CONTAGIONS OVER MULTIPLEX NETWORKS: THE ROLE OF AI AND COMPUTING

MADHAV MARATHE
& Biocomplexity Institute and Initiative Team
Network Systems Science and Advanced Computing Division (NSSAC)
Biocomplexity Institute and Initiative (BII)
& Dept. of Computer Science, University of Virginia
Email: marathe@virginia.edu
Phone: 540 808 6220

BII-TR#-2021-132

UNIVERSITY of VIRGINIA

BIOCOMPLEXITY INSTITUTE

Support:
NIH R01GM109718, NSF Grant CINES Grant: OAC-1916805, NSF DIBBS Grant ACI-1443054 and NSF BIG DATA Grant IIS-1633028.NSF Expeditions in Computing Grant CCF-1918656, NSF Virtual organization grant, CNS 2041952, NSF Eager CCF-1917819, US Centers for Disease Control and Prevention 75D30119C05935 and a collaborative seed grant from the UVA Global Infectious Disease Institute.

Thanks to team members of the Biocomplexity COVID-19 Response team & NSSAC, UVA research computing, Shawn Brown, John Towns, Tom Maiden and our partners at DTRA, CDC, VDH, NSF & Scenario Modeling Hub. We also sincerely thank the Scenario Modeling Hub Coordination team and in particular, Cecile Viboud, Katriona Shea, Michael Runge, Rebecca Borchering, Justin Lessler & Shaun Truelove for their incredible work on bringing the teams together.

Population Growth, Megacities & Small-World Effect

By 2050: 27 Megacities in the world with 10M+ people (up from 19 in 2022)

Multiplexed networks and the society they support co-evolve

Creates unprecedented opportunities. They also pose potential risks to the global environment, resilience, social stability, the national development of poorer countries and can result in inequitable allocation of resources putting the marginalized groups at further risk.

A brief history

1990 1995 2000 2005 2010 2015 2020 2025

TRANSIMS Program (1992-2001):
Urban transport planning

NISAC Program (2002-19): Interdependent infrastructure analysis for disaster response

NIH MIDAS Program (2004-present): Pandemic planning and response

DoD CNIMS program (2005-18):
Responding to WMD

IARPA OSI program (2013-17): Prediction of societal events using open-source indicators

NSF Expeditions Program (2020-present): Pandemic planning and response

NSF CINES Program (2019-present): Cyber-infrastructure for network science

We have been developing high performance computing-based based models to study complex multi-networks since 1992

How can one detect, assess and reduce the impact of a pandemic outbreak such as COVID-19?

https://covid19.biocomplexity.virginia.edu/

References

- *Papers:*
  - *CACM 2014:* Computational Epidemiology
  - *Nature 2004*: Modelling disease outbreaks in realistic urban social networks.
  - *PNAS 2008:* Modeling targeted layered containment of an influenza pandemic in the United States.
  - *PNAS 2014*: Opinion: Mathematical models: A key tool for outbreak response.
- Tutorials (KDD'14, AAAI'16, ICSB'17): https://covid19.biocomplexity.virginia.edu/publications
- COVID-19 resource page: https://covid19.biocomplexity.virginia.edu/
- New NSF Expeditions Project: https://computational-epidemiology.org

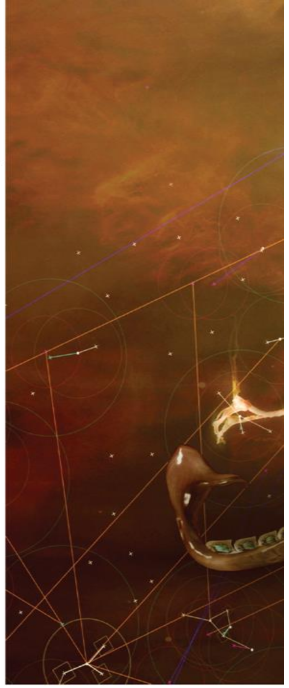
review articles

DOI:10.1145/2483852.2483871

The challenge of developing and using computer models to understand and control the diffusion of disease through populations.

BY MADHAV MARATHE AND ANIL KUMAR S. VULLIKANTI

Computational Epidemiology

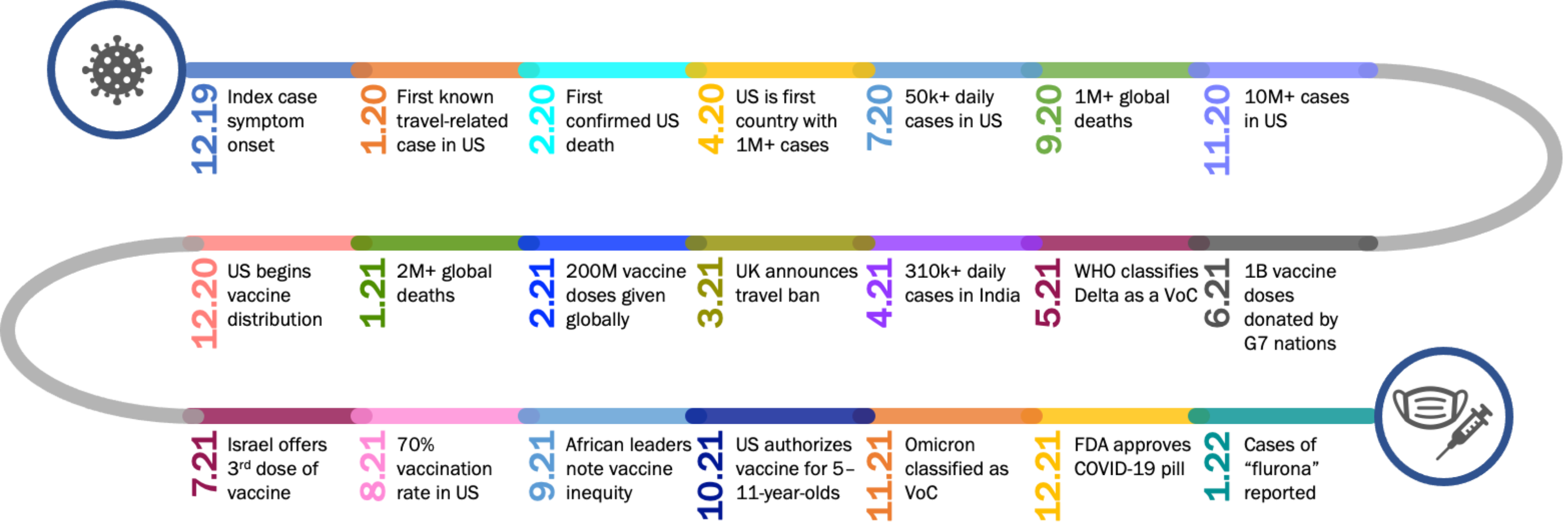


AN EPIDEMIC IS said to arise in a community or region when cases of an illness or other health-related events occur in excess of normal expectancy. Epidemics are considered to have influenced significant historical events, including the plagues in Roman times and Middle Ages, the fall of the Han empire in the 3rd century in China, and the defeat of the Aztecs in the 1500s, due to a smallpox outbreak.[9] The 1918 flu pandemic in the U.S. was responsible for more deaths than those due to World War I. The last 50 years have seen epidemics caused by HIV/AIDS, SARS, and influenza-like illnesses. Despite significant medical advances, according to the World Health Organization (WHO), infectious diseases account for more than 13 million deaths a year.[44]

Societal interest in controlling outbreaks is probably just as old as the diseases themselves. Interestingly, it appears the Indians and Chinese knew the idea of variolation to control smallpox as early as the 8th century A.D. Epidemiology is a formal branch of science focusing on the study of space-time patterns of illness in a population and the factors that contribute to these patterns. It plays an essential role in public health by

» key insights

- Controlling and responding to future pandemics will be challenging due to a number of emerging global trends including increased and denser urbanization, increased local as well as global travel, and a generally older and immuno-compromised population.
- Public health epidemiology is a complex system problem. Epidemics, social-contact networks, individual and collective behavior, and public policies coevolve during a pandemic—a system-level understanding must represent these components and their coevolution.
- Mathematical and computational models of social networks and epidemic spread and methods to analyze them are critical in public health epidemiology.
- Advances in computing, big data, and computational thinking have created entirely new opportunities to support real-time epidemiology.



COVID-19 : How it happened and how did society respond

Why did it become a pandemic

- Zoonosis
- Global connectivity where zoonosis happened
- Silent Spreaders
- Initial response was slow & not decisive (always hard)

Socio political & economic response

- Global unprecedented lock down & travel restrictions
- Trillions of dollars worth of economic stimulus
- Political disagreements led to uncoordinated and decisive response

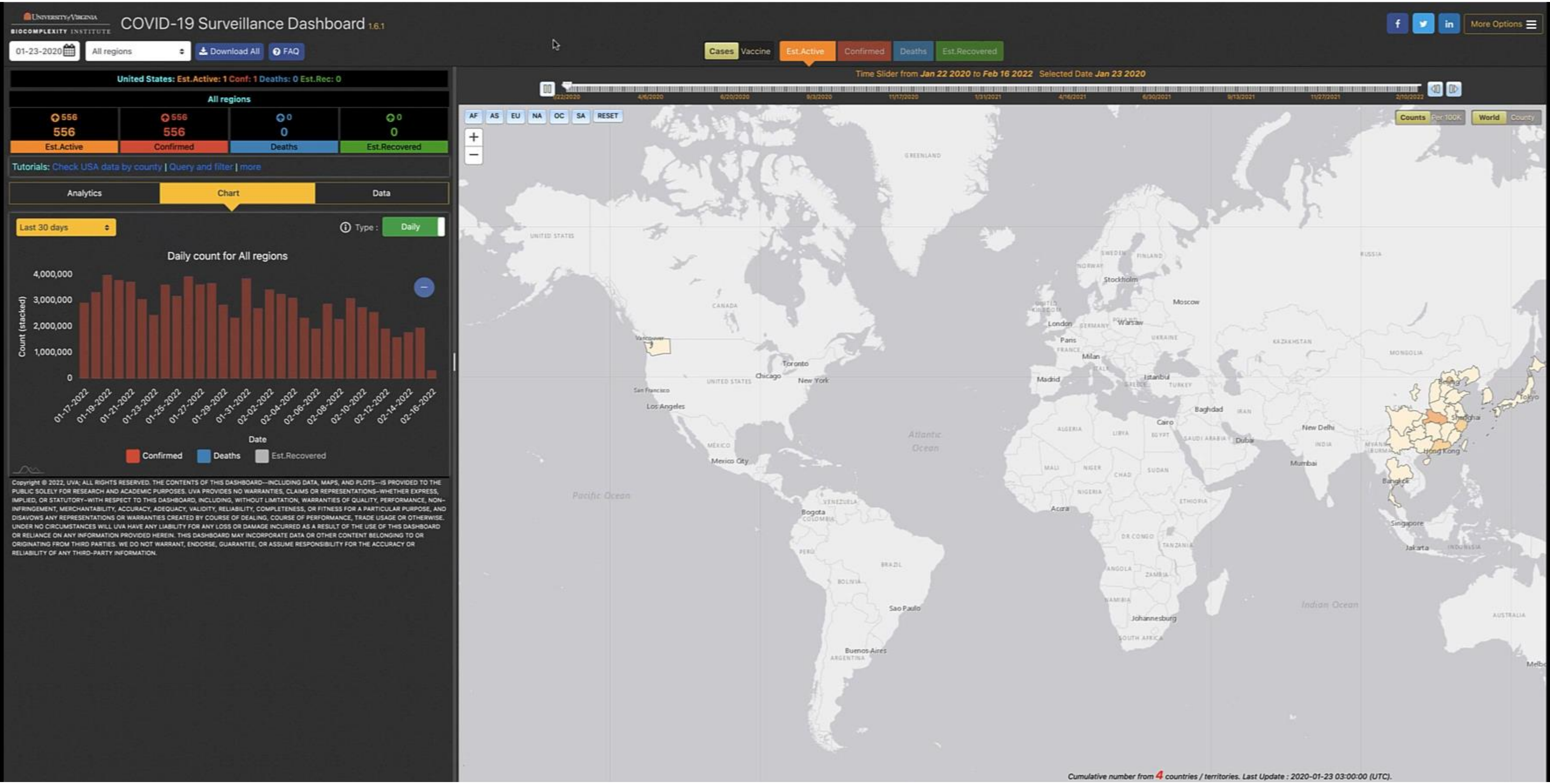
Scientific and medical response

- 530,000 Preprints from 03/20-09/21
- New treatments (anti-body, plasma)
- New treatment protocols
- Excellent vaccines in a year

Technological response

- New apps for contact tracing
- Genome sequencing completed in record time
- Innovative ways to work, educate, deliver, interact
- HPC consortium formed to support modeling

https://en.wikipedia.org/wiki/Timeline_of_the_COVID-19_pandemic;
https://www.science.org/content/article/no-revolution-covid-19-boosted-open-access-preprints-are-only-fraction-pandemic-papers
https://covid19primer.com/dashboard
https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines

