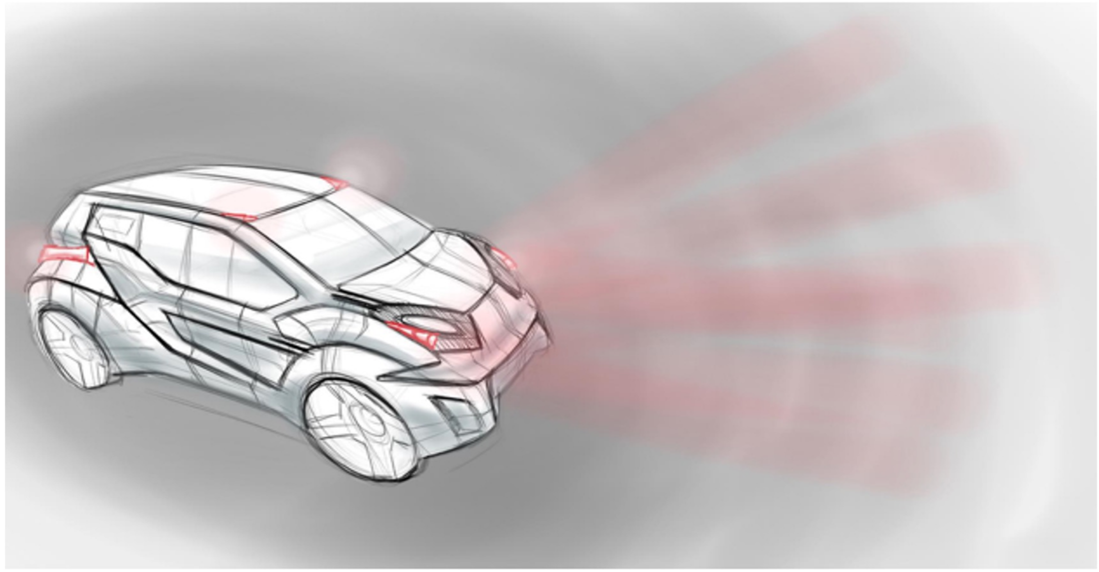
MIT Media Lab Rising Star, >20 papers,
Reach 50M globally: Health, Supply Chain

Seeing the Invisible

Scaling photonic simulations to image beyond individual line-of-sight



Look around corners?



See through fog?

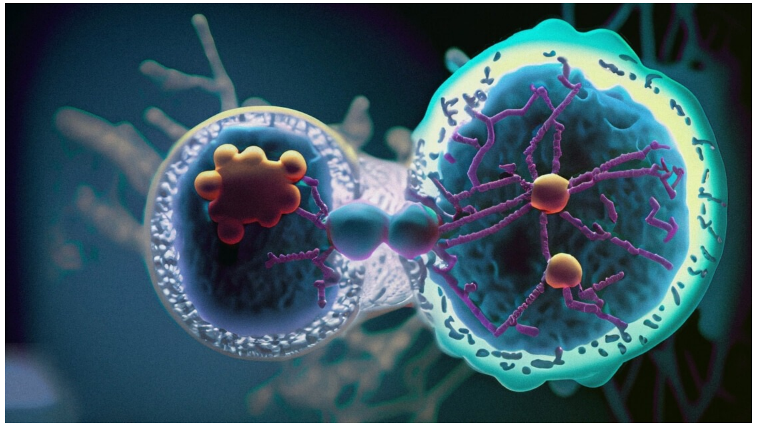
God’s Eye View

Scaling population simulations to orchestrate billions of interacting agents

Citizens

Disease Spread?

Cells



Immune Response?

Cities

Supply Chain Disruptions?

Inside the body, around us and beyond!

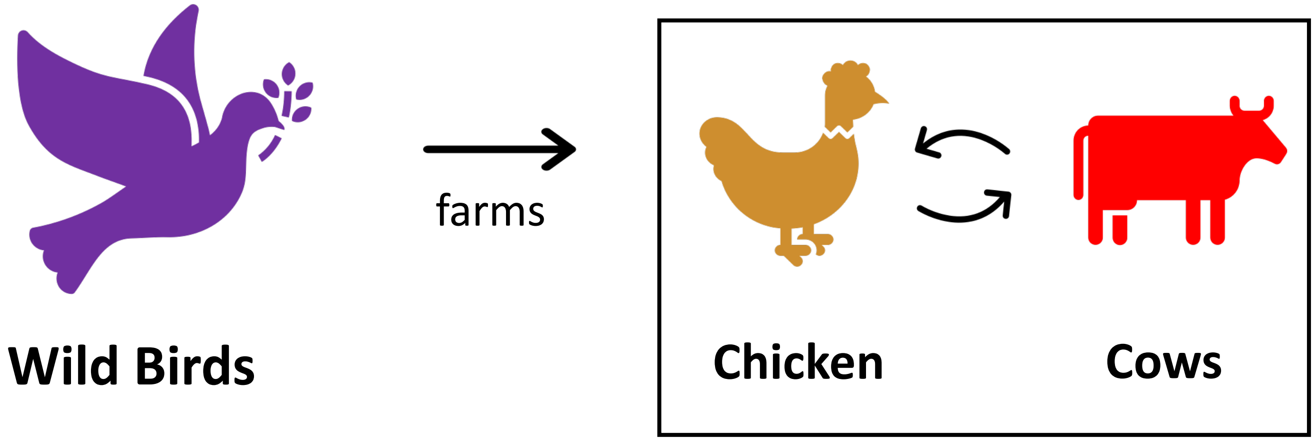

108 SQ FT PER HEN
ROTATED PASTURES
EST. 2007
Vital FARMS
TENDED BY HAND ON SMALL FAMILY FARMS
PASTURE-RAISED
EQUIVALENT TO 9 LARGE EGGS PER CARTON
KEEP REFRIGERATED
VF PTR RSD LRG
EGGS 18 CT
F13
346200
035-188
18CT
10.99
JUST CRACK
SCRAMBLE
FARM FRESH EGGS
Local

Eggs so expensive? Blame bird migration, not inflation…

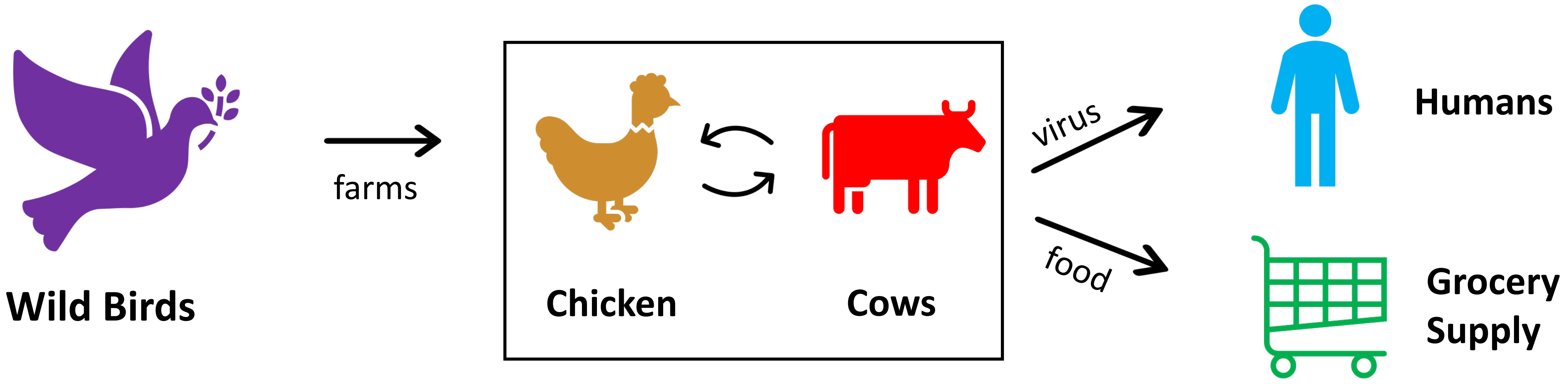
The H5N1 Crisis

Wild Birds

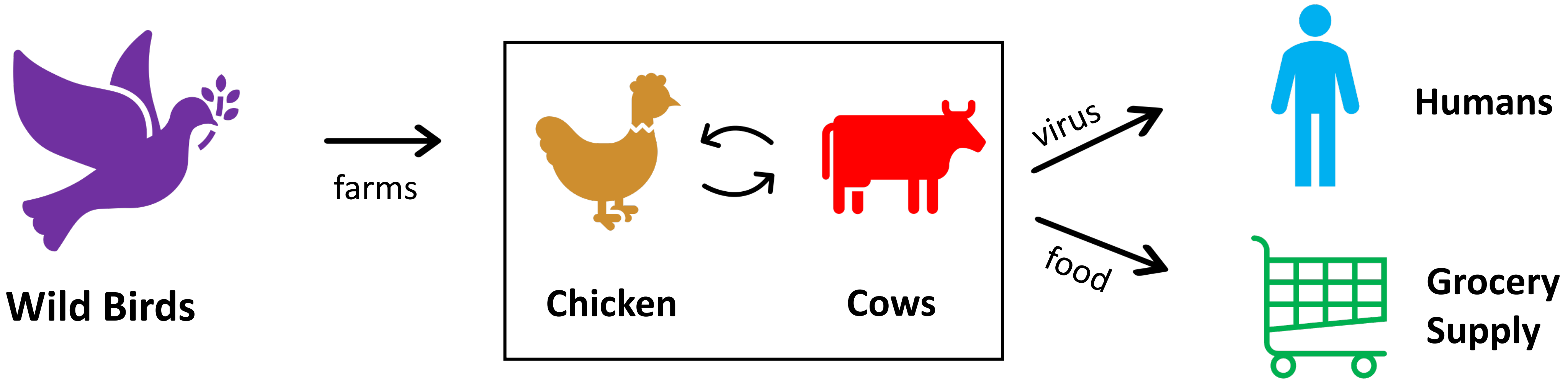
The H5N1 Crisis



The H5N1 Crisis



The H5N1 Crisis



The H5N1 Crisis

100 million chicken culled

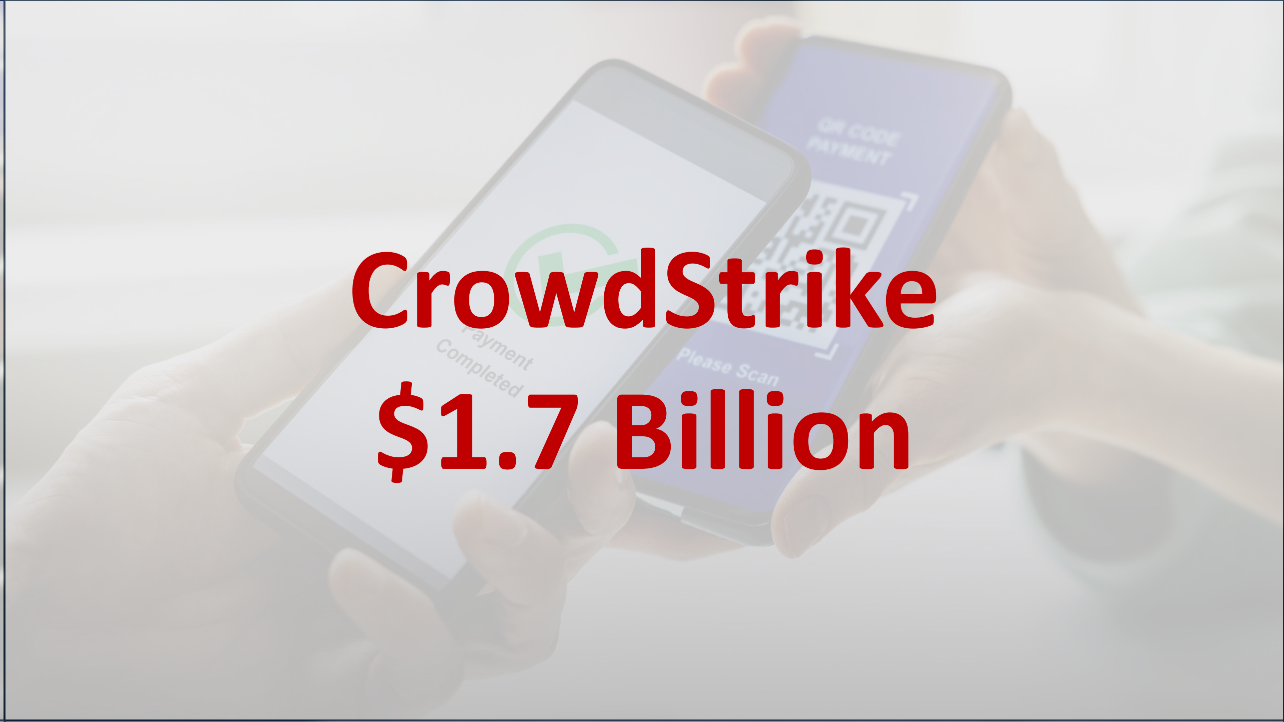
$4 billion revenue loss

Human-human infections!