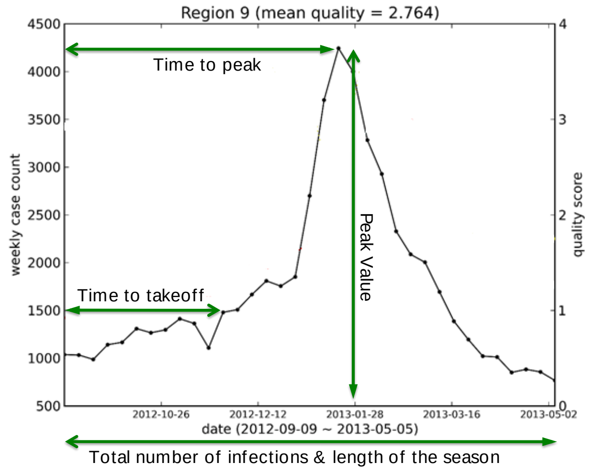

COVID-19 Surveillance Dashboard 1.6.1
01-23-2020
All regions
Download All
FAQ
Cases
Vaccine
Est.Active
Confirmed
Deaths
Est.Recovered
More Options
United States: Est.Active: 1 Conf: 1 Deaths: 0 Est.Rec: 0
All regions
556
556
Est.Active
556
556
Confirmed
0
0
Deaths
0
0
Est.Recovered
Tutorials: Check USA data by county | Query and filter | more
Analytics
Chart
Data
Last 30 days
Type :
Daily
Daily count for All regions
Count (stacked)
Date
Confirmed
Deaths
Est.Recovered
Time Slider from Jan 22 2020 to Feb 16 2022 Selected Date Jan 23 2020
AF
AS
EU
NA
OC
SA
RESET
Counts
Per 100K
World
County
Atlantic Ocean
Pacific Ocean
Indian Ocean
Cumulative number from 4 countries / territories. Last Update : 2020-01-23 03:00:00 (UTC).
Copyright © 2022, UVA; ALL RIGHTS RESERVED. THE CONTENTS OF THIS DASHBOARD—INCLUDING DATA, MAPS, AND PLOTS—IS PROVIDED TO THE PUBLIC SOLELY FOR RESEARCH AND ACADEMIC PURPOSES. UVA PROVIDES NO WARRANTIES, CLAIMS OR REPRESENTATIONS–WHETHER EXPRESS, IMPLIED, OR STATUTORY–WITH RESPECT TO THIS DASHBOARD, INCLUDING, WITHOUT LIMITATION, WARRANTIES OF QUALITY, PERFORMANCE, NON-INFRINGEMENT, MERCHANTABILITY, ACCURACY, ADEQUACY, VALIDITY, RELIABILITY, COMPLETENESS, OR FITNESS FOR A PARTICULAR PURPOSE, AND DISAVOWS ANY REPRESENTATIONS OR WARRANTIES CREATED BY COURSE OF DEALING, COURSE OF PERFORMANCE, TRADE USAGE OR OTHERWISE. UNDER NO CIRCUMSTANCES WILL UVA HAVE ANY LIABILITY FOR ANY LOSS OR DAMAGE INCURRED AS A RESULT OF THE USE OF THIS DASHBOARD OR RELIANCE ON ANY INFORMATION PROVIDED HEREIN. THIS DASHBOARD MAY INCORPORATE DATA OR OTHER CONTENT BELONGING TO OR ORIGINATING FROM THIRD PARTIES. WE DO NOT WARRANT, ENDORSE, GUARANTEE, OR ASSUME RESPONSIBILITY FOR THE ACCURACY OR RELIABILITY OF ANY THIRD-PARTY INFORMATION.

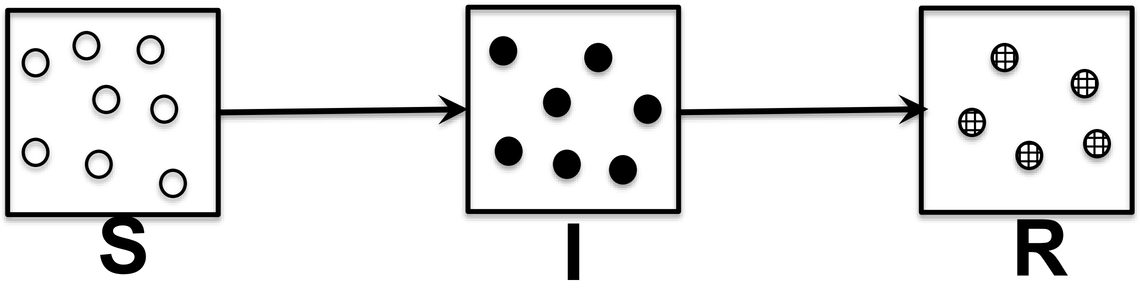
Mathematical Models to study Epidemics

Compartmental mass action models

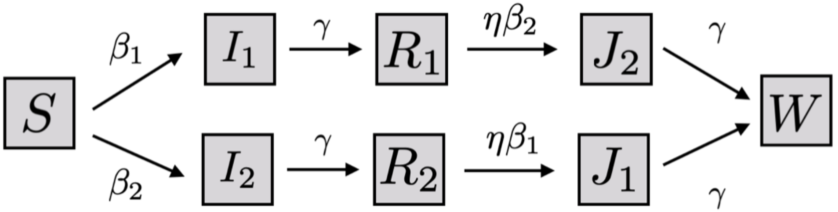
- Susceptible (S): An individual has never had the disease and is susceptible to being infected
- Infected (I): An individual who currently has the disease and can infect other individuals
- Recovered (R): An individual does not have the disease, cannot infect others, and cannot be infected

$$
\begin{aligned}
\frac{ds}{dt} &= -\beta i s \\
\frac{di}{dt} &= \beta i s - \gamma i \\
\frac{dr}{dt} &= \gamma i
\end{aligned}
$$

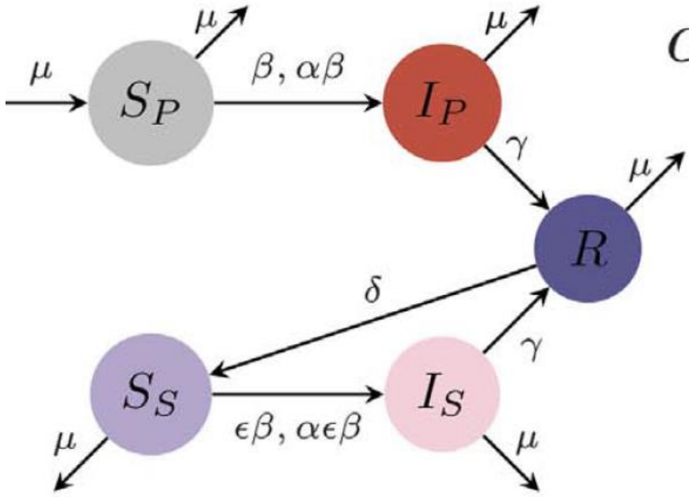




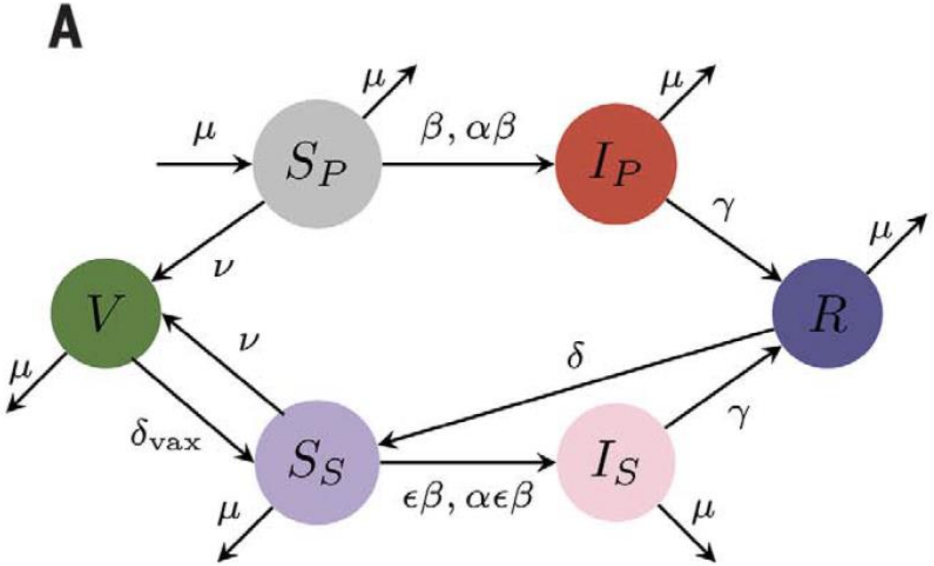
Compartmental mass action models: Extensions



A two strain model



Natural immunity and secondary infections



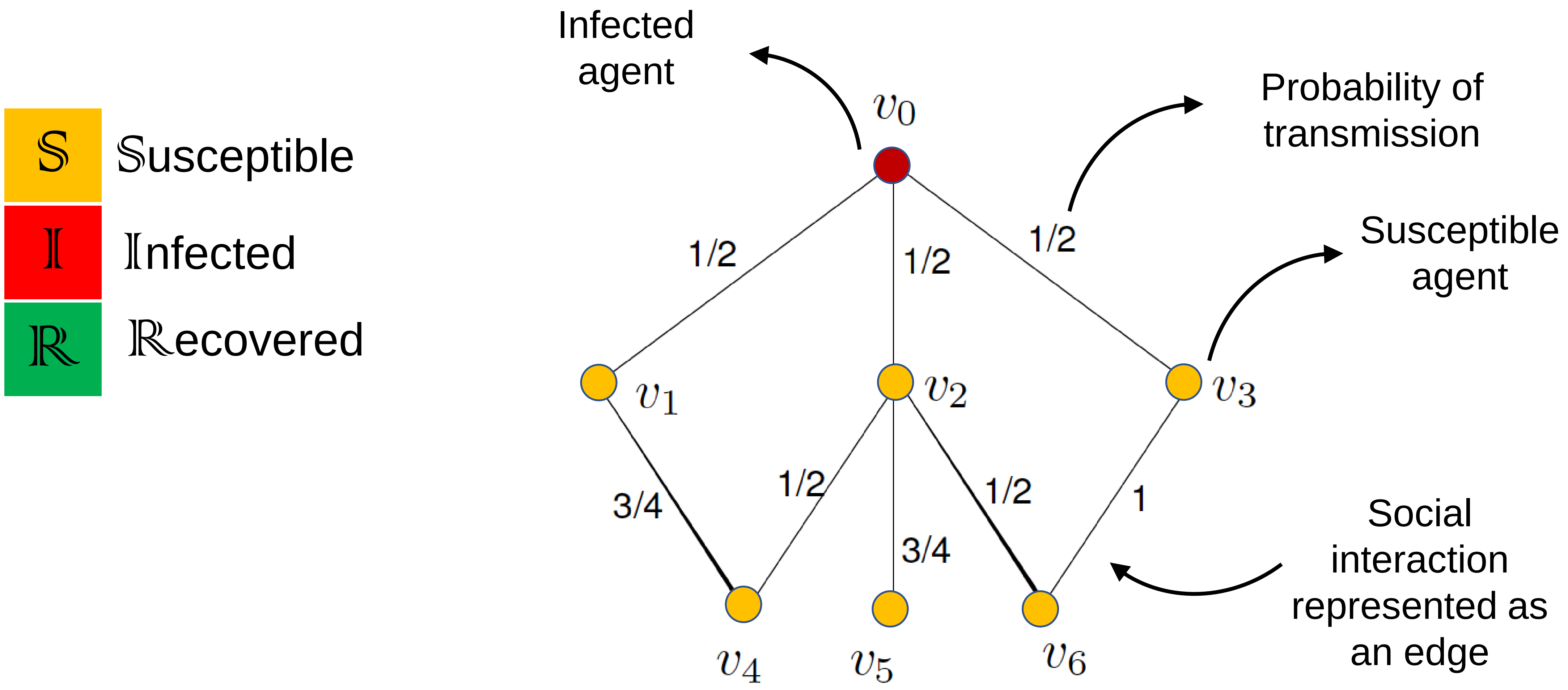
Incorporating vaccines

Directional arrows from one compartment to another represent flows

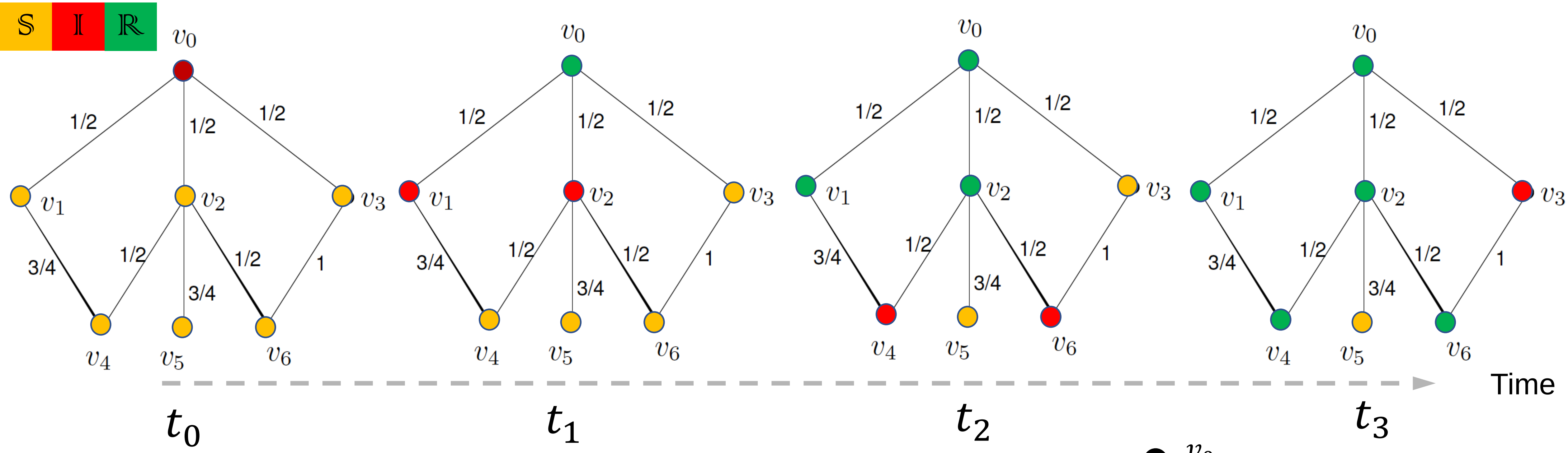
Labels on edges represent coefficients of interactions, e.g. in the third figure, S_P interacts with I_P and I_S with interaction rates β and $\alpha\beta$ respectively.

Baltazar et al. unpublished manuscript, 2021 & Chadi Roy et al. Immune life history, vaccination, and the dynamics of SARS-CoV-2 over the next 5 years, Science , 2020

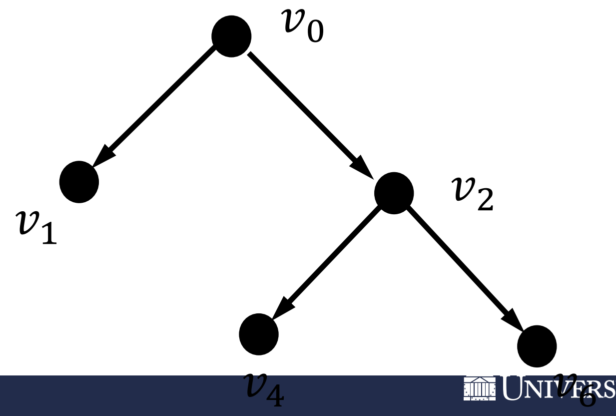
Networked epidemiology: Simple SIR model



Temporal evolution of disease spread

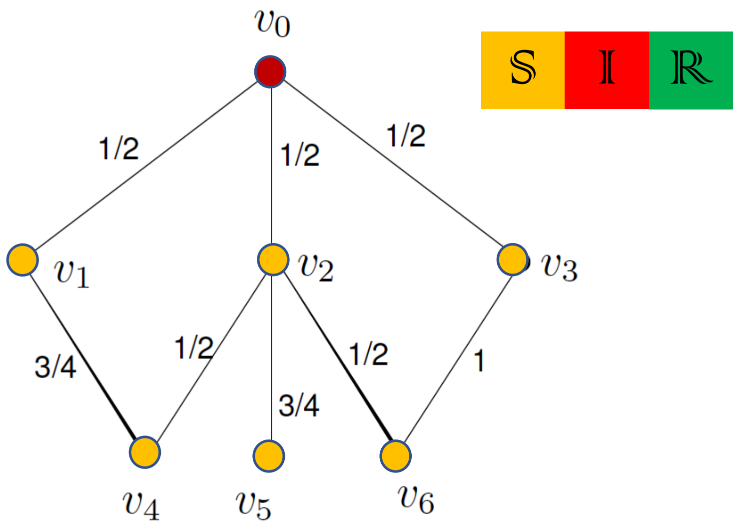


Time	Configuration
0	$(\mathbb{I}, \mathbb{S}, \mathbb{S}, \mathbb{S}, \mathbb{S}, \mathbb{S}, \mathbb{S})$
1	$(\mathbb{R}, \mathbb{I}, \mathbb{I}, \mathbb{S}, \mathbb{S}, \mathbb{S}, \mathbb{S})$
2	$(\mathbb{R}, \mathbb{R}, \mathbb{R}, \mathbb{S}, \mathbb{I}, \mathbb{S}, \mathbb{I})$
3	$(\mathbb{R}, \mathbb{R}, \mathbb{R}, \mathbb{I}, \mathbb{R}, \mathbb{S}, \mathbb{R})$
4	$(\mathbb{R}, \mathbb{R}, \mathbb{R}, \mathbb{R}, \mathbb{R}, \mathbb{S}, \mathbb{R})$



Simple SIR model yields a large Markov chain

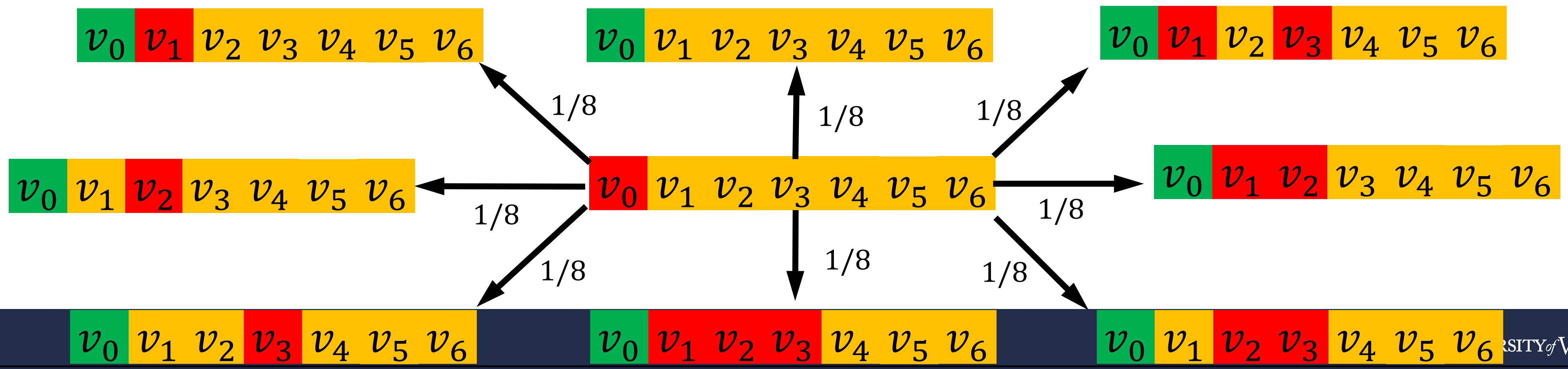
<u>All algorithmic problems come down to analysis of this Markov chain and thus often "hard"</u>



Markov chain (partial) exponentially larger than the network

A social network with $\boldsymbol{n}$ nodes & **3** disease states per nodes yields a chain of size $\mathbf{3}^{\boldsymbol{n}}$

Temporal Dynamics represented by Markov Chain (just a fragment is depicted)



Computational cost of analyzing SIR dynamics

- The reachability problem is #P-hard
  - *Implies that simulations are needed for understanding the dynamics*
  - Even short-term reachability is hard
  - Result due to stochastic nature of the dynamical process
- *Associated Markov chain size -- astronomical in the size of the description.*
  - *Network with* 10^6 *nodes & (3 states: S,I,R), the chain has* 3^{10^6} *nodes; impossible to do naïve search*

Simulations are therefore necessary to solve this problem.

Modeling challenges for operational epidemic science

- ***Challenge 1:*** Computational -– most natural problems are computationally hard.
- ***Challenge 2:*** Data --- Sparse, noisy and lagged and incomplete --> Algorithms need to be robust to this aspect
- ***Challenge 3:*** Near real-time --- Decisions need to be made in near real-time, often with conflicting objectives
- ***Challenge 4:*** Implementable interventions --- Interventions need to be implementable in the real world
- ***Challenge 5:*** Co-evolution --- Decisions and epidemic co-evolves – need robust optimization framework

AI+HPC Driven Real-time, Operational Epidemic Science

https://nssac.github.io/covid-19/index

Science
AAAS

Harvey V. Fineberg is president of the Institute of Medicine.

Mary Elizabeth Wilson is associate professor of Global Health and Population at the Harvard School of Public Health and associate clinical professor at Harvard Medical School, Boston, MA.

Epidemic Science in Real Time

FEW SITUATIONS MORE DRAMATICALLY ILLUSTRATE THE SALIENCE OF SCIENCE TO POLICY THAN AN epidemic. The relevant science takes place rapidly and continually, in the laboratory, clinic, and community. In facing the current swine flu (H1N1 influenza) outbreak, the world has benefited from research investment over many years, as well as from preparedness exercises and planning in many countries. The global public health enterprise has been tempered by the outbreak of severe acute respiratory syndrome (SARS) in 2002–2003, the ongoing threat of highly pathogenic avian flu, and concerns over bioterrorism. Researchers and other experts are now able to make vital contributions in real time. By conducting the right science and communicating expert judgment, scientists can enable policies to be adjusted appropriately as an epidemic scenario unfolds.

In the past, scientists and policy-makers have often failed to take advantage of the opportunity to learn and adjust policy in real time. In 1976, for example, in response to a swine flu outbreak at Fort Dix, New Jersey, a decision was made to mount a nationwide immunization program against this virus because it was deemed similar to that responsible for the 1918–1919 flu pandemic. Immunizations were initiated months later despite the fact that not a single related case of infection had appeared by that time elsewhere in the United States or the world (www.iom.edu/swinefluaffair). Decision-makers failed to take seriously a key question: What additional information could lead to a different course of action? The answer is precisely what should drive a research agenda in real time today.

In the face of a threatened pandemic, policy-makers will want real-time answers in at least five areas where science can help: pandemic risk, vulnerable populations, available interventions, implementation possibilities and pitfalls, and public understanding. Pandemic risk, for example, entails both spread and severity. In the current H1N1 influenza outbreak, the causative virus and its genetic sequence were identified in a matter of days. Within a couple of weeks, an international consortium of investigators developed preliminary assessments of cases and mortality based on epidemic modeling.*

Specific genetic markers on flu viruses have been associated with more severe outbreaks. But virulence is an incompletely understood function of host-pathogen interaction, and the absence of a known marker in the current H1N1 virus does not mean it will remain relatively benign. It may mutate or acquire new genetic material. Thus, ongoing, refined estimates of its pandemic potential will benefit from tracking epidemiological patterns in the field and viral mutations in the laboratory. If epidemic models suggest that more precise estimates on specific elements such as attack rate, case fatality rate, or duration of viral shedding will be pivotal for projecting pandemic potential, then these measurements deserve special attention. Even when more is learned, a degree of uncertainty will persist, and scientists have the responsibility to accurately convey the extent of and change in scientific uncertainty as new information emerges.

A range of laboratory, epidemiologic, and social science research will similarly be required to provide answers about vulnerable populations; interventions to prevent, treat, and mitigate disease and other consequences of a pandemic; and ways of achieving public understanding that avoid both over- and underreaction. Also, we know from past experience that planning for the implementation of such projects has often been inadequate. For example, if the United States decides to immunize twice the number of people in half the usual time, are the existing channels of vaccine distribution and administration up to the task? On a global scale, making the rapid availability and administration of vaccine possible is an order of magnitude more daunting.

Scientists and other flu experts in the United States and around the world have much to occupy their attention. Time and resources are limited, however, and leaders in government agencies will need to ensure that the most consequential scientific questions are answered. In the meantime, scientists can discourage irrational policies, such as the banning of pork imports, and in the face of a threatened pandemic, energetically pursue science in real time.

– Harvey V. Fineberg and Mary Elizabeth Wilson

10.1126/science.1176297

*C. Fraser *et al.*, *Science* 11 May 2009 (10.1126/science.1176062).

An integrated program for real-time COVID-19 response

- Early models for importatio n risk
- Early projections for VDH & VDEM; decisions on field hospitals
- Forecasts for CDC-coordinated COVID-19 Hubs
- University operations scenarios
- Holiday effects and winter surge
- Vaccines and variants