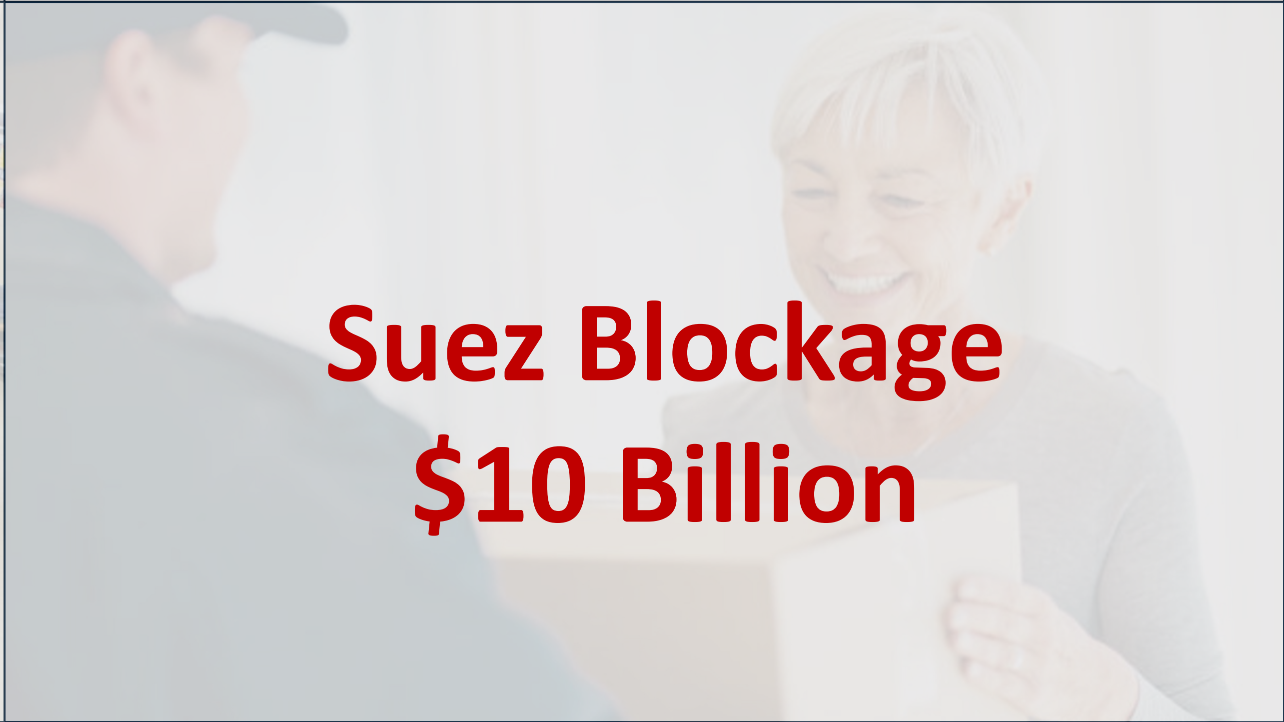

PUSHING DAISIES


Payment
Completed
QR CODE
PAYMENT
Please Scan

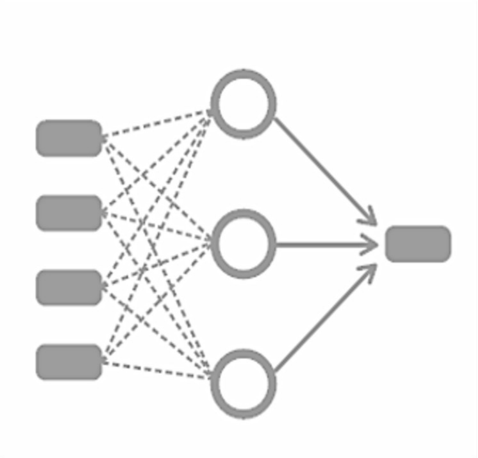

COVID
$3.5 Trillion


CrowdStrike
$1.7 Billion


Hurricane Ida
$80 Billion

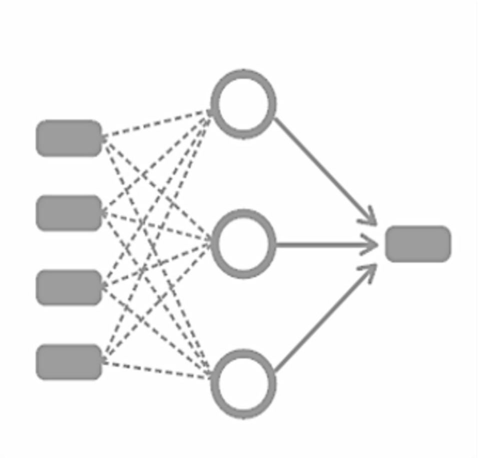

Suez Blockage
$10 Billion

Imagine if… God's Eye View



Capture	Analyze	Act
Bird Migration Disease Transmission	Farmer Incentives Supply Disruptions	Stop Spread Control Prices

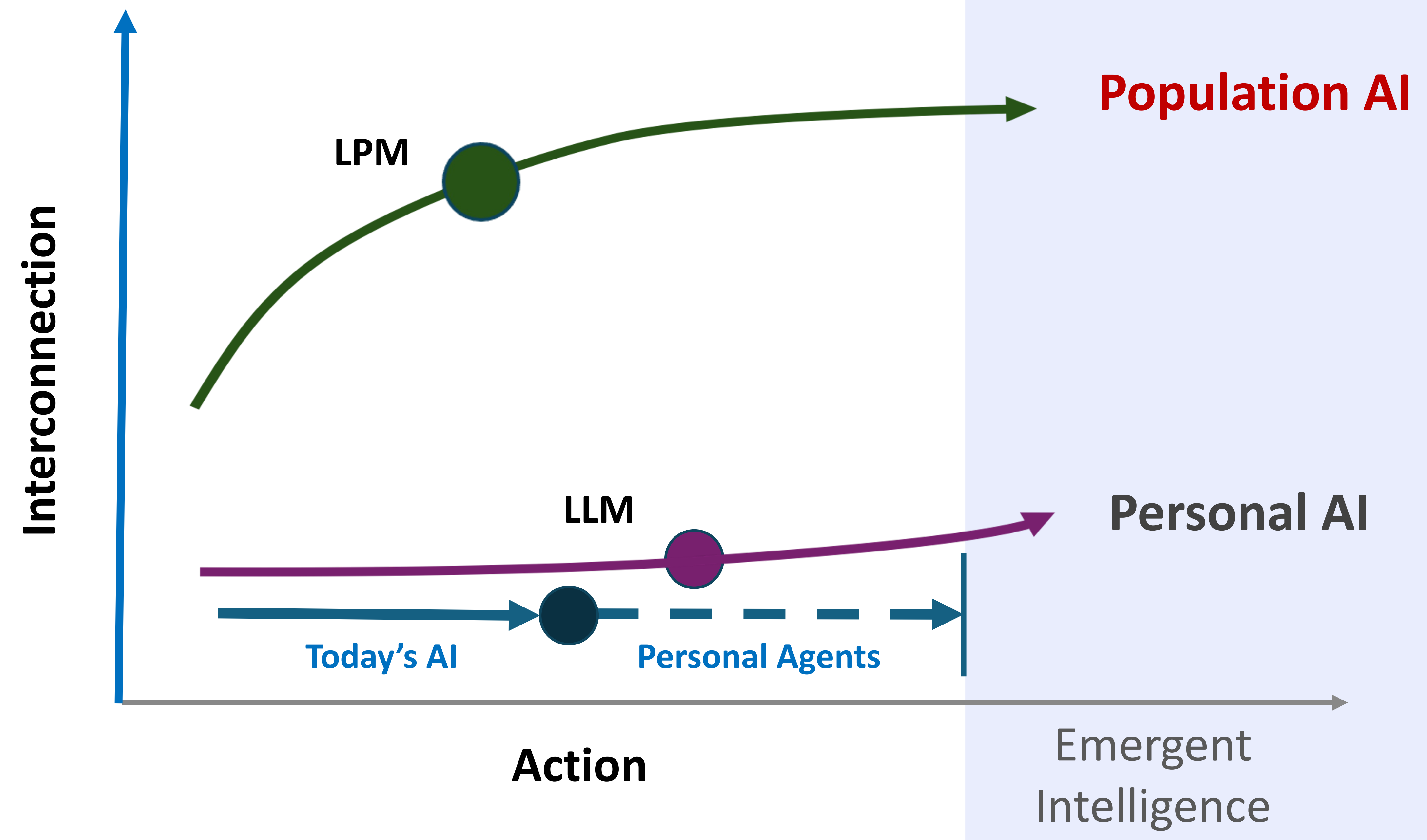
God’s Eye View is Not Trivial



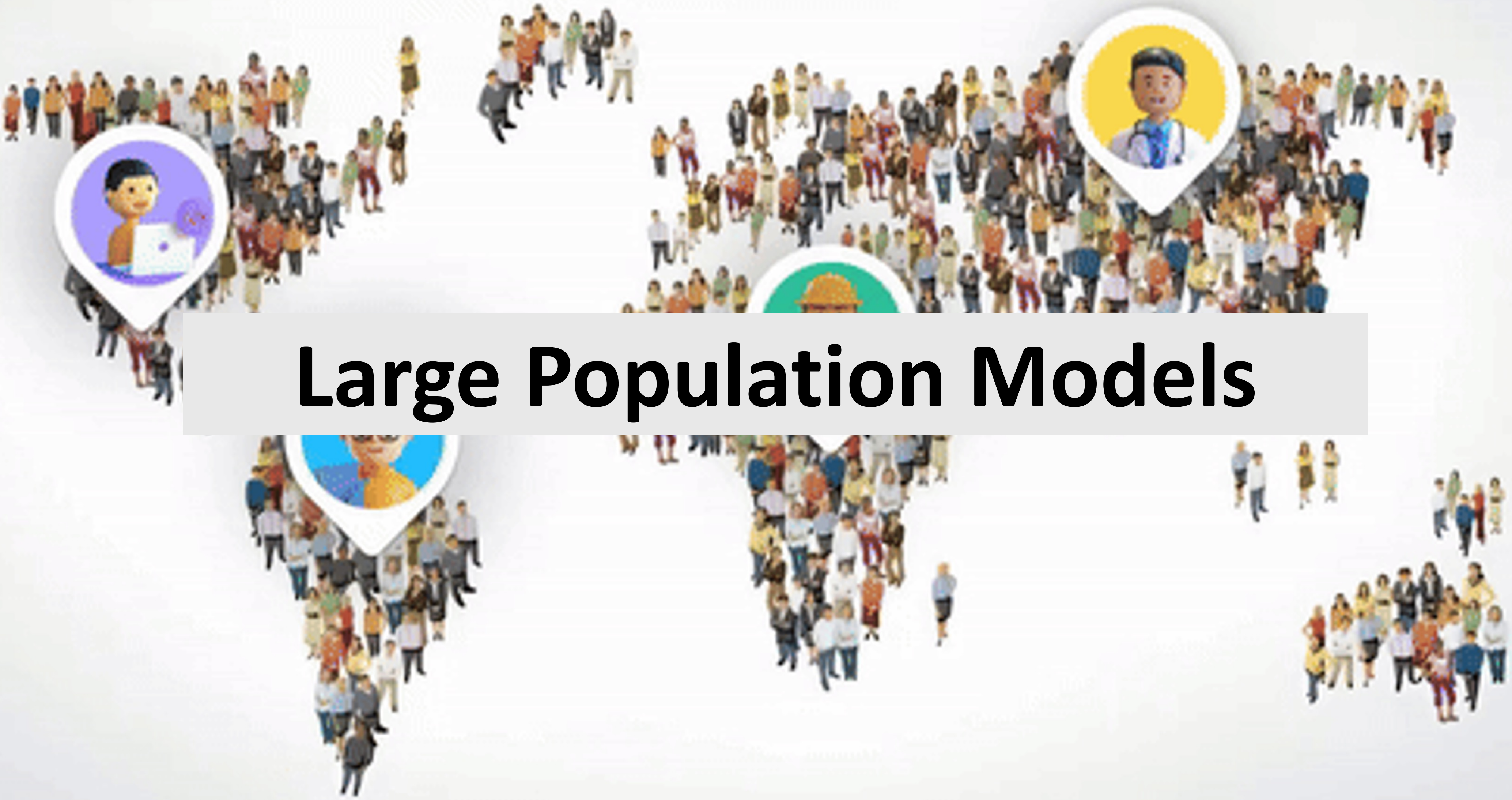
Capture	Analyze	Act
Bird Migration Disease Transmission	Farmer Incentives Supply Disruptions	Stop Spread Control Prices
Web of Interconnections	Heterogeneous Behaviors	Multi-scale Decisions

Large ~~Language~~ Population Models

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

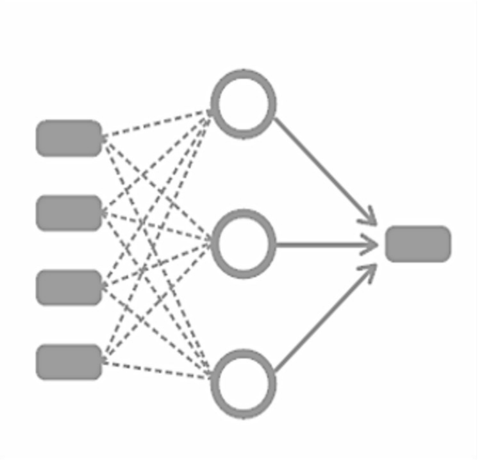

Population AI
LPM
Interconnection
LLM
Personal AI
Today's AI
Personal Agents
Action
Emergent
Intelligence

Large Population Models



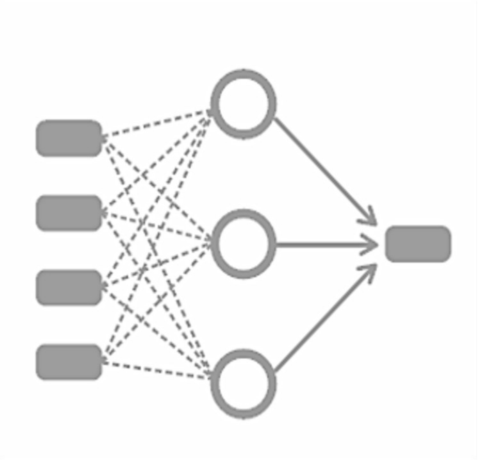
lpm.media.mit.edu

LPMs: Building the God's Eye View



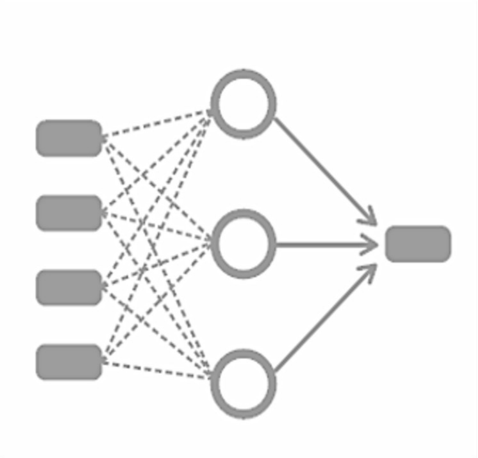
Capture	Analyze	Act
Bird Migration Disease Transmission	Farmer Incentives Supply Disruptions	Stop Spread Control Prices
Web of Interconnections	Heterogeneous Data	Multi-scale Decisions