80+ scientists, staff & students

90+ weekly model updates

NSF

DEPARTMENT OF HEALTH & HUMAN SERVICES · USA ·

DEPARTMENT OF DEFENSE · UNITED STATES OF AMERICA

CDC

VIRGINIA · SIC SEMPER TYRANNIS

VDH VIRGINIA DEPARTMENT OF HEALTH

VIRGINIA DEPARTMENT OF EDUCATION

Virginia Department of Emergency Management

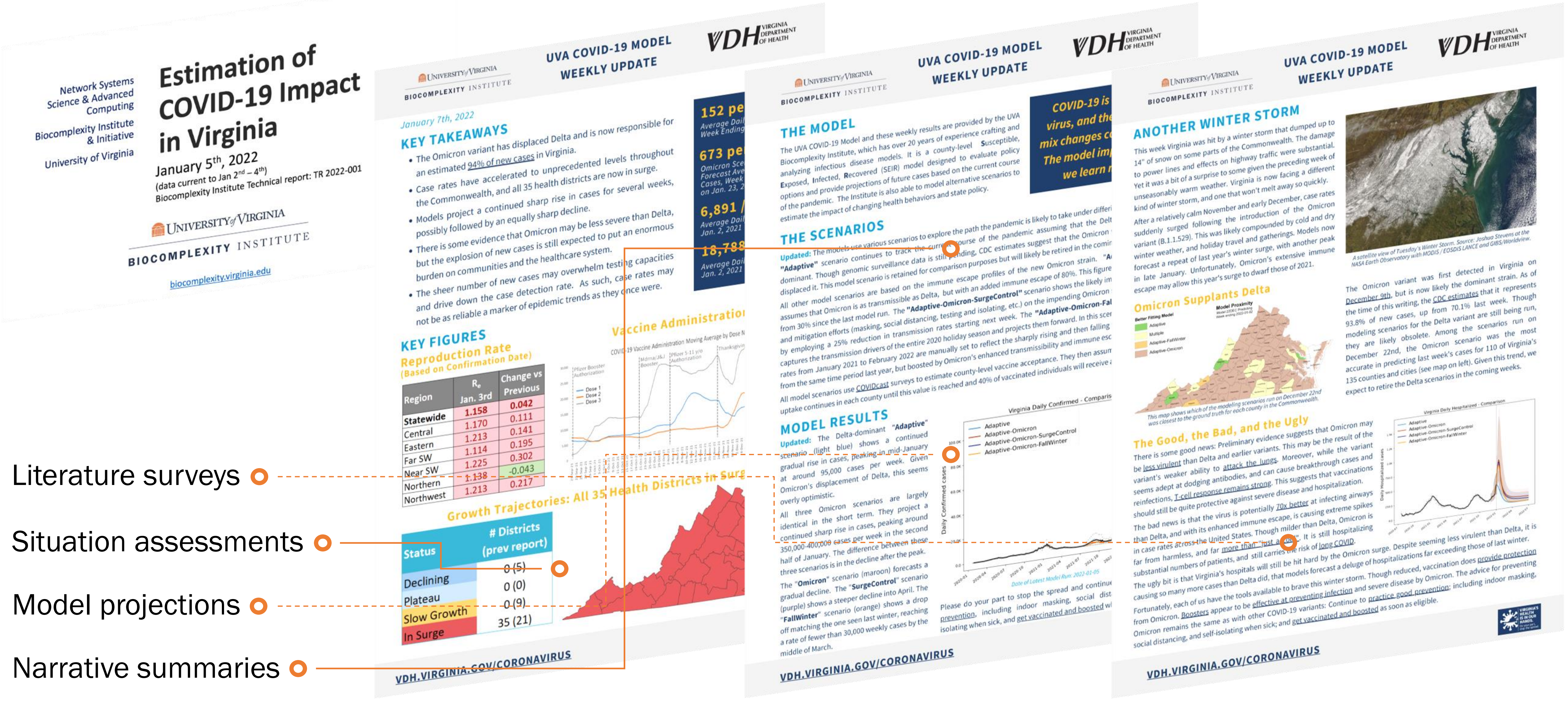
UNIVERSITY of VIRGINIA

JMU JAMES MADISON UNIVERSITY

VIRGINIA COMMONWEALTH UNIVERSITY · MCV · RPI · 1838

VIRGINIA HOSPITAL & HEALTHCARE ASSOCIATION

Weekly updates to state and federal agencies since February 2020



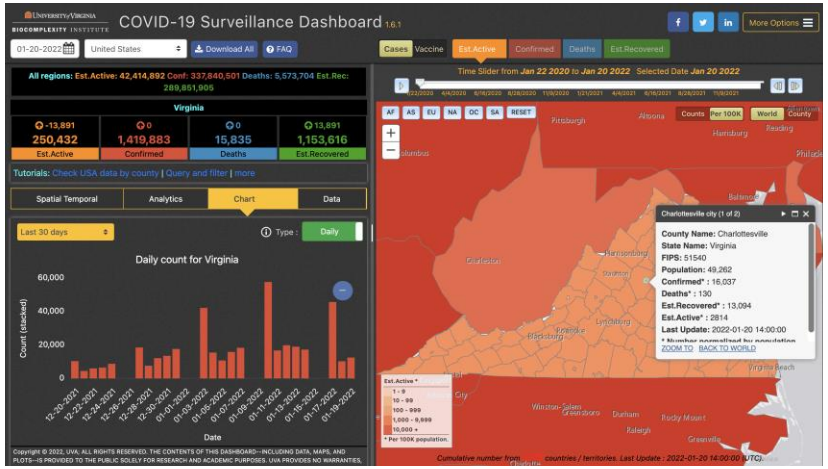
Literature surveys

Situation assessments

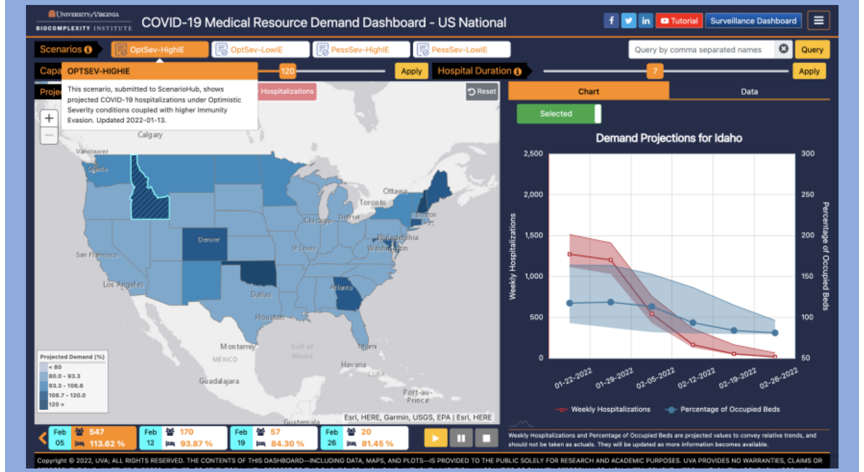
Model projections

Narrative summaries

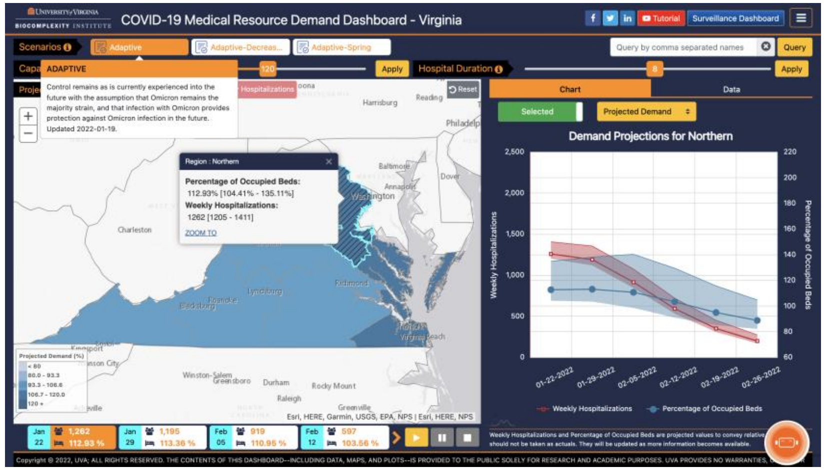
COVID-19 Surveillance



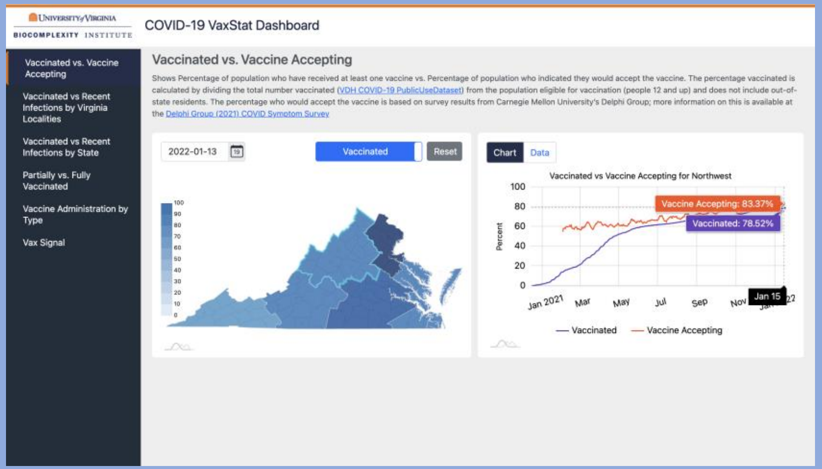
US Medical Resource Demand



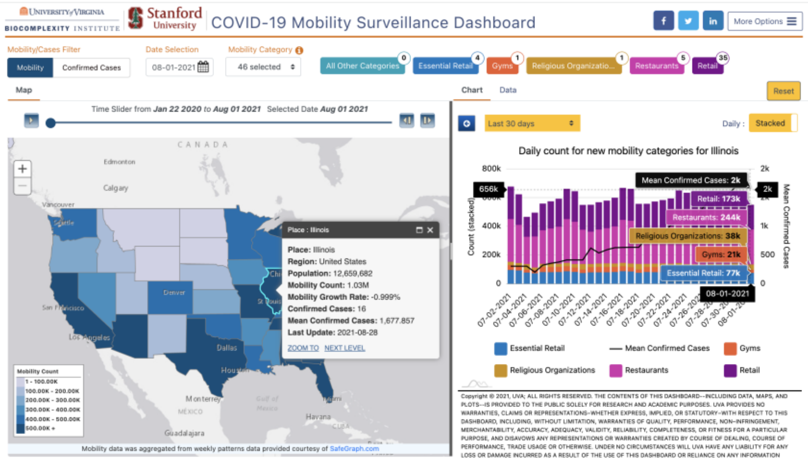
VA Medical Resource Demand



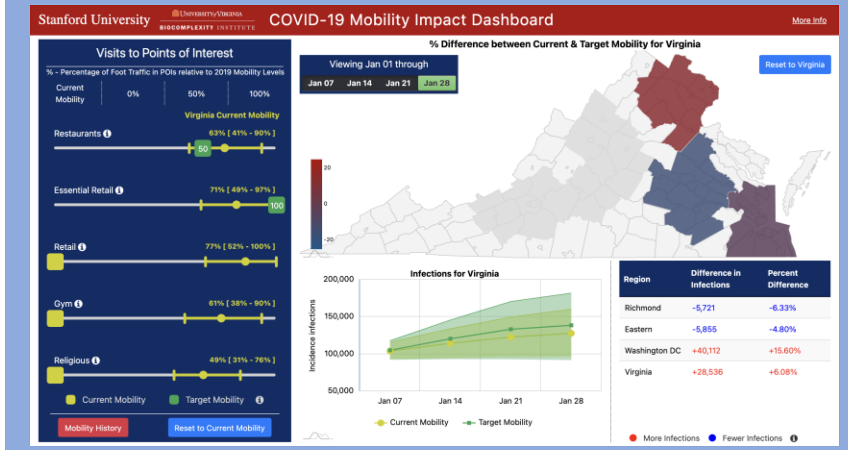
Vaccine Status



COVID-19 Mobility Surveillance



COVID-19 Mobility Impact



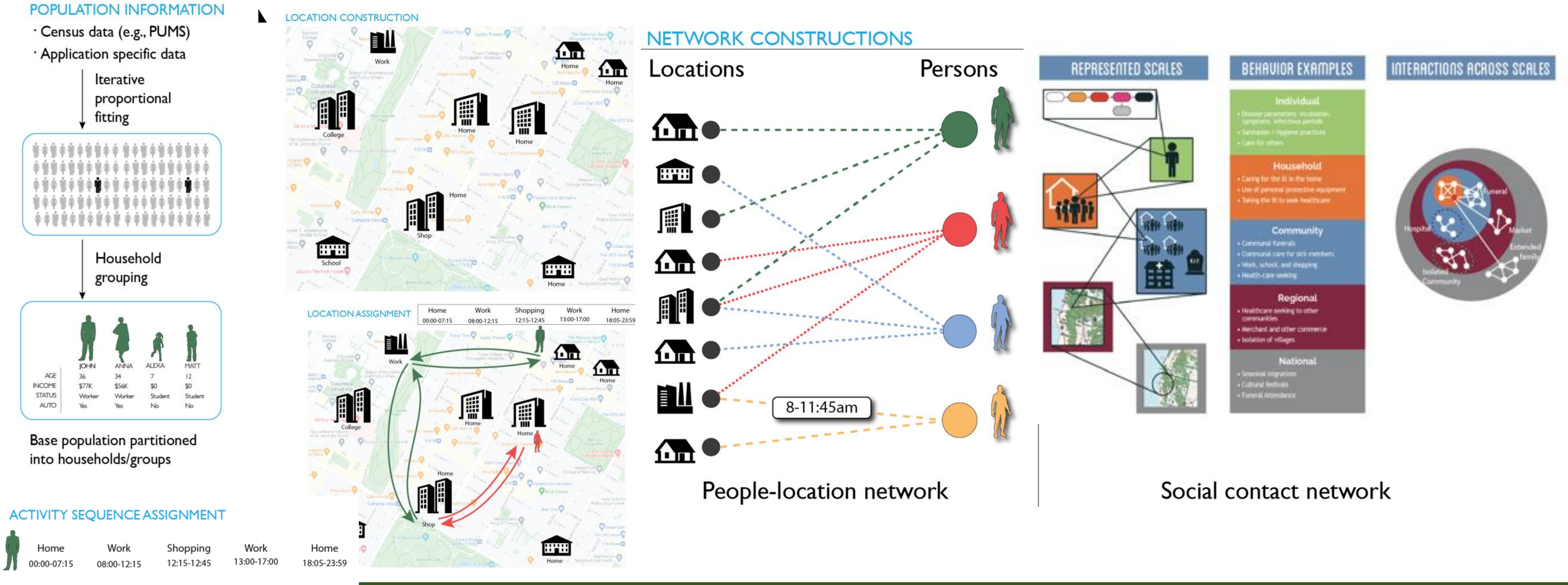
Decision support dashboards

https://nssac.github.io/covid-19/index

UNIVERSITY of VIRGINIA
BIOCOMPLEXITY INSTITUTE

Overall approach

Component 1: Digital twin of US social contact network



Digital twin captures heterogeneous spatially explicit, multi-scale, multi-theory, multi-level social contact network representations

Component 2: EpiHiper -- a parallel agent-based socio-epidemic simulator

Inputs: disease model + Social contact network from Digital Twin

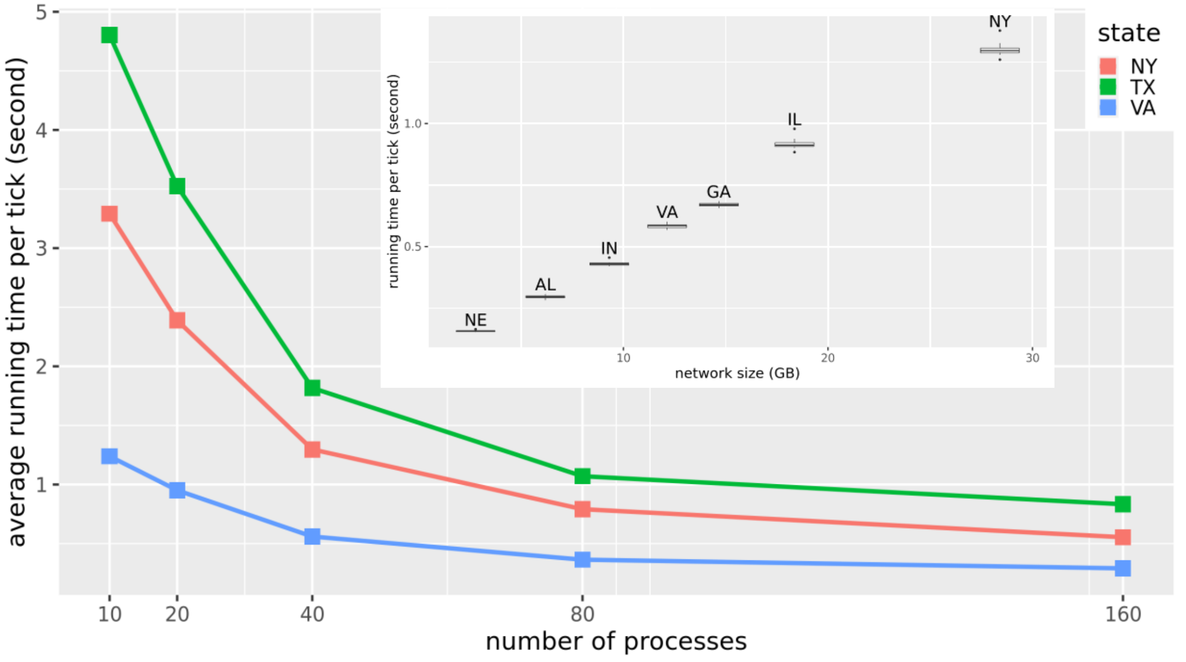
Agent-based discret time/event simulatio model