LPM: Technical Pillars



Capture	Analyze	Act
Multi-scale Dynamics Stochastic Protocols	Multi-modal data Adaptive behavior	Multi-objective Decisions Real-time Response
Compositional Simulations	Differentiable Learning	Decentralized Analysis

LPM: Technical Pillars



Capture	Analyze	Act
Multi-scale Dynamics Stochastic Protocols	Multi-modal data Adaptive behavior	Multi-objective Decisions Real-time Response
Compositional Simulations	Differentiable Learning	Decentralized Analysis
AAMAS’23, AAMAS'24, Nature Medicine’25	AAMAS’24 (Best Paper Runner-Up), AAMAS’23	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

Chopra et al – a framework for learning in agent-based models (AAMAS'24, Oral)

8.4 million agents

AgentTorch

Large Population Models: Reaching millions around the world

World Health Organization

MAYO CLINIC

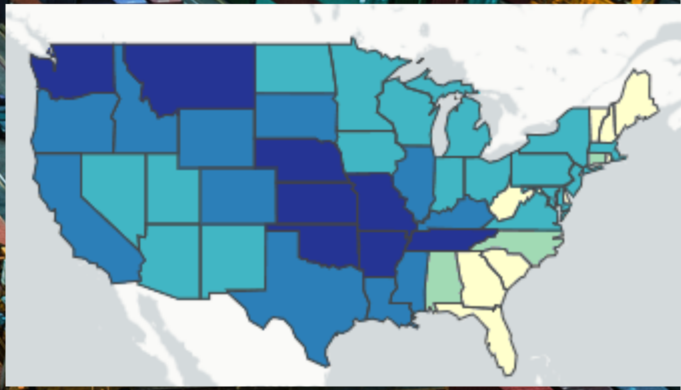
Te Kāwanatanga o Aotearoa
New Zealand Government

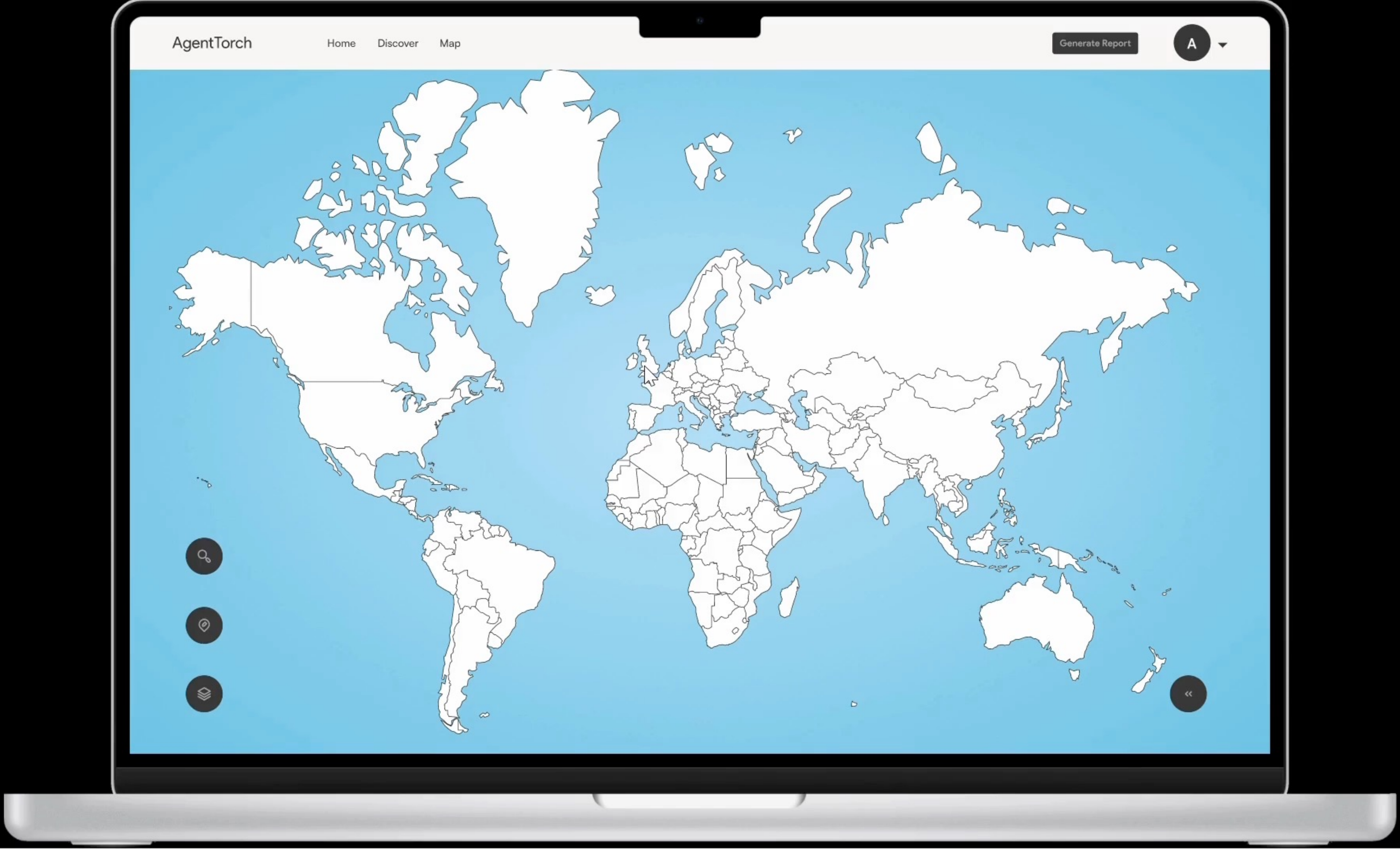

SIMCITY™
BUILDIT

5 million citizens

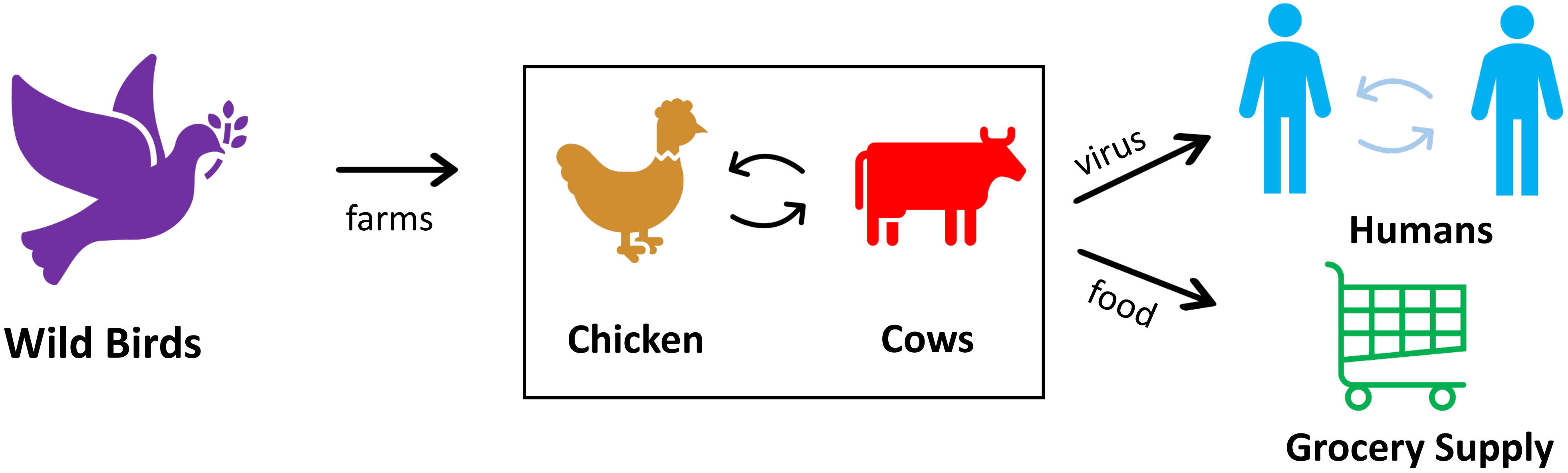
30 million households

$ billion supply chains

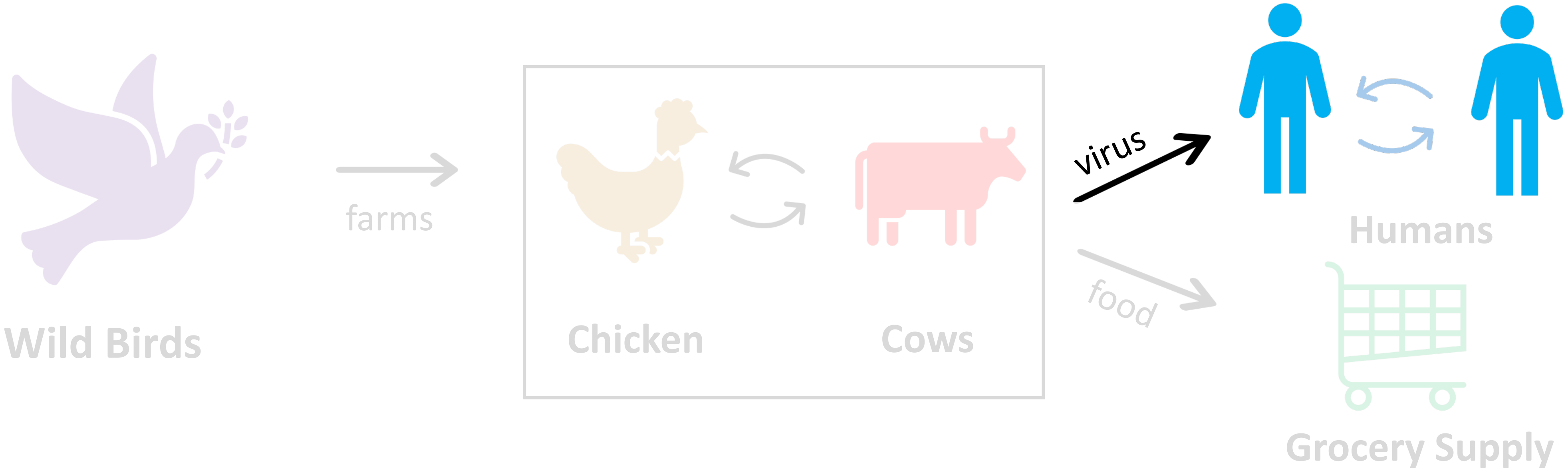


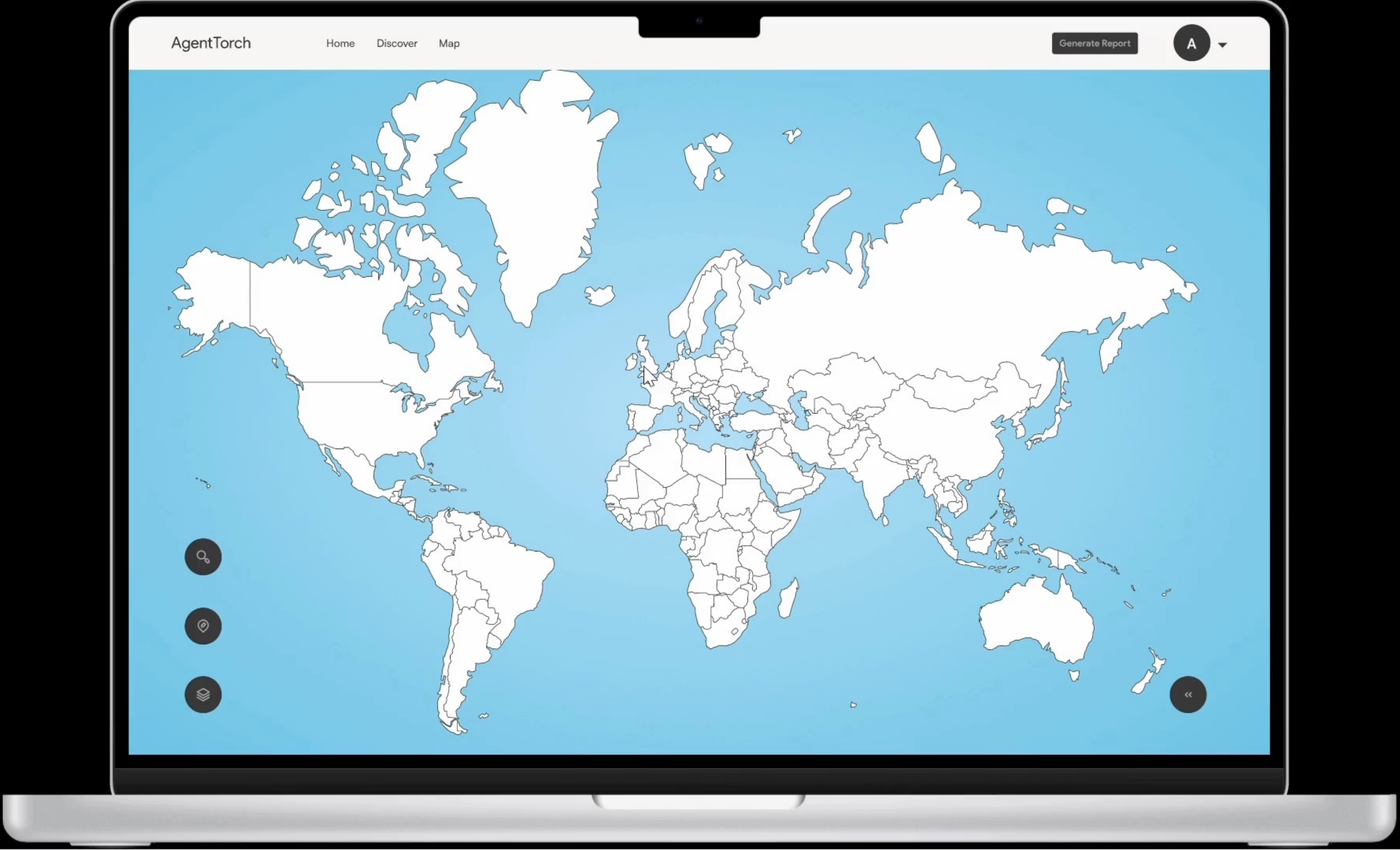

AgentTorch
Home
Discover
Map
Generate Report
A

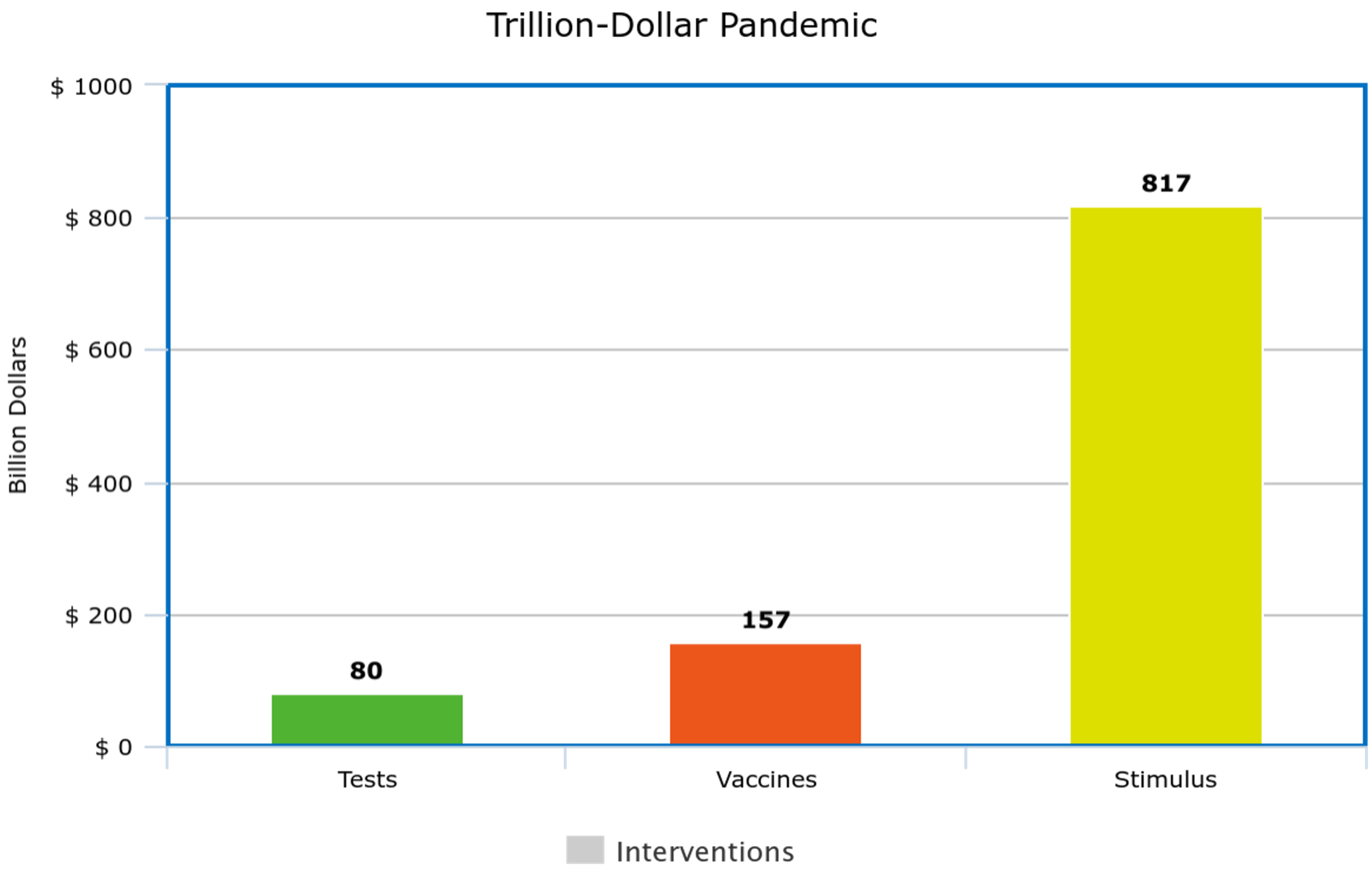
The H5N1 Crisis



Human Disease Transmission?




AgentTorch
Home
Discover
Map
Generate Report
A


Trillion-Dollar Pandemic
Billion Dollars
$ 1000
$ 800
$ 600
$ 400
$ 200
$ 0
80
157
817
Tests
Vaccines
Stimulus
Interventions

"What kind of a test do you want?"

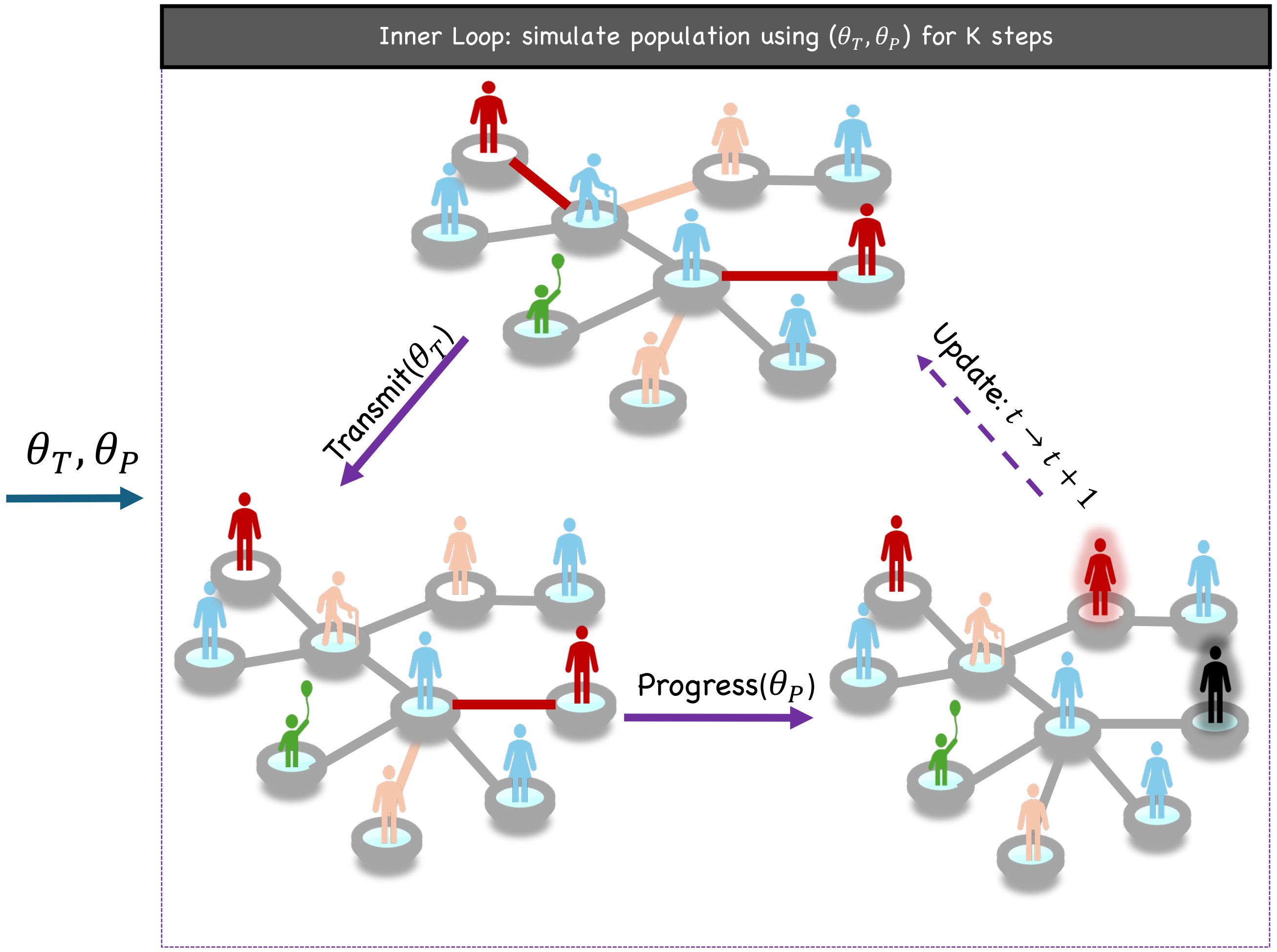
10% false-positive
3 day result

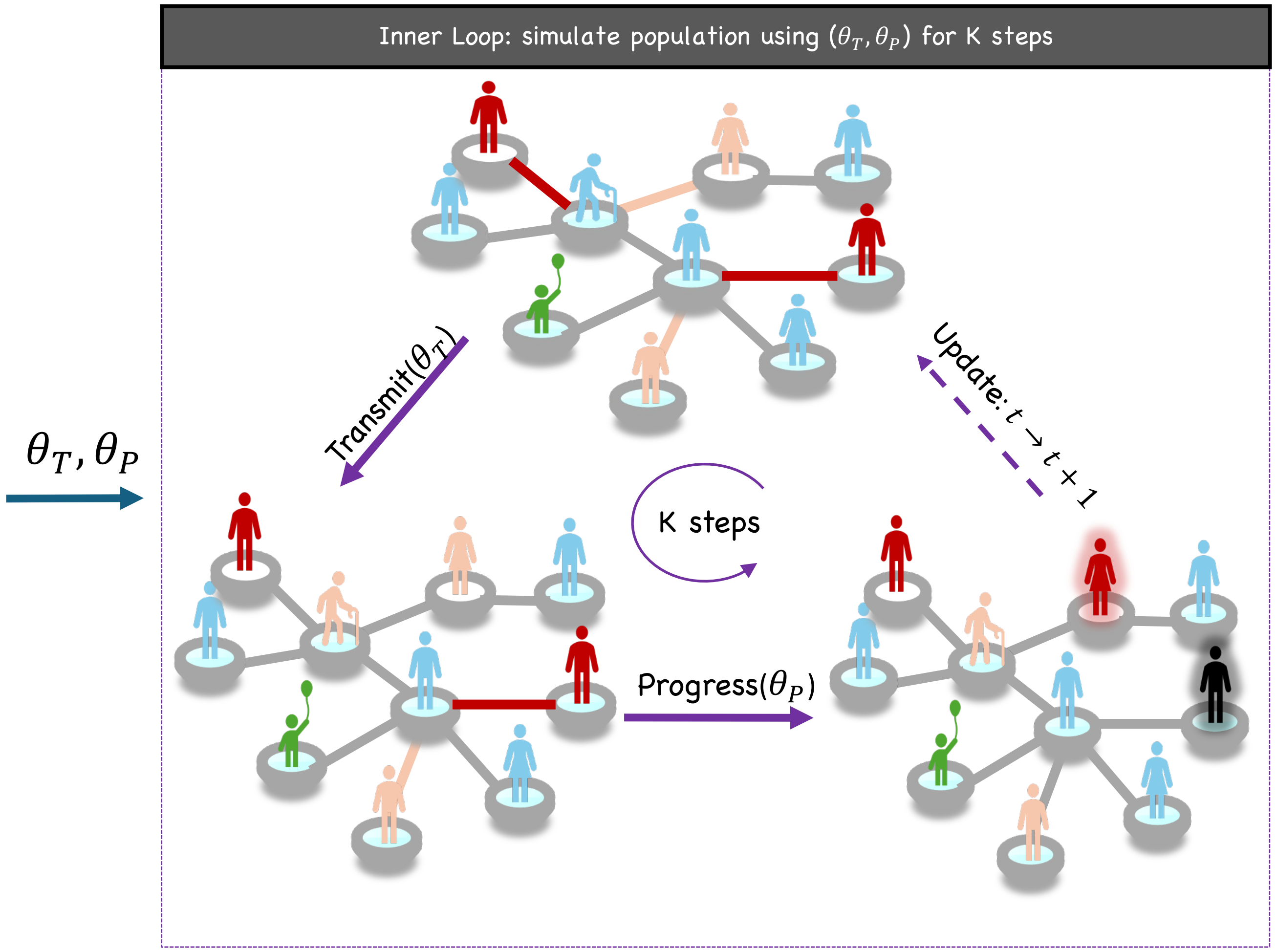
40% false positive
1 day result

New York City

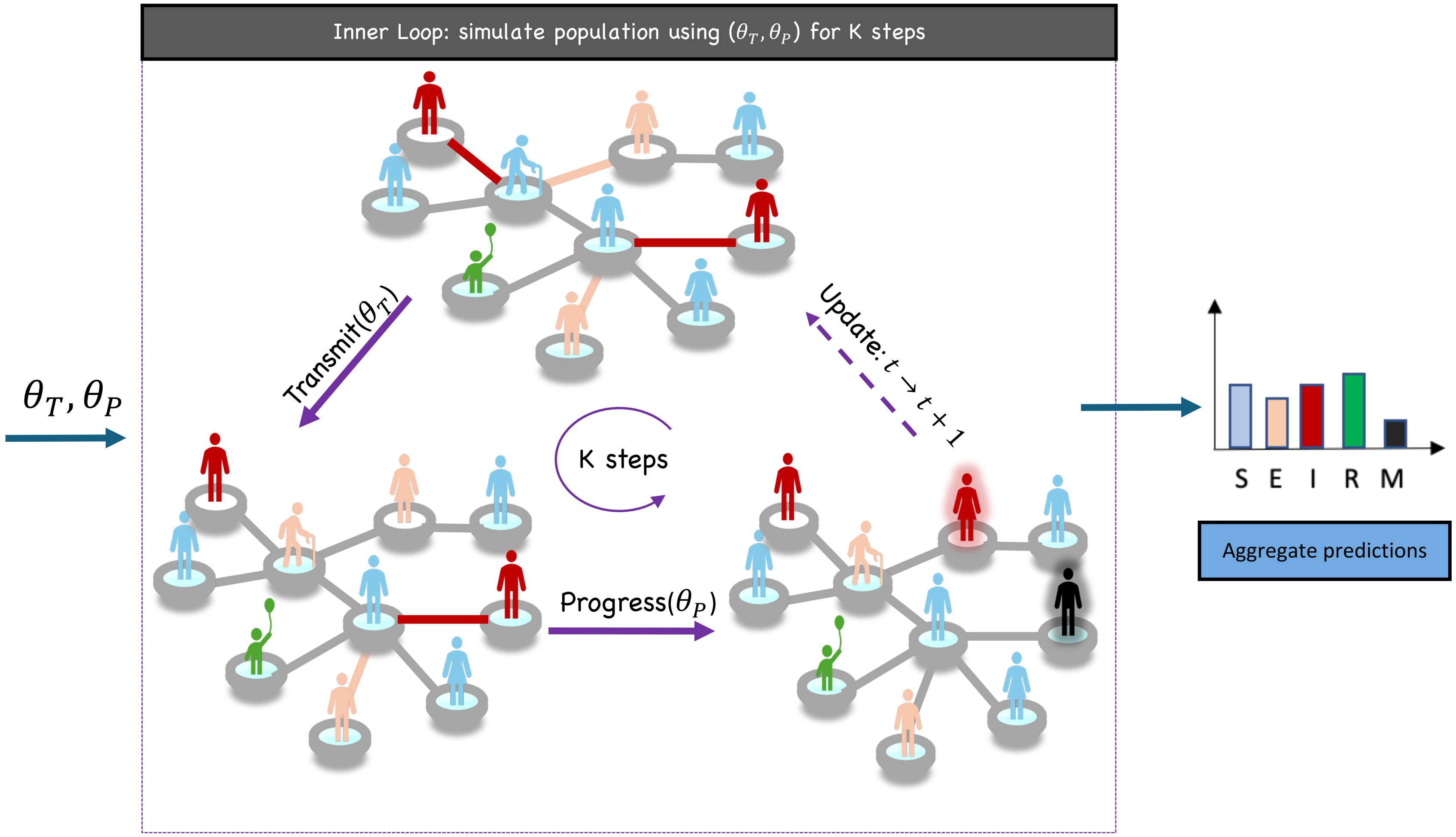
- 5 boroughs
- 8.4 million people
- 3.3 million households
- 200,000 small businesses