Contact network partitioned onto MPI ranks

Distributed memory program written in C++/MPI

Custom domain-specific language for programming NPI scenarios

Shared inputs served v PostgreSQL database



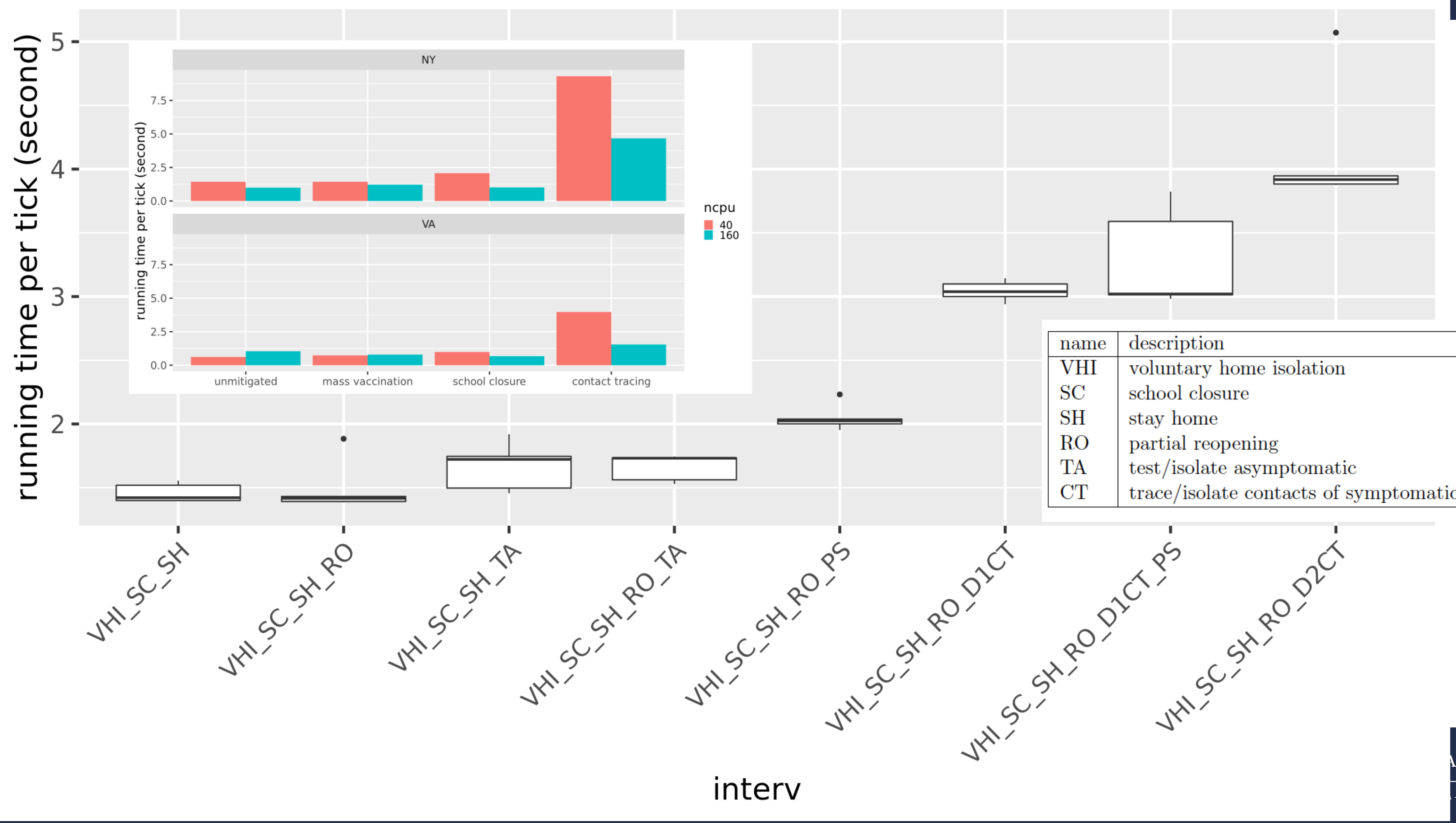
Interventions

- Pharmaceutical (PI) vs non-pharmaceutical (NPI)
  - Vaccination, Anitvirals
  - Isolation, Sequestration, Quarantine, social-distance
- Public health vs individual
  - Work/school closure, Wearing a mask, self isolation
- Pre-specified vs online
- Static (independent of simulation outcome) vs dynamic (dependent on simulation progress)
- Targeted vs un-targeted

if **conditions (triggers)** are satisfied.

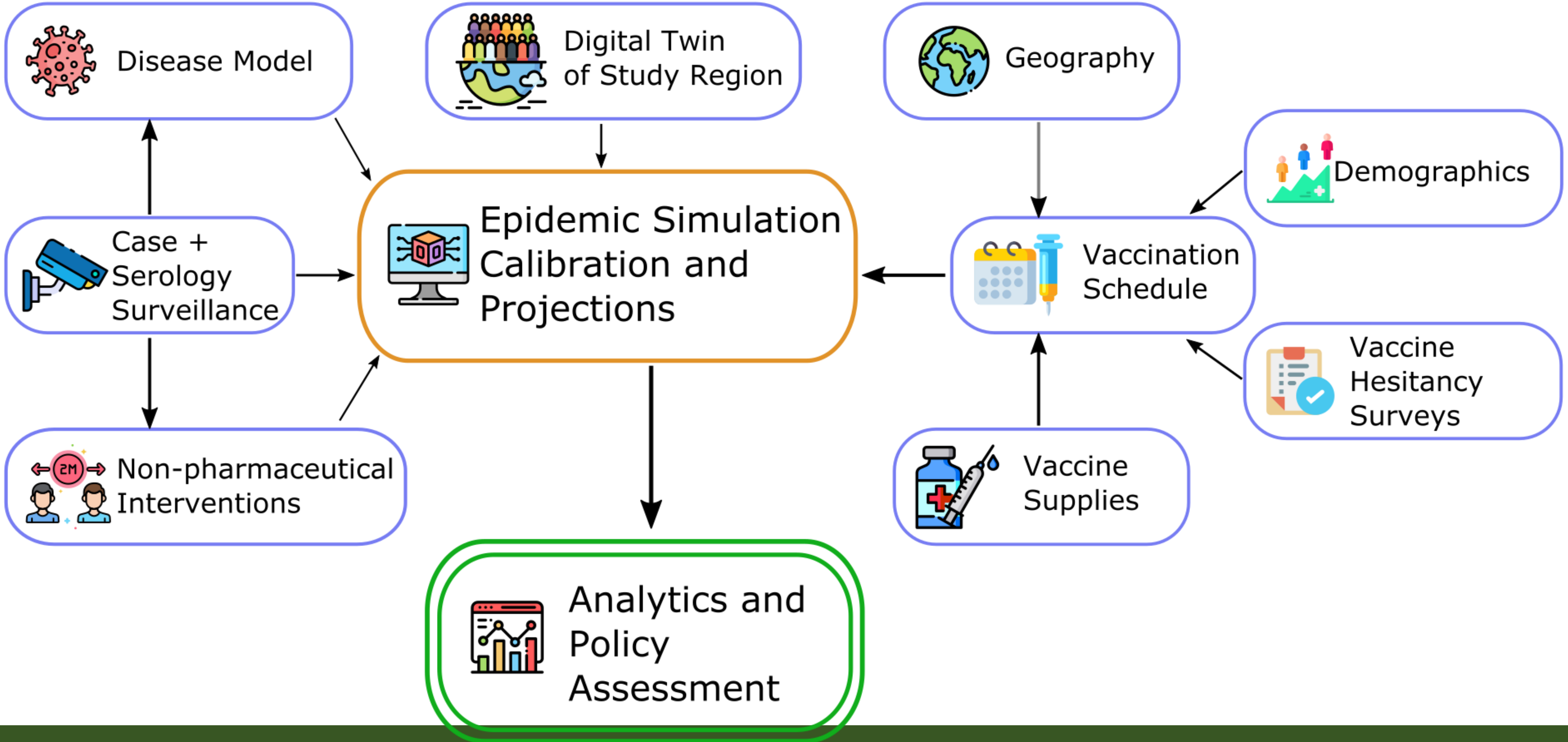
apply **action** (with *parameters*) to **subnetwork** (chosen with *constraints*)

Interventions are expensive



name	description
VHI	voluntary home isolation
SC	school closure
SH	stay home
RO	partial reopening
TA	test/isolate asymptomatic
CT	trace/isolate contacts of symptomatic cases

Component 3: Live digital twin (DT)



Bringing the digital twin to life: Contextualize the DT with current ground conditions

Analytical workflows for COVID-19

Risk of importation of disease to various parts of the world

Forecasting the future course of the pandemic with and without interventions

Resource augmentation, e.g. building field hospitals and resource sharing, e.g. ventilators

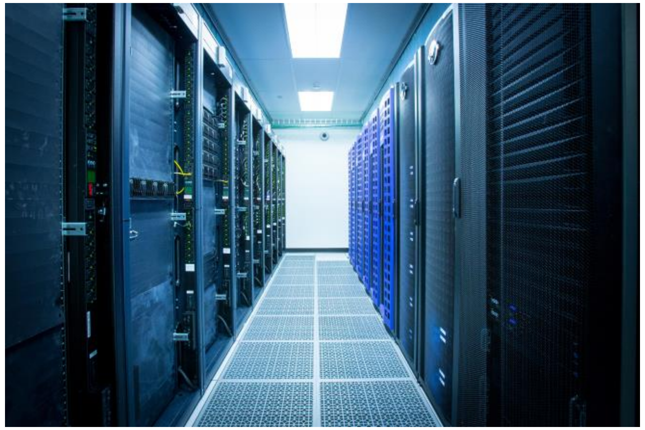
Counterfactual analysis to study how the pandemic will evolve under various scenarios

Projecting the healthcare demands including PPE, ventilators, beds

Economic impact analysis: cost/benefit analysis of interventions

Integrated workflow orchestration across two supercomputers

20,000 cores



2,000 cores

Incredible support by UVA Computing, PSC and XSEDE: Rick, Shawn Brown, John Towns and their team made this happen and provided unprecedented access and real-time support

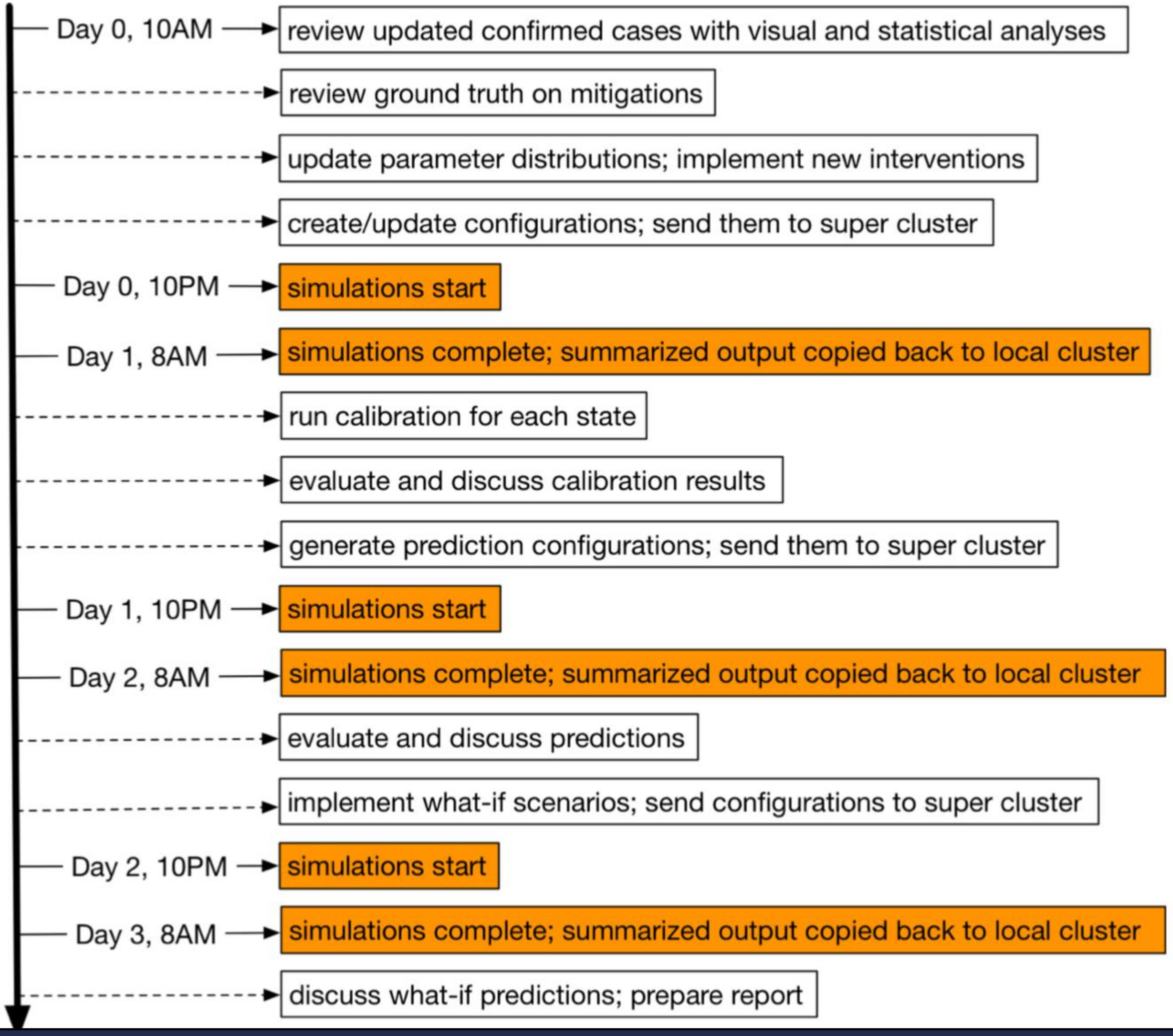
	Remote Cluster Bridges@PSC	Home Cluster Rivanna@UVA
#Allocated nodes	720	50
#CPUs/node	2	2
#Cores/CPU	14	20
RAM/node	128GB (DDR4)	384GB (DDR4)
CPU	Intel Haswell E5-2695 v3	Intel Xeon Gold 6148
Network	Intel Omni path-1	Mellanox ConnectX-5
Filesystem	Lustre	Lustre

Scalable epidemiological workflows to support covid-19 planning and response, to appear in IPDPS, 2021, D Machi,

HPC-Enabled workflows for real-Time epidemiology

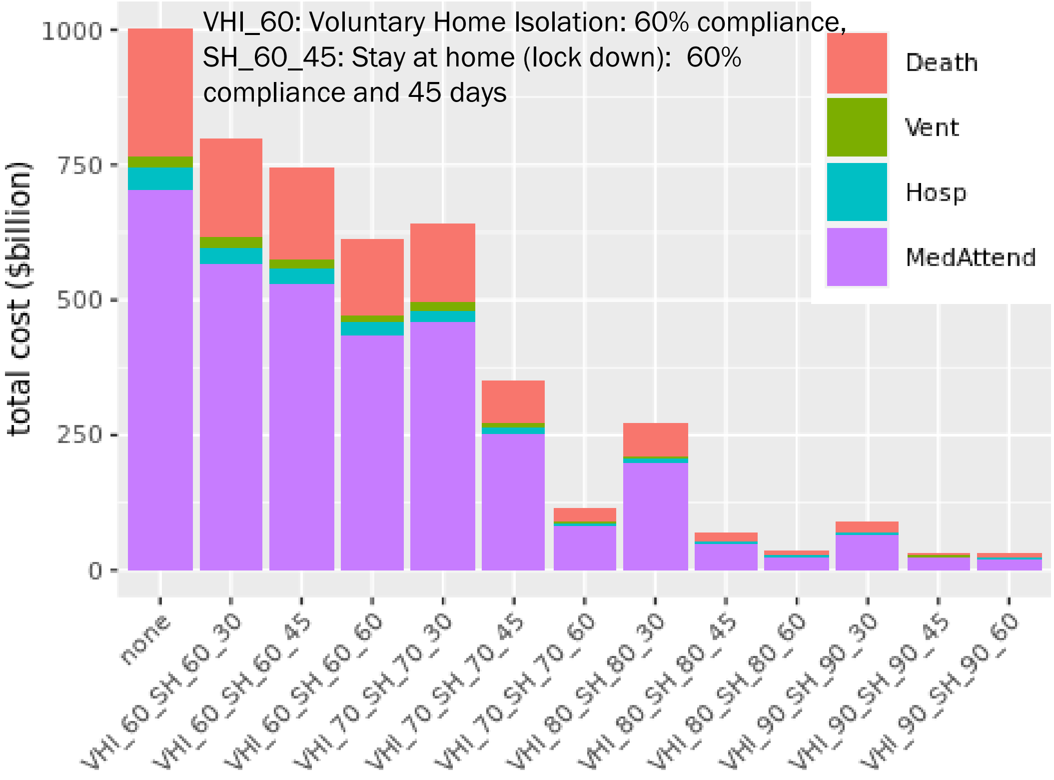
During March- November 2020

- Typically, per night, the pipeline runs 5,000-17,900 regional or national simulations
- Network with 300 million nodes and 7.9 billion edges partitioned across all 50 states
- The simulations yield ensemble models for prediction of epidemic incidence curves at the US county level (3243 counties).



Computing medical costs

- Objective:
  - Analyze economic impact of infections and NPI's (Medical costs)
- Study settings:
  - calibrated COVID-19 disease model
  - NPI's with parameter sweep on compliance and duration
  - medical costs due to hospitalizations and deaths
  - GDP reduction due to cascading effects between sectors from workforce interruptions not computed
- Findings:
  - Total medical cost decreases with NPI compliance and duration.
  - NPI's have significant impact on Medical costs



www.nature.com/scientificreports

scientific reports

OPEN **Epidemiological and economic impact of COVID-19 in the US**

Jiangzhuo Chen[1], Anil Vullikanti[1,2], Joost Santos[3], Srinivasan Venkatramanan[1], Stefan Hoops[1], Henning Mortveit[1,4], Bryan Lewis[1], Wen You[5], Stephen Eubank[1,5], Madhav Marathe[1,2], Chris Barrett[1,2] & Achla Marathe[1,5]

This research measures the epidemiological and economic impact of COVID-19 spread in the US under different mitigation scenarios, comprising of non-pharmaceutical interventions. A detailed disease model of COVID-19 is combined with a model of the US economy to estimate the direct impact of labor supply shock to each sector arising from morbidity, mortality, and lockdown, as well as the indirect impact caused by the interdependencies between sectors. During a lockdown, estimates of jobs that are workable from home in each sector are used to modify the shock to labor supply. Results show trade-offs between economic losses, and lives saved and infections averted are non-linear in compliance to social distancing and the duration of the lockdown. Sectors that are worst hit are not the labor-intensive sectors such as the Agriculture sector and the Construction sector, but the ones with high valued jobs such as the Professional Services, even after the teleworkability of jobs is accounted for. Additionally, the findings show that a low compliance to interventions can be overcome by a longer shutdown period and vice versa to arrive at similar epidemiological impact but their net effect on economic loss depends on the interplay between the marginal gains from averting infections and deaths, versus the marginal loss from having healthy workers stay at home during the shutdown.

www.nature.com/scientificreports

scientific reports

OPEN **Medical costs of keeping the US economy open during COVID-19**

Jiangzhuo Chen[1], Anil Vullikanti[1,2], Stefan Hoops[1], Henning Mortveit[1,3], Bryan Lewis[1], Srinivasan Venkatramanan[1], Wen You[4], Stephen Eubank[1,4], Madhav Marathe[1,2], Chris Barrett[1,2] & Achla Marathe[1,4]

We use an individual based model and national level epidemic simulations to estimate the medical costs of keeping the US economy open during COVID-19 pandemic under different counterfactual scenarios. We model an unmitigated scenario and 12 mitigation scenarios which differ in compliance behavior to social distancing strategies and in the duration of the stay-home order. Under each scenario we estimate the number of people who are likely to get infected and require medical attention, hospitalization, and ventilators. Given the per capita medical cost for each of these health states, we compute the total medical costs for each scenario and show the tradeoffs between deaths, costs, infections, compliance and the duration of stay-home order. We also consider the hospital bed capacity of each Hospital Referral Region (HRR) in the US to estimate the deficit in beds each HRR will likely encounter given the demand for hospital beds. We consider a case where HRRs share hospital beds among the neighboring HRRs during a surge in demand beyond the available beds and the impact it has in controlling additional deaths.

Chen, J., Vullikanti, A., Hoops, S., Mortveit, H., Lewis, B., Venkatramanan, S., You, W., Eubank, S., Marathe, M., Barrett, C. and Marathe, A., 2020. Medical costs of keeping the US economy open during COVID-19. *Scientific reports*, *10*(1), pp.1-10.

Supporting the scenario modeling hub

Problem: Real time scenario projections in an evolving global pandemic

- **December 2020**: Optimally allocating a limited supply of vaccines
- **February 2021**: Optimally and fairly allocating vaccines among different states
- **April 2021**: Understanding risks of new and emerging variants
- **June 2021**: Understanding impacts and risks of vaccine hesitancy and reduced NPIs
- **July 2021**: Planning for multiple evolving strains
- **August 2021**: Understanding impacts of waning immunity from vaccination and past infections
- **September 2021**: Understanding impacts of vaccinating 5-11 year-olds
- **October 2021**

<u>The scenario projection problem</u>: Given the current conditions on the ground, and a set of possible future scenarios, the goal is to assess the likelihood of epidemiological outcomes for each of these scenarios by analyzing the output of our simulations.

Assessing the impact of waning immunity (Example)

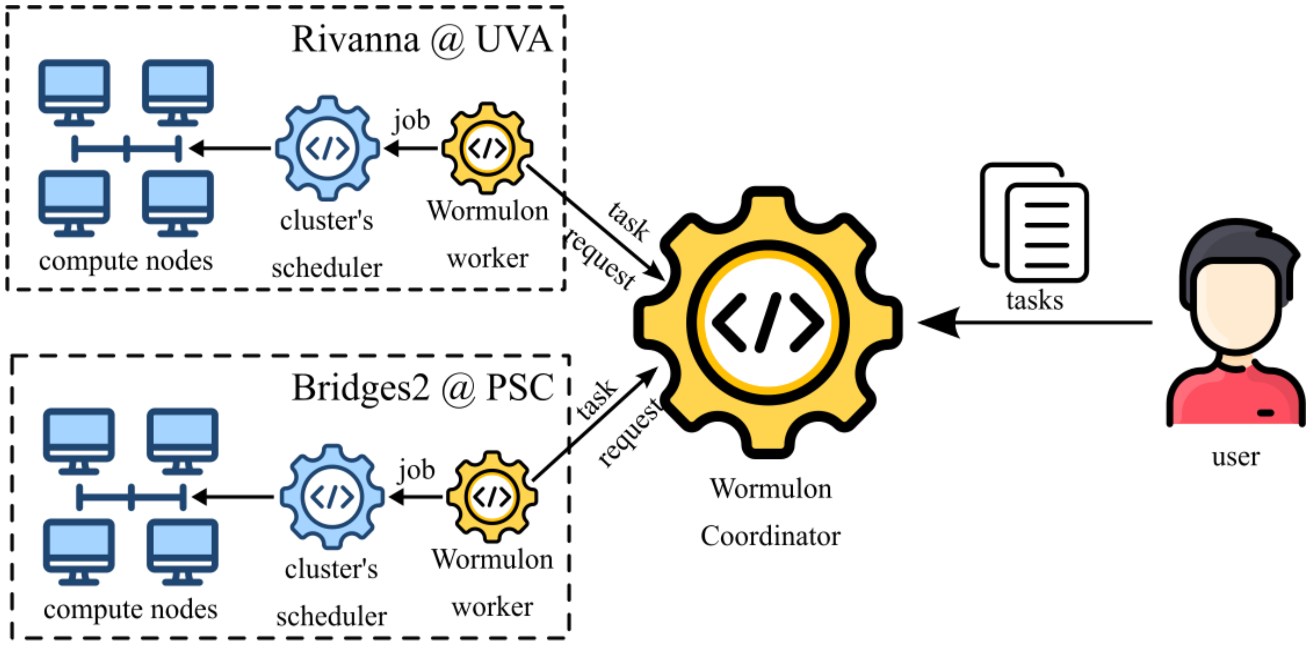
COVID-19 ScenarioModelingHub

Scenario projections from 9 teams are harmonized to assist decision makers with COVID-related planning for the next 3-6 months

Each model projects six different targets (variables) for the US at state- or county-level resolution

Round 12	**Higher immune escape:** 80% of previously immune are susceptible to infection	**Lower immune escape:** 50% of previously immune are susceptible to infection
Lower severity: 70% reduction in severity of Omicron infection, relative to Delta (all-age risk of hospitalization and death times 0.3) among all immune classes.	Scenario A Optimistic severity, High immune escape	Scenario B Optimistic severity, Low immune escape
Higher severity: 30% reduction in severity of Omicron infection, relative to Delta (all-age risk of hospitalization and death times 0.7) among all immune classes.	Scenario C Pessimistic severity, High immune escape	Scenario D Pessimistic severity, Low immune escape

Hybrid HPC pipeline: orchestrating two supercomputers



- Leverages HPC schedulers at HPC clusters & runs on modern secure clusters
- Supports task dependency and dynamic task creation
- Supports task semantic-aware fault detection and retries
- Supports disparate HPC site-specific configurations
- Coordinator-Worker architecture: Single coordinator & a cluster worker per HPC cluster

Data intensive, real-time, end-to-end HPC-workflows

Inputs

3100 US counties
500 days
confirmed case count data

~20GB synthetic population
person, household, location, and activities data

300 million nodes and 12 Billion edges

Simulations

~10200 simulation instances
factorial design

2550 simulation instances
calibration

33 hours
9000 cores
spread over two machines

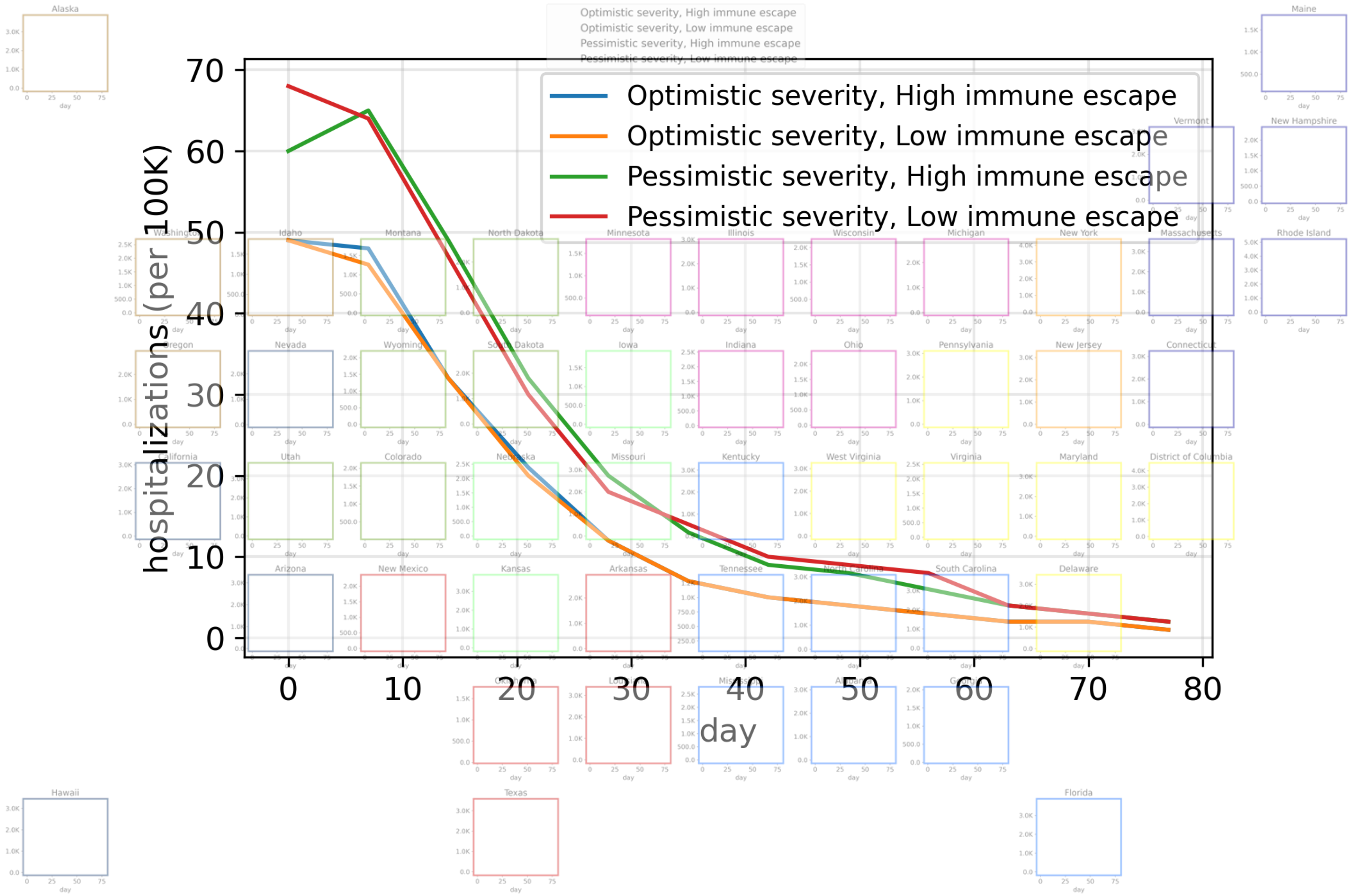
Outputs

~1.5B entries, ~4GB aggregate output data
250 days, 200 health states, 3 counts

~4TB individual-level output data
multi-million state transitions = multi-billion entries