2020-2022


Inner Loop: simulate population using (θ_T, θ_P) for K steps
θ_T, θ_P
Transmit(θ_T)
Update: $t \rightarrow t+1$
Progress(θ_P)

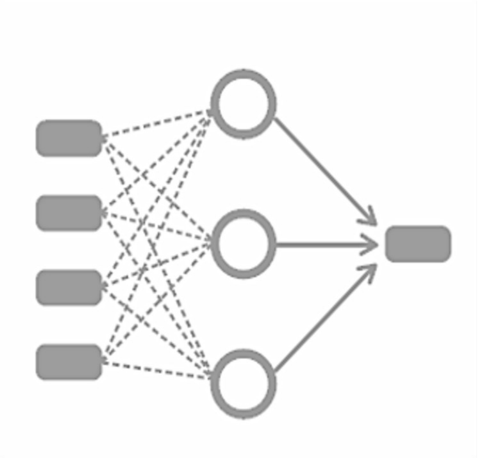


Inner Loop: simulate population using (θ_T, θ_P) for K steps
θ_T, θ_P
Transmit(θ_T)
Update: $t \rightarrow t+1$
K steps
Progress(θ_P)



Inner Loop: simulate population using (θ_T, θ_P) for K steps
θ_T, θ_P
Transmit(θ_T)
Update: $t \rightarrow t+1$
K steps
Progress(θ_P)
S E I R M
Aggregate predictions

LPMs for Public Health



Capture	Analyze	Act
Disease Transmission Viral Evolution	Wastewater Signal Compliance Behavior	Test: Speed vs Accuracy? $1M: how Vaccinate?
Simulate disease spread across 8 million agents	Calibrate to disease and intervention data	Optimize policies while preserving privacy

LPMs for Public Health

Capture	Analyze	Act
Disease Transmission Viral Evolution	Wastewater Signal Compliance Behavior	Test: Speed vs Accuracy? $1M: how Vaccinate?
Simulate disease spread across 8 million agents	Calibrate to disease and intervention data	Optimize policies will preserving privacy

Capture

Disease Transmission
Viral Evolution

Simulate disease spread across 8 million agents

Get Started in 3 lines of code

Custom Templates, High-res visualizations

```
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')
```

Capture

Disease Transmission
Viral Evolution

Simulate disease spread across 8 million agents

Compose: LLMs

Scale LLMs to prompt millions of 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()
```

Capture

Disease Transmission
Viral Evolution

Simulate disease spread across 8 million agents

Compose: ODEs

Simulate across multiple time scales!

```python
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

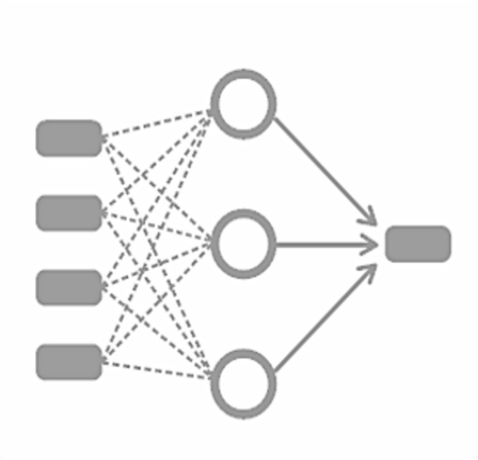
Simulate disease spread across 8 million agents

Implement Stochastic and Conditional Policies

Ensure Differentiability via custom operators

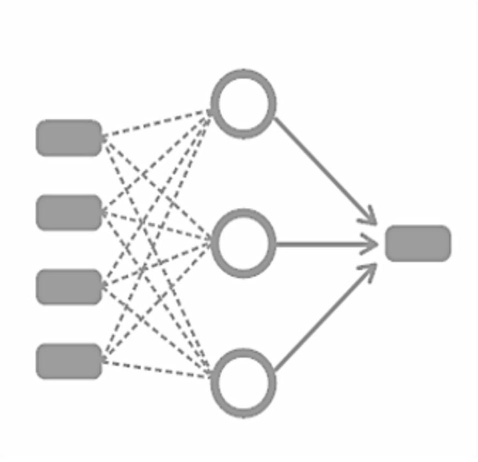
```
from torch.distributions import Bernoulli
from agent_torch.distributions import Bernoulli as AT_Bernoulli

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

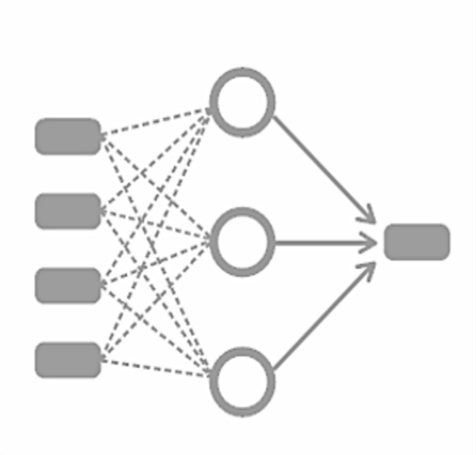


Capture	Analyze	Act
Disease Transmission Viral Evolution	Wastewater Signal Compliance Behavior	Test: Speed vs Accuracy? $1M: how Vaccinate?
Simulate disease spread across 8 million agents	Calibrate to disease and intervention data	Optimize policies while preserving privacy

LPMs for Public Health



Capture	Analyze	Act
Disease Transmission Viral Evolution	Wastewater Signal Compliance Behavior	Test: Speed vs Accuracy? $1M: how Vaccinate?
Simulate disease spread across 8 million agents	Calibrate to disease and intervention data	Optimize policies will preserving privacy



Analyze

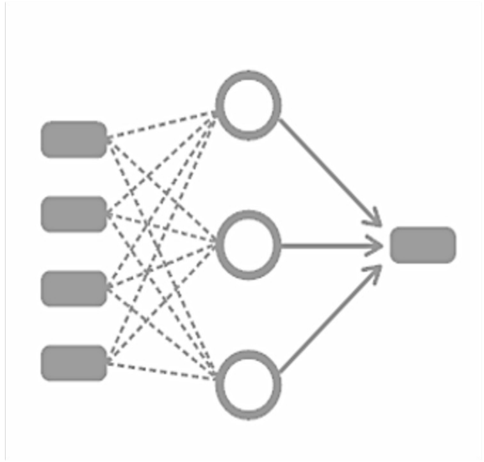
Wastewater Signal
Compliance Behavior

Calibrate to disease and intervention data

Compose end-to-end with NNs

Calibrate the simulator instead of a surrogate!

```python
# 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()
```



Analyze

Wastewater Signal
Compliance Behavior

Calibrate to disease and intervention data

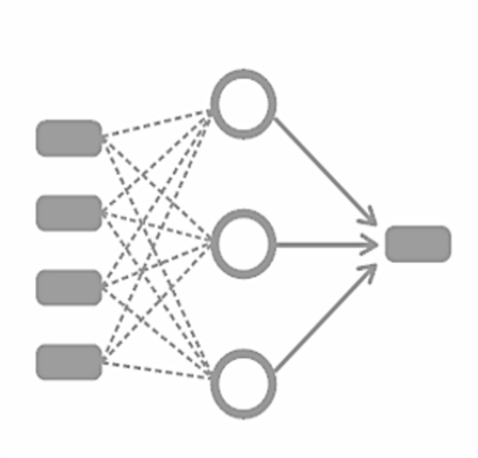
AgentTorch simulators can be calibrated directly, *without* a surrogate

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	Analyze	Act
Disease Transmission Viral Evolution	Wastewater Signal Compliance Behavior	Test: Speed vs Accuracy? $1M: how Vaccinate?
Simulate disease spread across 8 million agents	Calibrate to disease and intervention data	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

Act

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

Optimize policies while preserving privacy

Optimize Policies in Natural Language

Custom analyzers for AgentTorch API

```python
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 is isolation time is increased and virus become more transmissible?")
```

Act

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

Optimize policies while preserving privacy

Deploy decision-making to the edge

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()
```