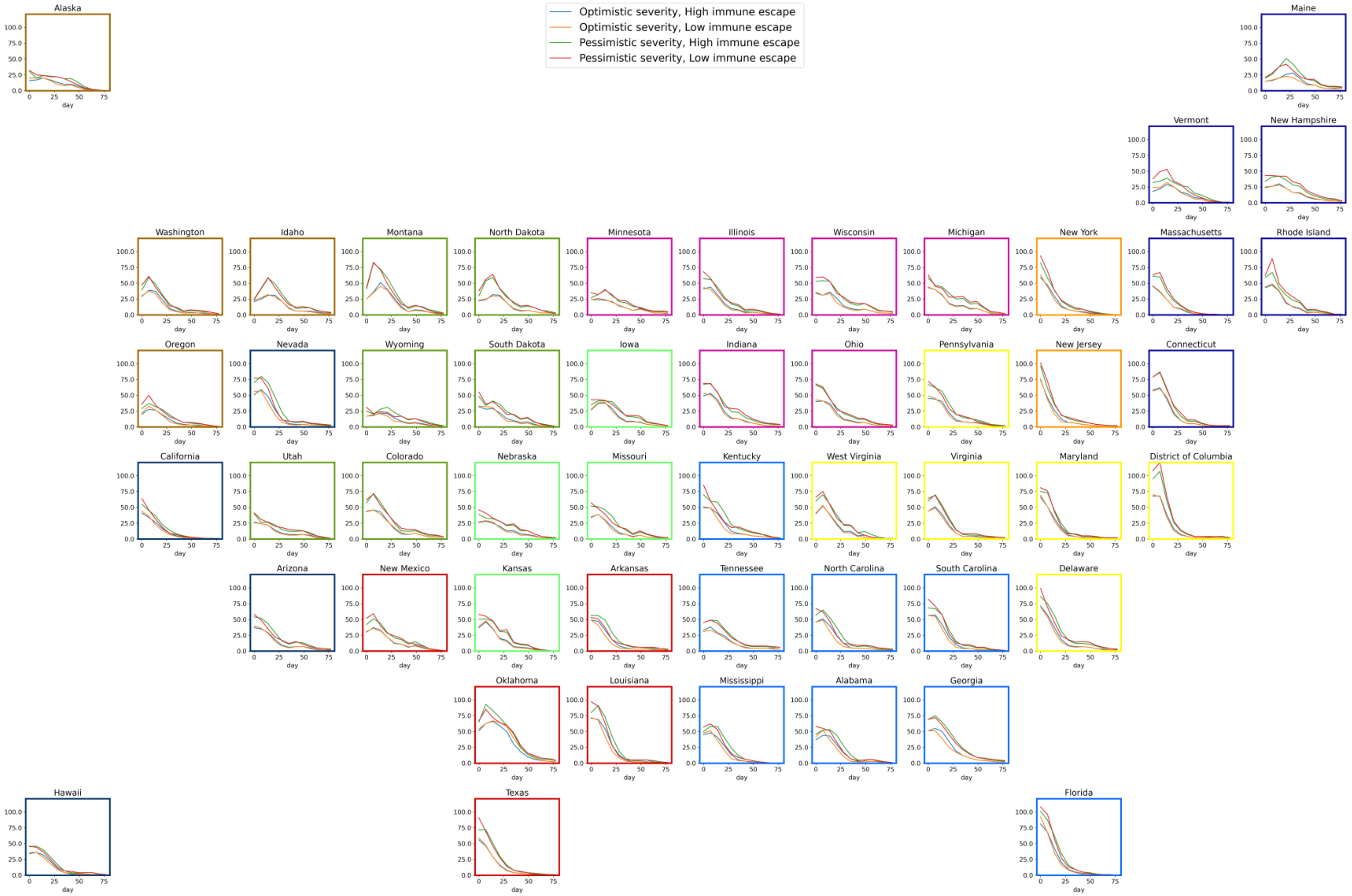
Spatio-temporal variability of the projection: hospitalizations in US

- Pessimistic severity scenarios result in larger number of hospitalizations
- Effect of immune escape is relatively small
- Hospitalizations rapidly decrease next month in all scenarios



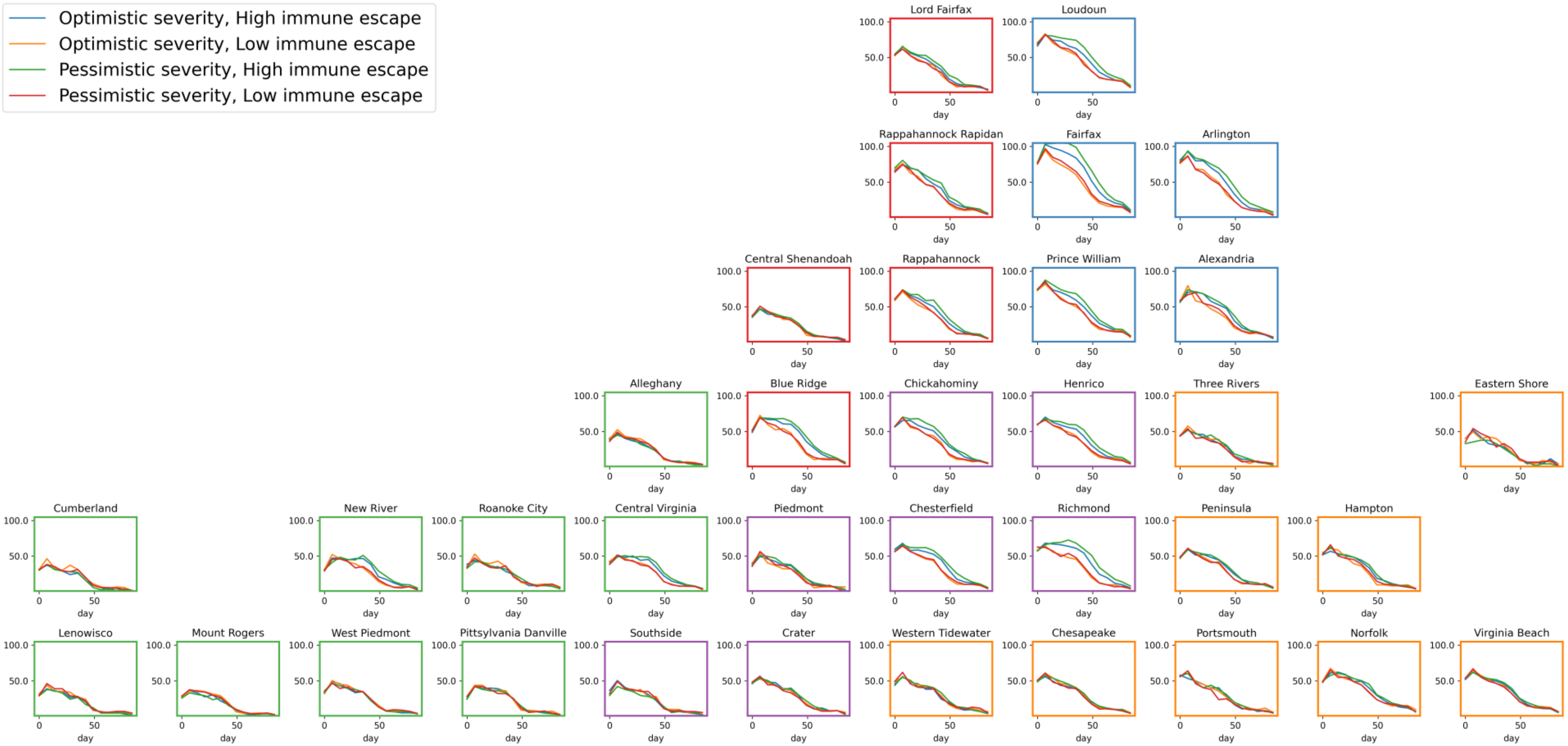
Spatio-temporal variability of the projection: hospitalizations in US

- Pessimistic severity scenarios result in larger numbers of hospitalizations
- Hospitalizations rapidly decrease in next month in all scenarios
- Significant differences between states



Spatio-temporal variability of the projection: hospitalizations in Virginia health districts

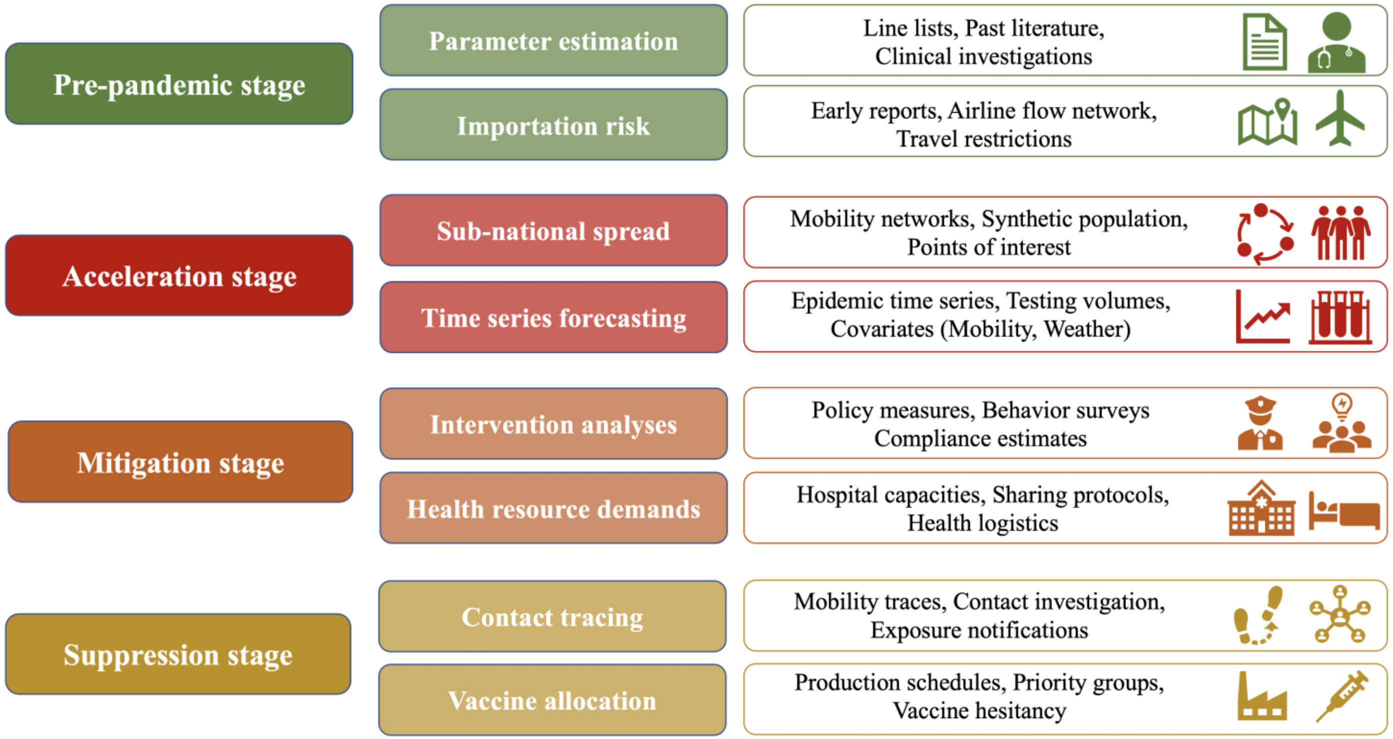
- Hospitalizations peak in January 2022 in all districts, then rapidly decrease next month in all scenarios
- Significant differences between districts
- Higher number of hospitalizations (normalized by population) in Northern VA



Challenges for Data Scientists and AI

Data Challenge

- Size of Data Sets
  - Observational and simulated data for global problems
- Noisy and sparse, Time lagged, Non Stationary and often not in computable form nor linked
  - Reasoning across these data sets becomes challenging
- Privacy, anonymity, security and commercial concerns
  - Individualized privacy preserving data is hard to get

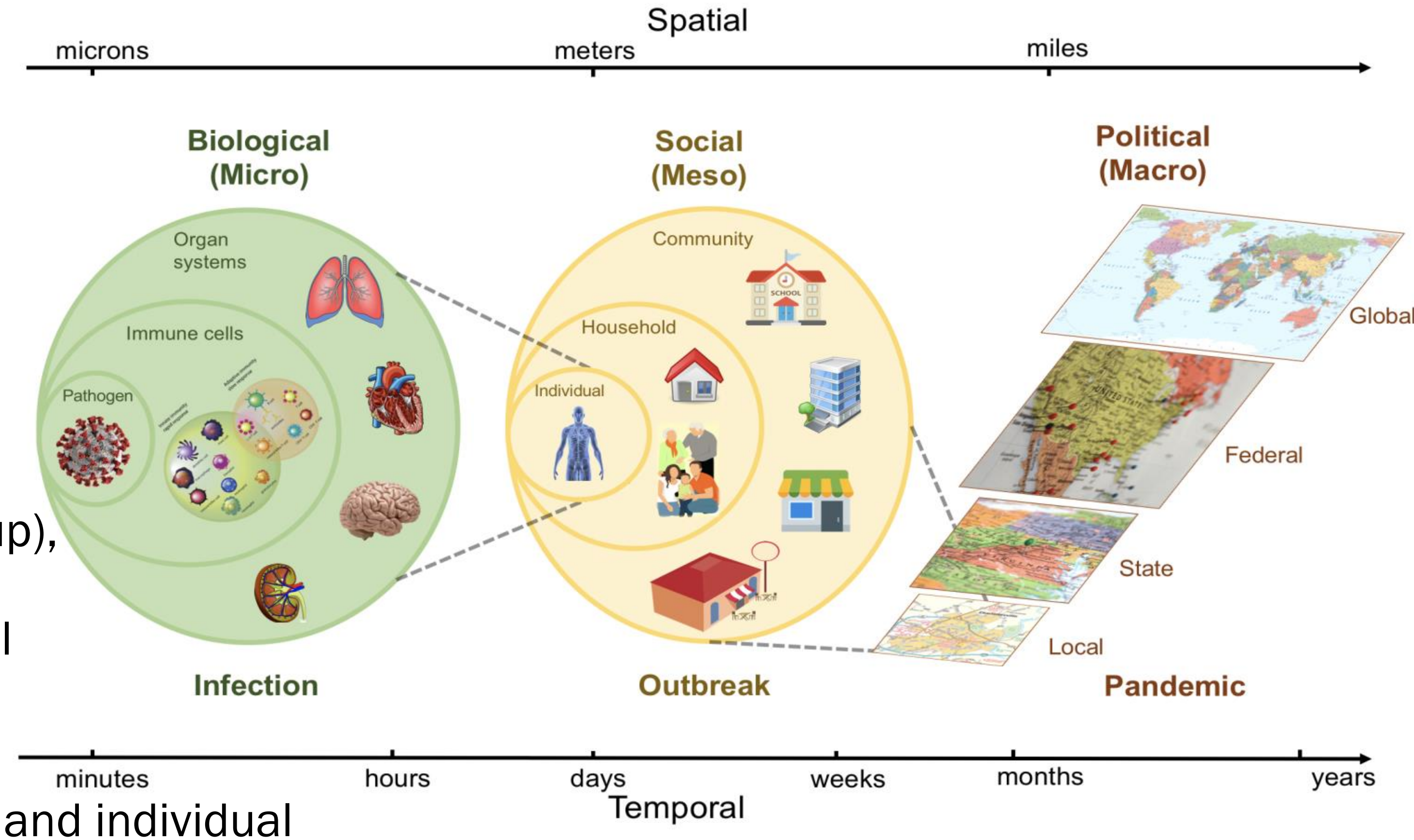


Need: Federated and generalizable machine Learning models trained with limited and noisy data.

Leveraging Data Science to Combat COVID-19: A Comprehensive Review, Latif et al. IEEE Tran.in AI
Data-driven modeling for different stages of pandemic response, Adiga et al. J. IISc, 2020

Scale Challenge: Spatial, temporal and social scales

- Processes unfolding across temporal scales
  - Within host disease progression
  - Between host transmission
  - Individual behavioral change
  - Community outbreak response
  - Public health control measures
  - Global response coordination
  - Seasonality and waning immunity
- Spatial and Social Interaction Scale
  - House (building), neighborhood (block group), city; state, country and region
  - Household, neighborhood and country level contact network
- Individual and collective behavior
  - Organizations (months), Community (days) and individual (days)



Need: Multi-scale, multi-theory, multi-level network representations, analytical tools and simulations

Leveraging Data Science to Combat COVID-19: A Comprehensive Review, Latif et al. IEEE Tran.in AI
Data-driven modeling for different stages of pandemic response, Adiga et al. J. IISc, 2020

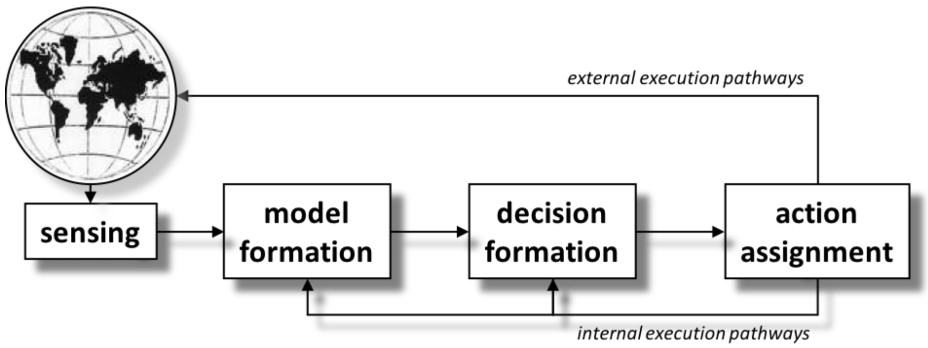
Trust and Adequacy challenge

- Retrospective and predictive validity are not as useful in crisis situations when data is limited.
- External Validation
  - validate past predictions
  - update future projections
- Internal Validation
  - ensure structurally correct
- How do we gather and incorporate relevant data in real-time to:
  - when modeling assumptions cause models to fail to capture real-world dynamics?

Additional challenges in: governance, ethics, logistics, diplomacy and one health are important but could not be covered due to time constraints

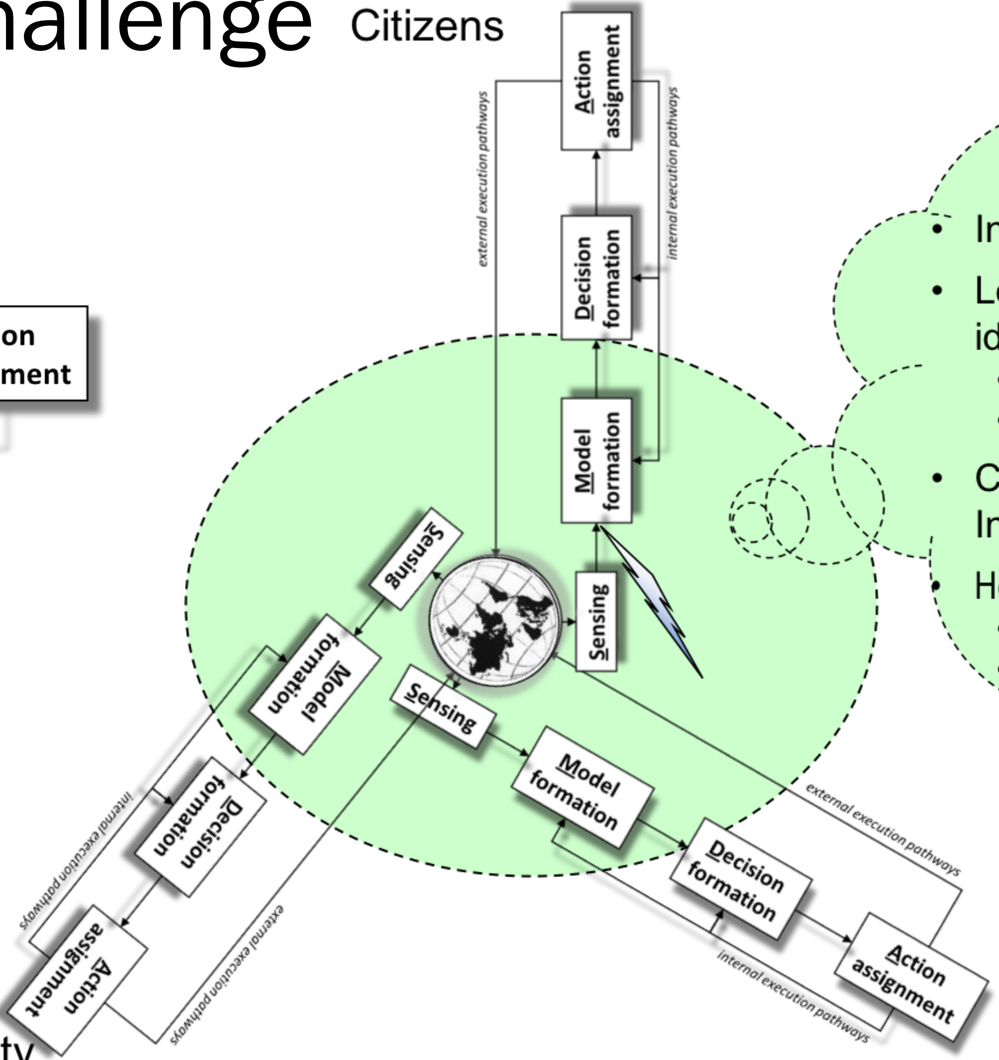
Need: New approaches for V&V, UQ and model adaptation for co-evolving networks

Decision making challenge



Individual decision maker

Citizens



US, State & Local Authority

On the ground responders: Doctors, Nurses, Educators, …

- Information and actions are distributed
- Local models are compatible, but not identical or simple transformations
  - Differences may be important
  - Dynamics
- Context Specific Shared Synthetic Information
- Horizontal Integration
  - Non-attribution of sources
  - Privacy/ security

Need: Robust human-centered, decision making that is operational, works with diverse decision makers and is real-time

Pervasive, Personalized and Precision (P^3) analytics for multiplexed networked systems

Pervasive: *Enable decision maker to make decisions at any, place, anytime and any device*

Personalized: *Enable decision maker to get personalized information that reflects her context*

Precise: *The decision maker should have precise information in space, time and context.*

PERSPECTIVE

doi:10.1038/nature12047

Globally networked risks and how to respond

Dirk Helbing[1,2]

Today's strongly connected, global networks have produced highly interdependent systems that we do not understand and cannot control well. These systems are vulnerable to failure at all scales, posing serious threats to society, even when external shocks are absent. As the complexity and interaction strengths in our networked world increase, man-made systems can become unstable, creating uncontrollable situations even when decision-makers are well-skilled, have all data and technology at their disposal, and do their best. To make these systems manageable, a fundamental redesign is needed. A 'Global Systems Science' might create the required knowledge and paradigm shift in thinking.

Globalization and technological revolutions are changing our planet. Today we have a worldwide exchange of people, goods, money, information, and ideas, which has produced many new opportunities, services and benefits for humanity. At the same time, however, the underlying networks have created pathways along which dangerous and damaging events can spread rapidly and globally. This has increased systemic risks[1] (see Box 1). The related societal costs are huge.

When analysing today's environmental, health and financial systems or our supply chains and information and communication systems, one finds that these systems have become vulnerable on a planetary scale. They are challenged by the disruptive influences of global warming, disease outbreaks, food (distribution) shortages, financial crashes, heavy solar storms, organized (cyber-)crime, or cyberwar. Our world is already facing some of the consequences: global problems such as fiscal and economic crises, global migration, and an explosive mix of incompatible interests and cultures, coming along with social unrests, international and civil wars, and global terrorism.

In this Perspective, I argue that systemic failures and extreme events are consequences of the highly interconnected systems and networked risks humans have created. When networks are interdependent[2,3], this makes them even more vulnerable to abrupt failures[4–6]. Such interdependencies in our "hyper-connected world"[1] establish "hyper-risks" (see Fig. 1). For example, today's quick spreading of emergent epidemics is largely a result of global air traffic, and may have serious impacts on our global health, social and economic systems[6–9]. I also argue that initially beneficial trends such as globalization, increasing network densities, sparse use of resources, higher complexity, and an acceleration of institutional decision processes may ultimately push our anthropogenic (man-made or human-influenced) systems[10] towards systemic instability—a state in which things will inevitably get out of control sooner or later.

Many disasters in anthropogenic systems should not be seen as 'bad luck', but as the results of inappropriate interactions and institutional settings. Even worse, they are often the consequences of a wrong understanding due to the counter-intuitive nature of the underlying system behaviour. Hence, conventional thinking can cause fateful decisions and the repetition of previous mistakes. This calls for a paradigm shift in thinking: systemic instabilities can be understood by a change in perspective from a component-oriented to an interaction- and network-oriented view. This also implies a fundamental change in the design and management of complex dynamical systems.

The FuturICT community[11] (see http://www.futurict.eu), which involves thousands of scientists worldwide, is now engaged in establishing a 'Global Systems Science', in order to understand better our information society with its close co-evolution of information and communication technology (ICT) and society. This effort is allied with the "Earth system science"[10] that now provides the prevailing approach to studying the physics, chemistry and biology of our planet. Global Systems Science wants to make the theory of complex systems applicable to the solution of global-scale problems. It will take a massively data-driven approach that builds on a serious collaboration between the natural, engineering, and social sciences, aiming at a grand integration of knowledge. This approach to real-life techno-socio-economic-environmental systems[8] is expected to enable new response strategies to a number of twenty-first century challenges.

BOX 1

Risk, systemic risk and hyper-risk

According to the standard ISO 31000 (2009; http://www.iso.org/iso/catalogue_detail?csnumber=43170), risk is defined as "effect of uncertainty on objectives". It is often quantified as the probability of occurrence of an (adverse) event, times its (negative) impact (damage), but it should be kept in mind that risks might also create positive impacts, such as opportunities for some stakeholders.

Compared to this, systemic risk is the risk of having not just statistically independent failures, but interdependent, so-called 'cascading' failures in a network of N interconnected system components. That is, systemic risks result from connections between risks ('networked risks'). In such cases, a localized initial failure ('perturbation') could have disastrous effects and cause, in principle, unbounded damage as N goes to infinity. For example, a large-scale power blackout can hit millions of people. In economics, a systemic risk could mean the possible collapse of a market or of the whole financial system. The potential damage here is largely determined by the size N of the networked system.

Even higher risks are implied by networks of networks[4,5], that is, by the coupling of different kinds of systems. In fact, new vulnerabilities result from the increasing interdependencies between our energy, food and water systems, global supply chains, communication and financial systems, ecosystems and climate[10]. The World Economic Forum has described this situation as a hyper-connected world[1], and we therefore refer to the associated risks as 'hyper-risks'.

[1]ETH Zurich, Clausiusstrasse 50, 8092 Zurich, Switzerland. [2]Risk Center, ETH Zurich, Swiss Federal Institute of Technology, Scheuchzerstrasse 7, 8092 Zurich, Switzerland.

Some background information