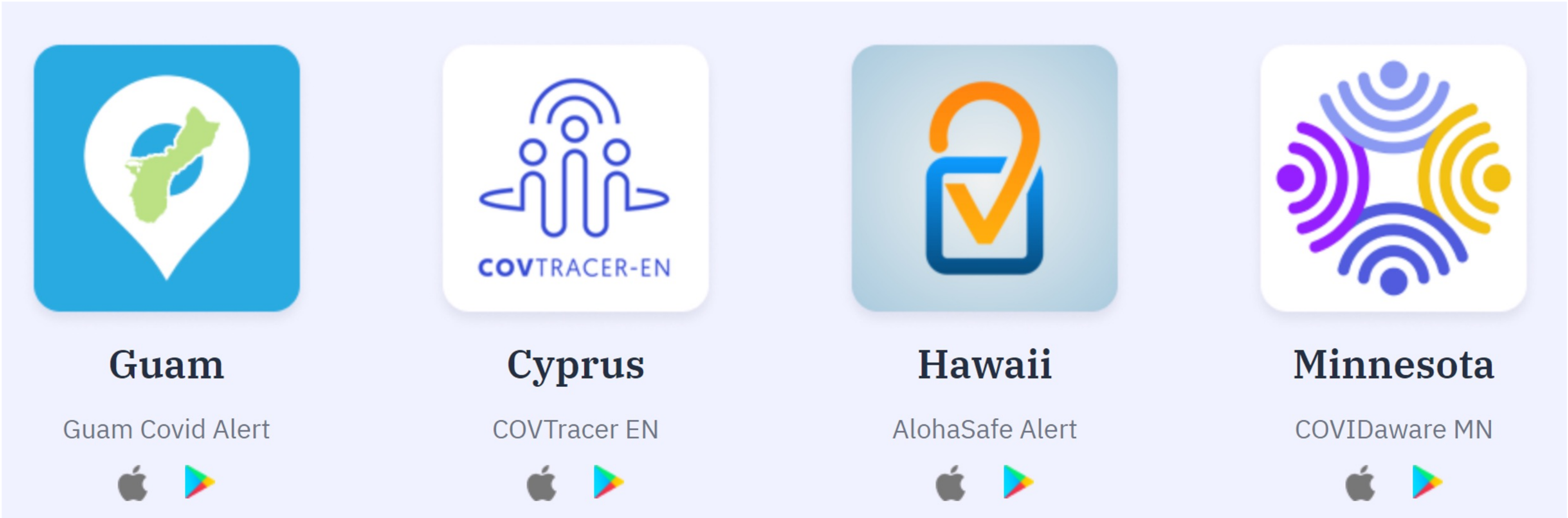
Act

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



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

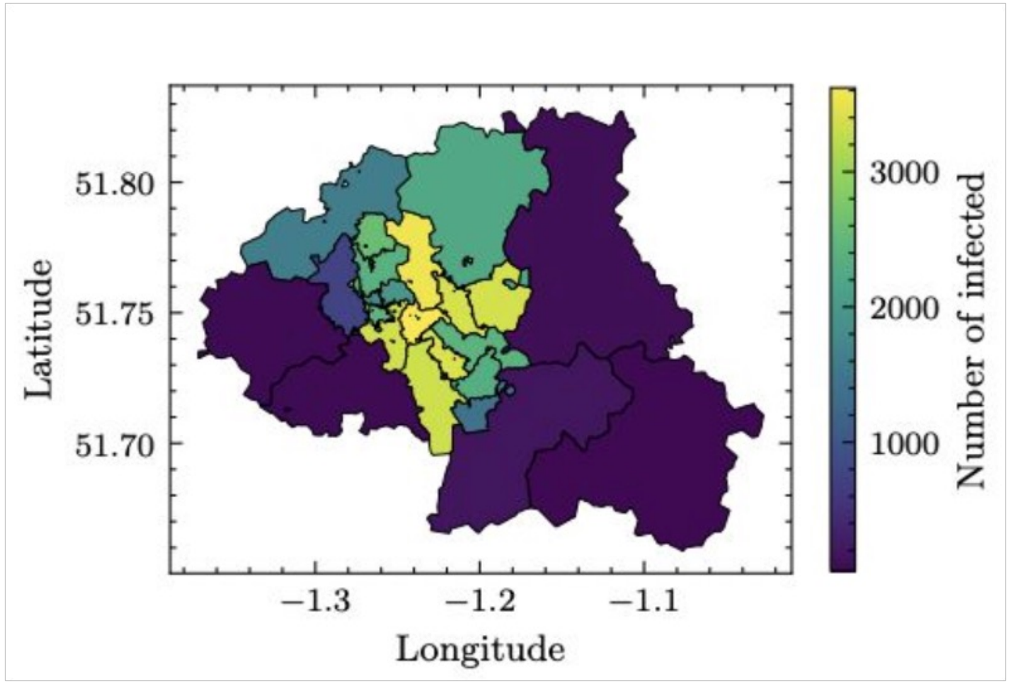
Act

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

Optimize policies while preserving privacy

Close Sim2Real Gap: Contact Intelligence

Synchronize simulation and protocol for real-time and private computation

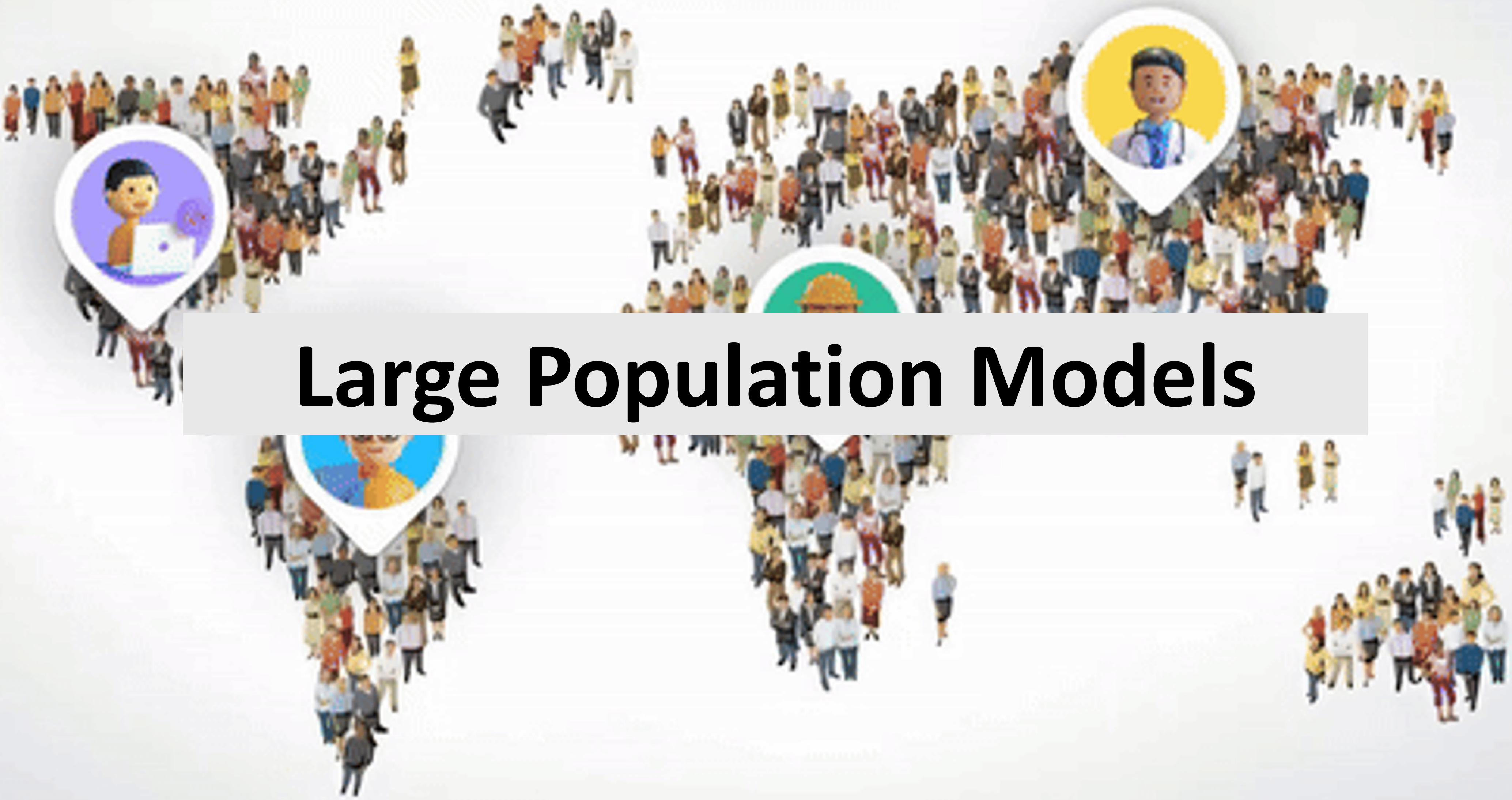


Track disease without leaking individual demographic, disease status or geolocation

AgentTorch

Imagine… your enterprise a living lab

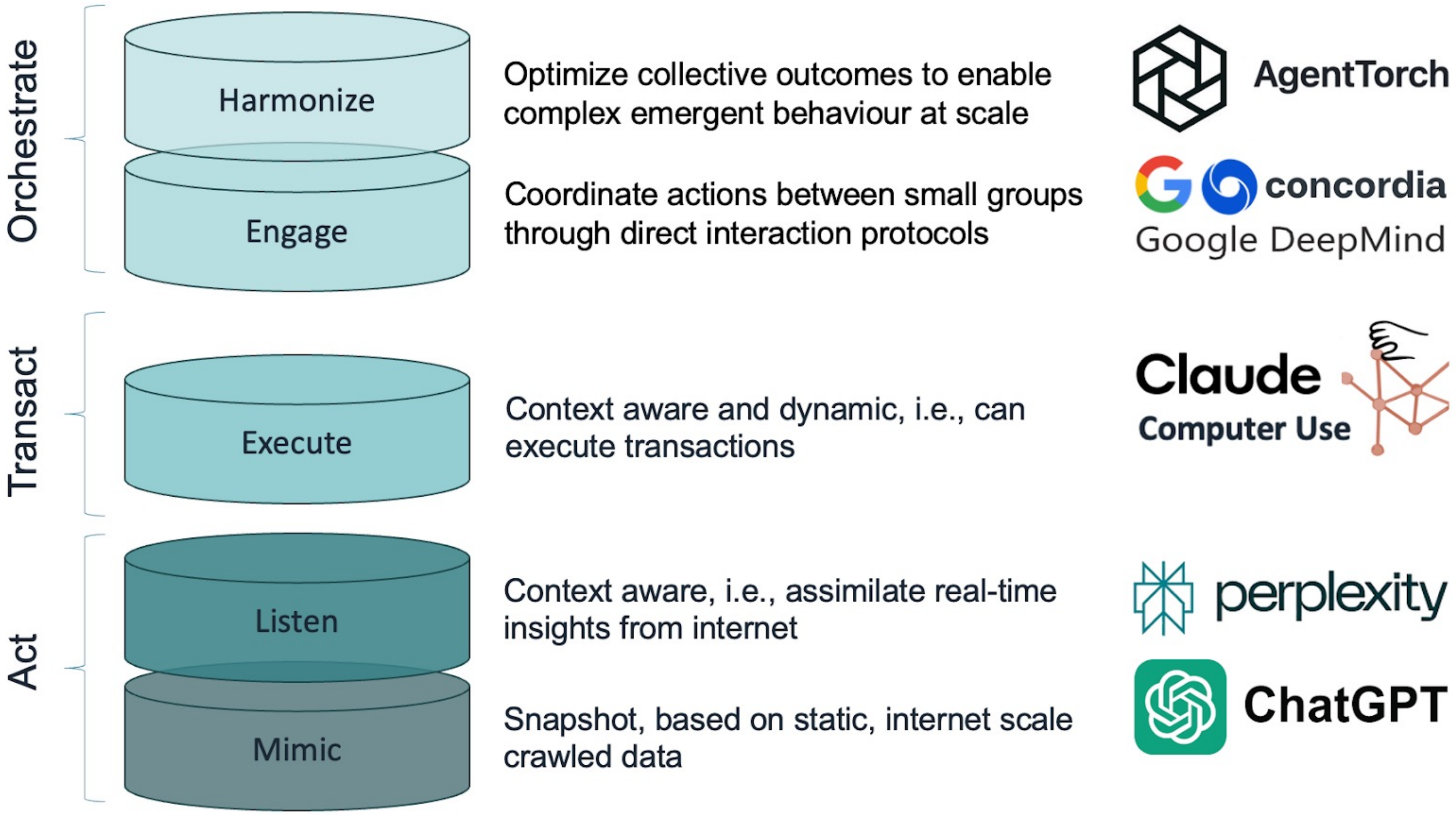
Large Population Models



lpm.media.mit.edu

Appendix

Levels of Agentic Systems

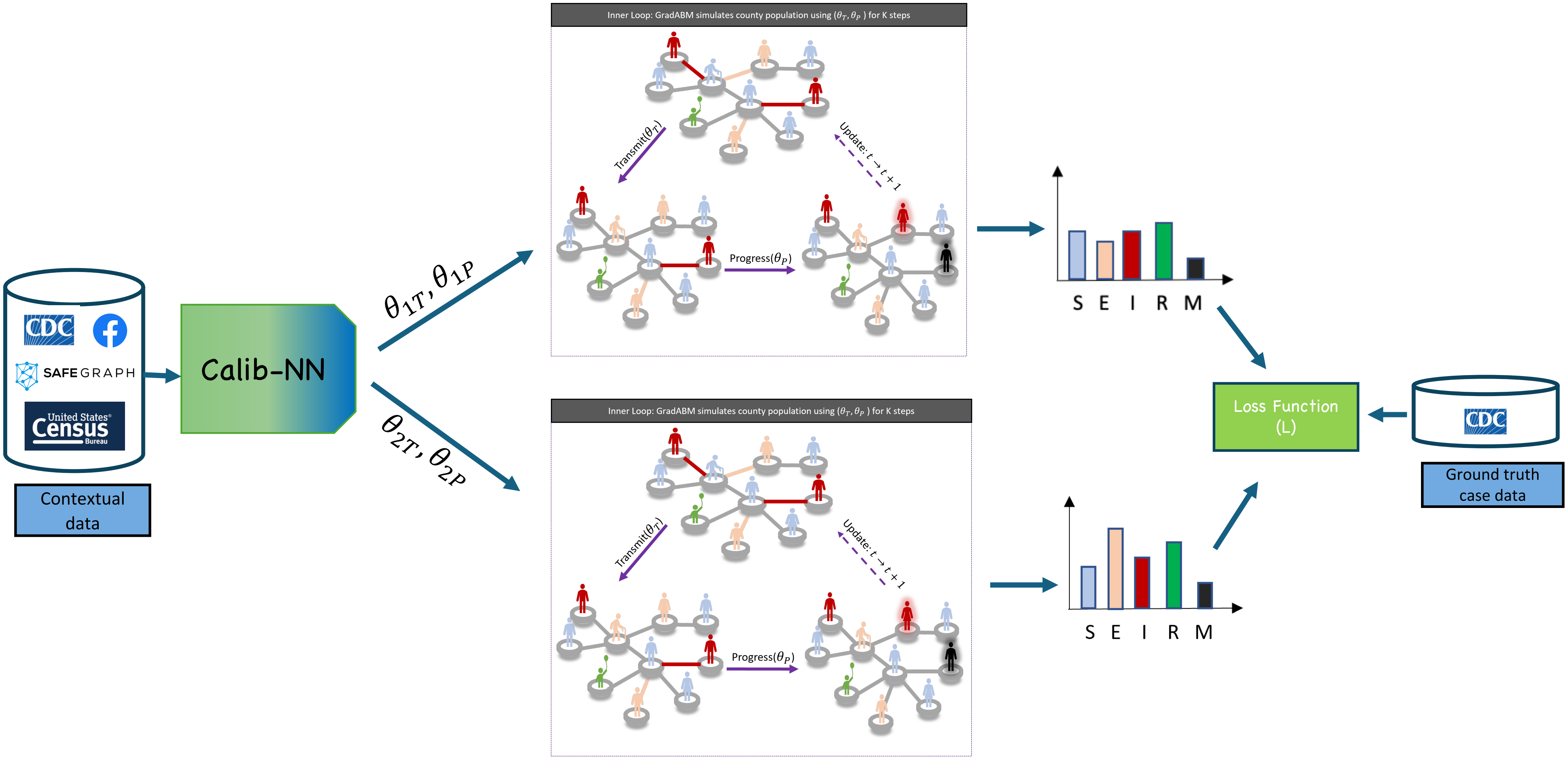


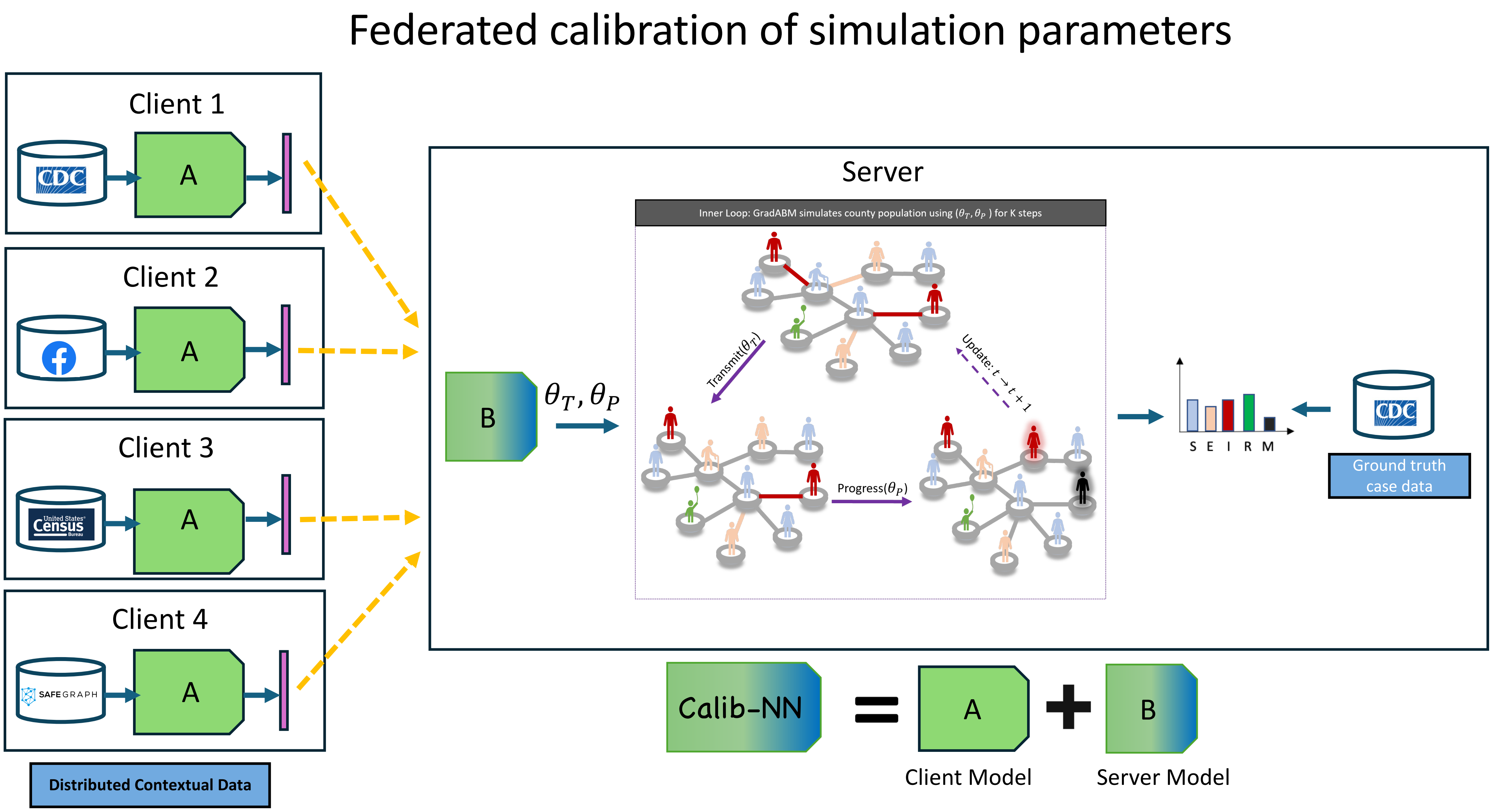
Comparison with other frameworks

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

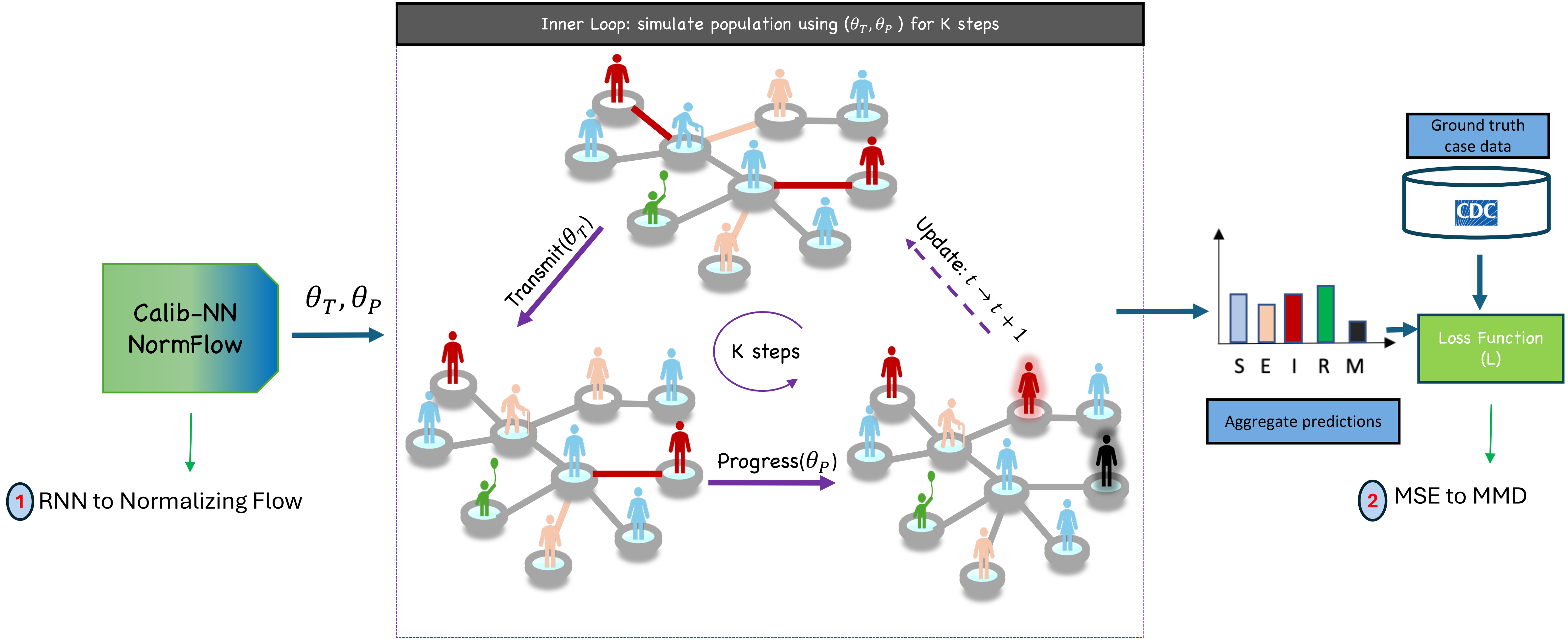
Data for Differentiable Learning

Multi-task learning over simulation environments, reduce overfitting

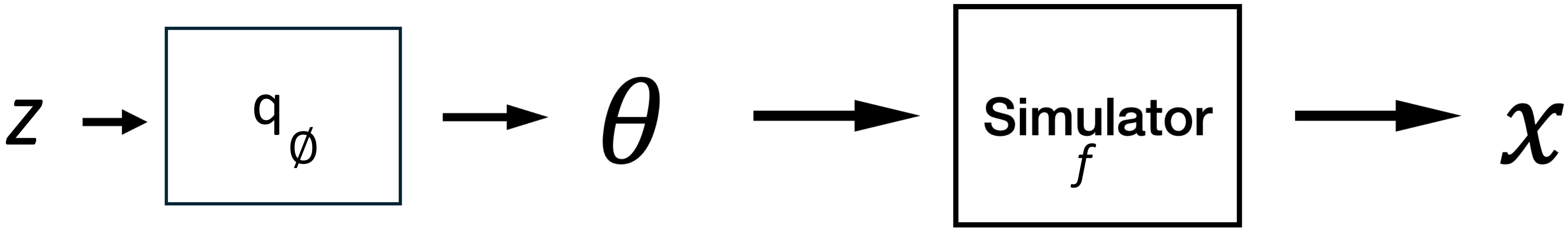


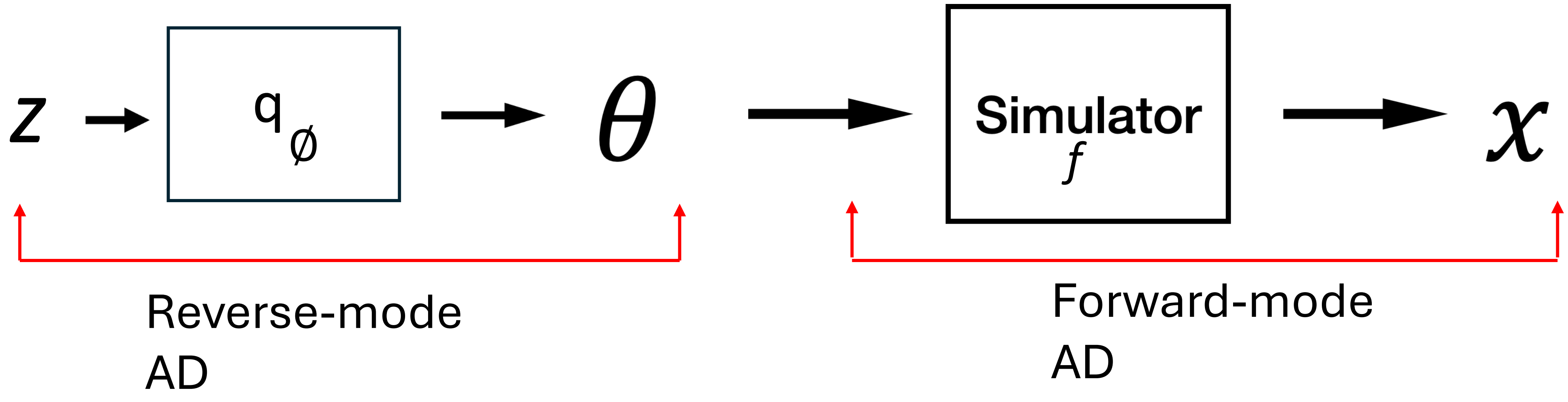

Federated calibration of simulation parameters
Client 1
A
Client 2
A
Client 3
A
Client 4
A
Distributed Contextual Data
Server
Inner Loop: GradABM simulates county population using (θ_T, θ_P) for K steps
Transmit(θ_T)
Update: $t \to t+1$
Progress(θ_P)
B
θ_T, θ_P
S E I R M
Ground truth case data
Calib-NN
A
B
Client Model
Server Model

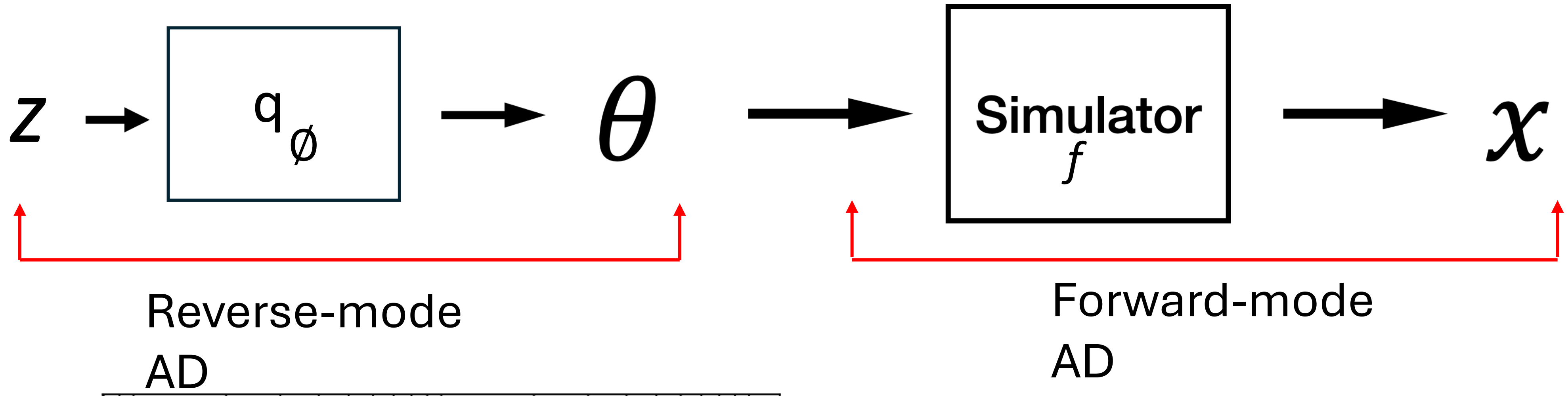
Estimate posteriors over simulation parameters

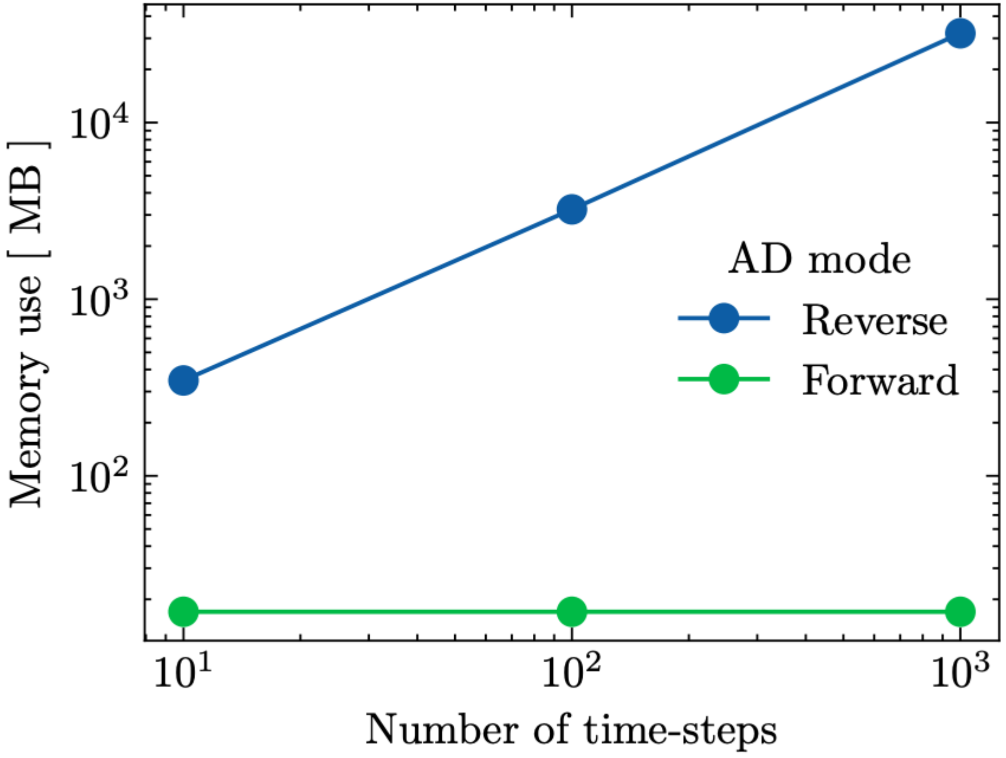


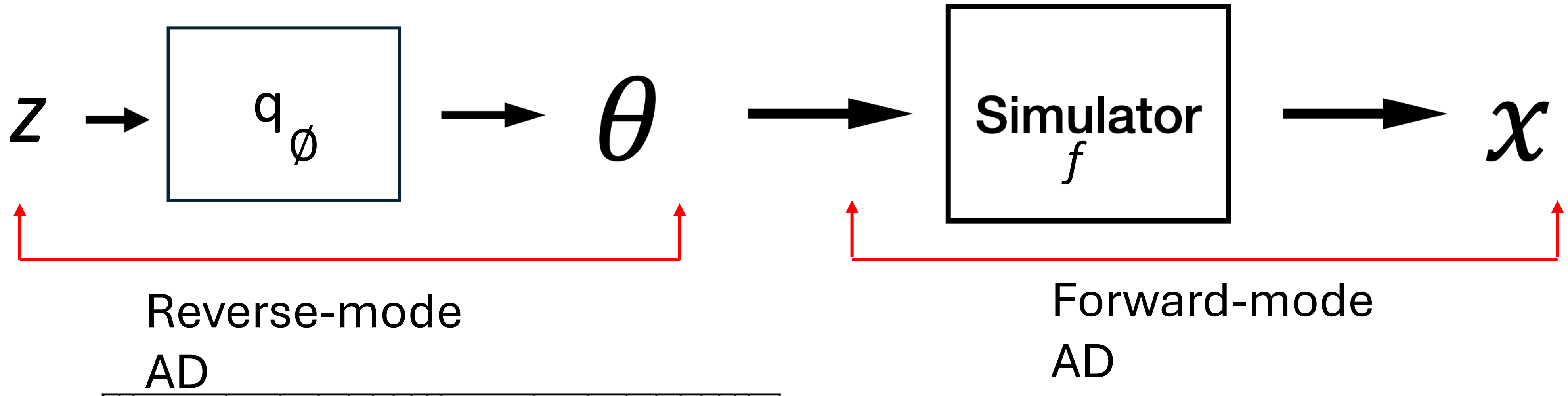
Advanced Differentiable Learning: Memory Optimization

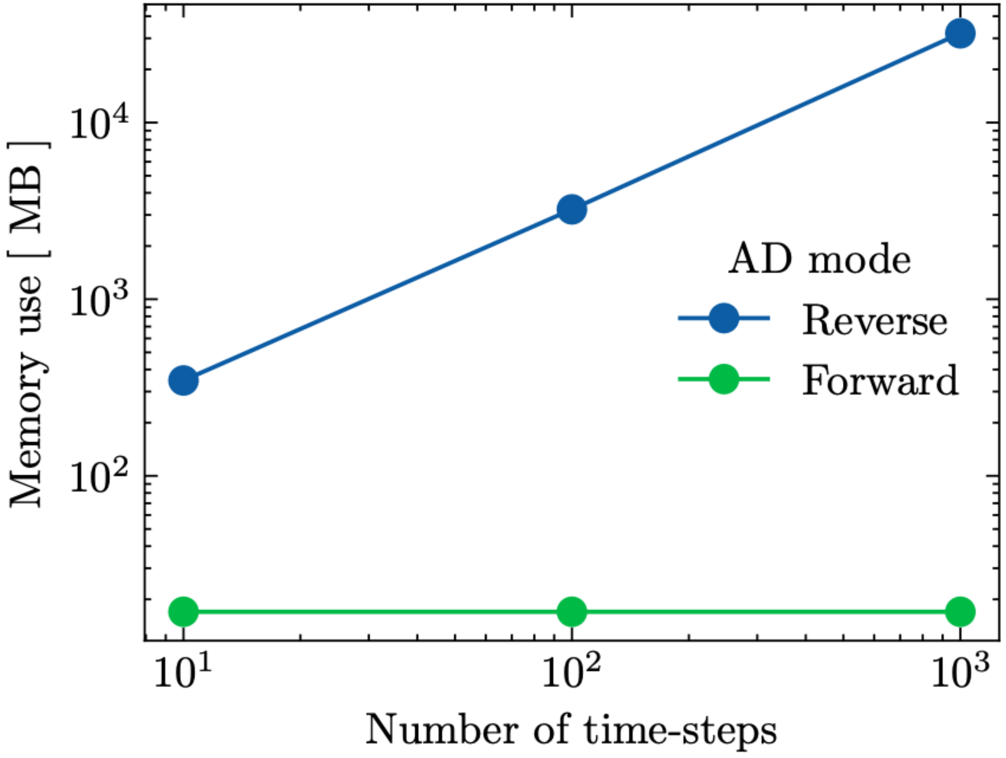

z
q
ø
θ
Simulator
f
x


z
q
ø
θ
Simulator
f
x
Reverse-mode
AD
Forward-mode
AD


z
q_ϕ
θ
Simulator
f
x
Reverse-mode AD
Forward-mode AD


Memory use [MB]
10^2
10^3
10^4
Number of time-steps
10^1
10^2
10^3
AD mode
Reverse
Forward





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

$$J_f = \frac{\partial f(\theta)}{\partial \theta} \qquad g_q = \frac{\partial q_\phi(z)}{\partial \phi}$$

Advanced Differentiable Learning: Zero-shot Sensitivity Analysis

Sensitivity Analysis is critical for validation

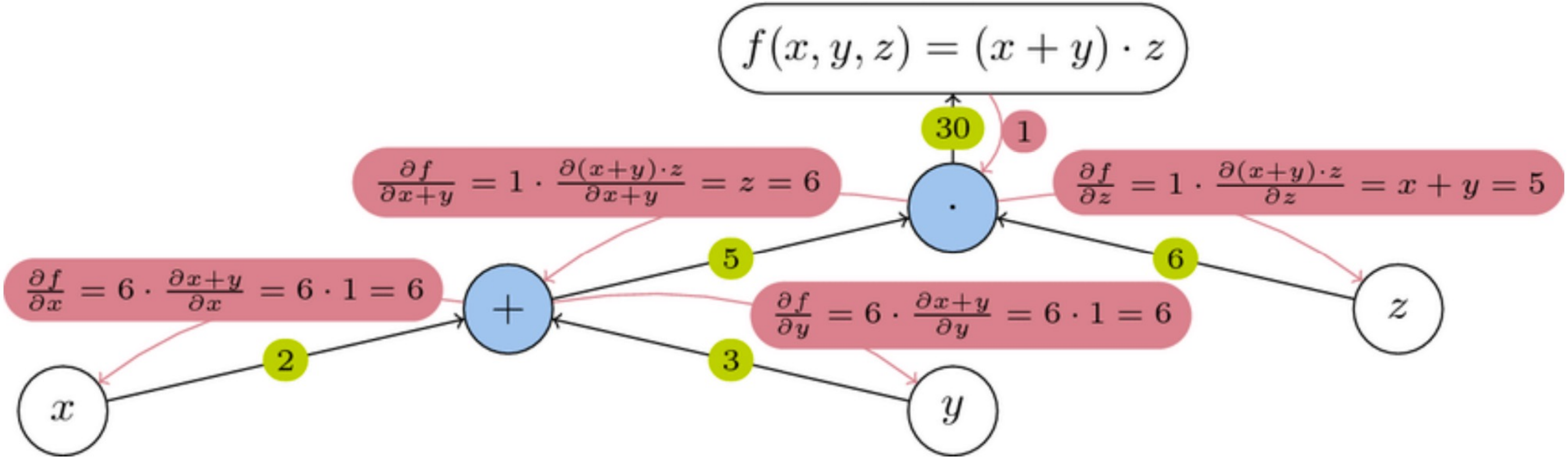
The impact of uncertainty on predictions of the CovidSim epidemiological code

Wouter Edeling[1], Hamid Arabnejad[2], Robbie Sinclair[3], Diana Suleimenova[2], Krishnakumar Gopalakrishnan[3], Bartosz Bosak[4], Derek Groen[2], Imran Mahmood[2], Daan Crommelin[1,5] and Peter V. Coveney[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.

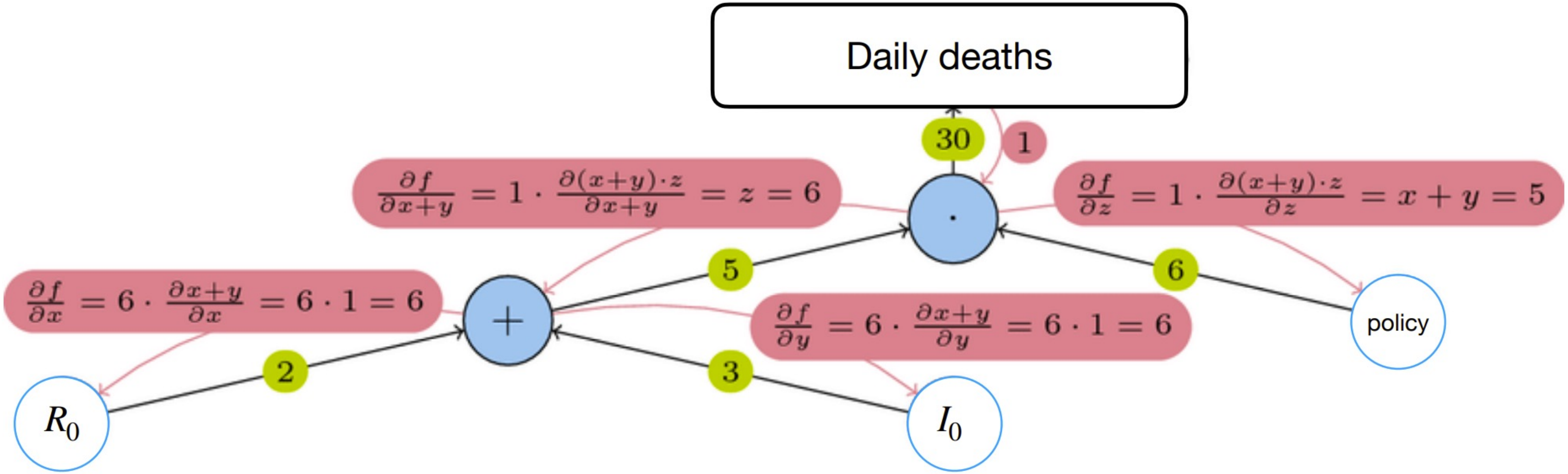
Ensemble execution. Consequently, through the use of adaptive methods we make the uncertainty analysis of CovidSim tractable, but our analysis nevertheless required us to perform thousands of runs, each with its own unique set of input parameters. Specifically, we used the Eagle supercomputer at the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

Recap: Reverse-mode automatic differentiation



Sensitivity analysis via reverse-mode automatic differentiation

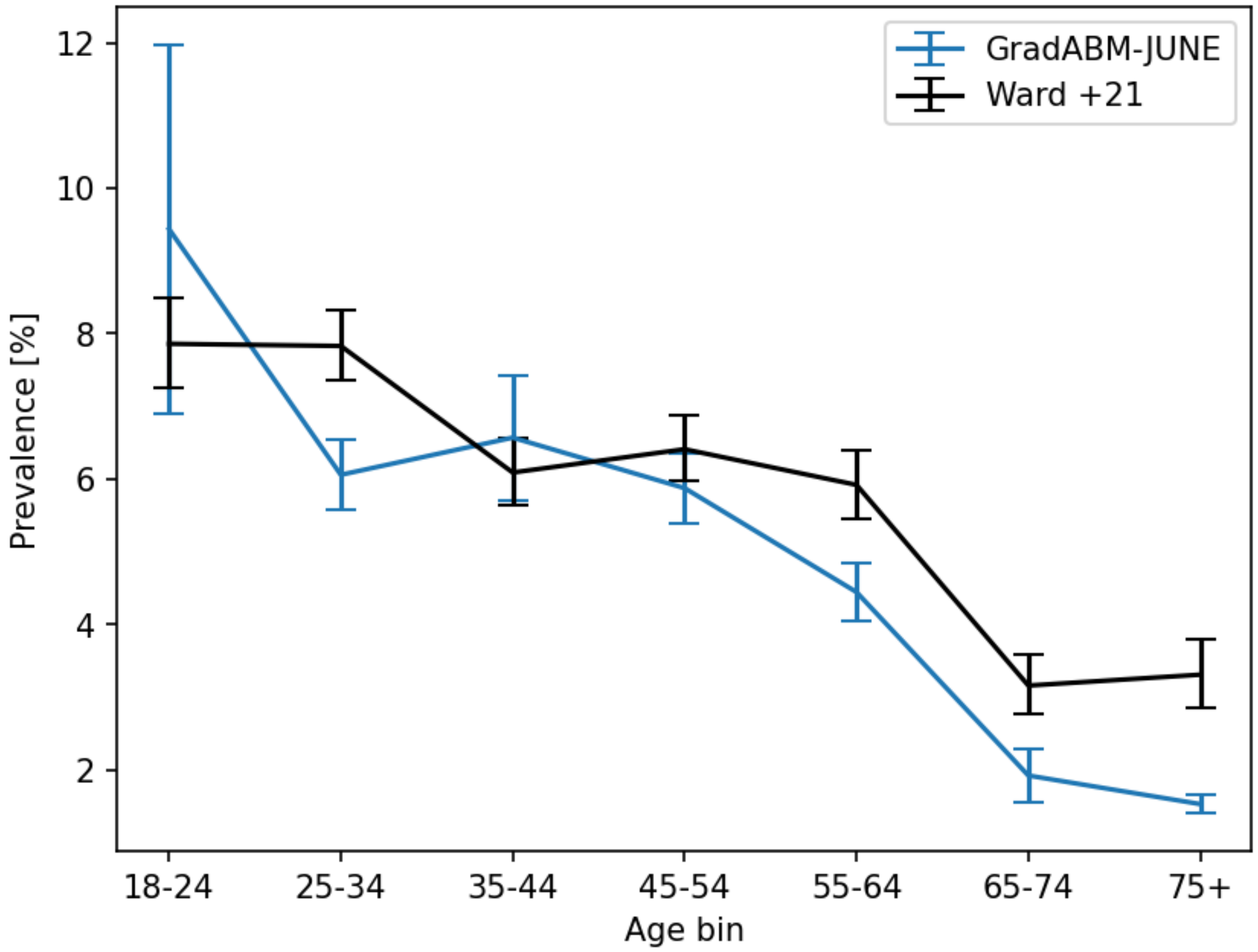
Reverse-mode automatic differentiation is independent of the number of parameters!!



Quera Bofarull et al

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!