COMPUTATIONAL MODELING AND DATA STRATEGIES FOR PREDICTING AND RESPONDING TO PANDEMICS

SRINI VENKATRAMANAN

srini@virginia.edu
@sriniv_venkat

27 May 2020

UNIVERSITY of VIRGINIA

BIOCOMPLEXITY INSTITUTE

biocomplexity.virginia.edu

ACKNOWLEDGMENTS

Biocomplexity Institute & Initiative at the University of Virginia

- 60+ researchers spanning multiple disciplines studying massively interactive systems
- "From molecules to policy"
- Sustained program in computational epidemiology for nearly two decades funded by multiple US federal agencies (DoD, NIH, NSF, CDC)
- Real-time decision support for multiple outbreaks/epidemics: H1N1, Ebola, Zika, COVID-19

Network Systems Science & Advanced Computing Division

- Using the lens of networks to understand complex systems
- Simulation science aided by high-performance computing

National Institutes of Health

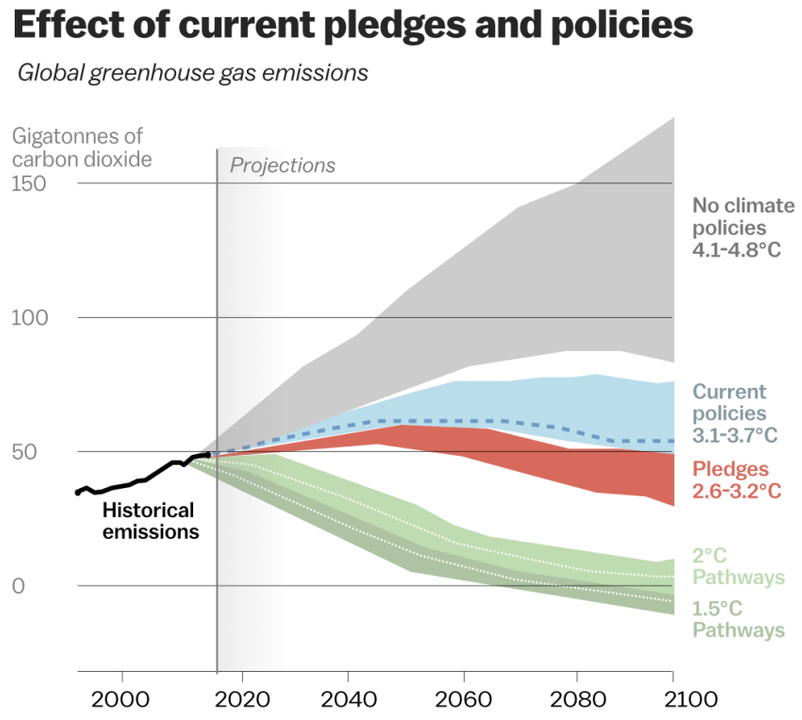
CDC Centers for Disease Control and Prevention

AccuWeather

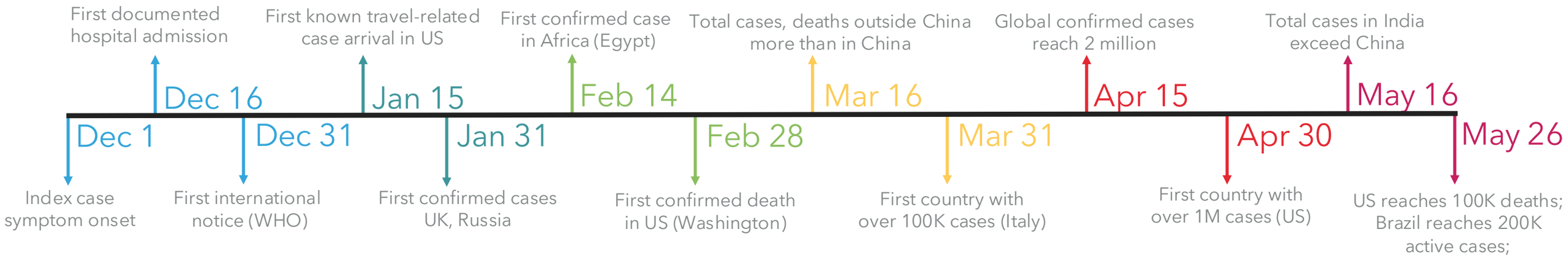
University of Virginia
BIOCOMPLEXITY INSTITUTE

LET'S TALK CLIMATE CHANGE!

- Climate change
  - Still the biggest existential threat facing humanity
  - "Flattening the Carbon Curve"
  - Alternative:
    - Bush fires, Hurricanes, Droughts
    - Loss of biodiversity, infectious diseases, chronic health risks
    - Food security, socio-economic instability, amplified disparities

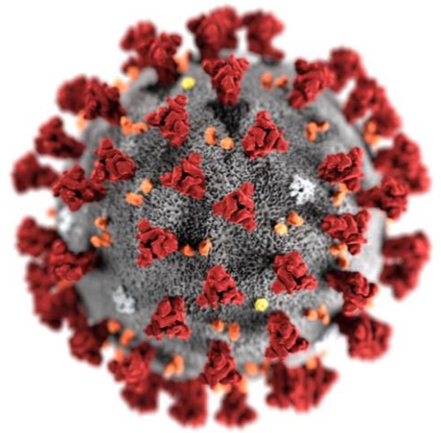


A BRIEF TIMELINE OF COVID-19



Why it became a pandemic?

- Novelty: All humans are susceptible; Delay in diagnostics;
- Transmissibility: Silent spreaders; Multiple modes;
- Connectivity: Dense population; Global travel;



Socio-political response

- Unprecedented lockdowns
  - Wuhan (Jan 23)
  - Italy (Mar 9)
  - ~Half of humanity (Apr 3)
- $2 trillion economic relief package in the United States

Scientific response

- 16000+ scientific articles; 6000+ on preprint servers between Jan and Apr 2020
- 161 vaccine and therapy candidates in R&D pipeline (mid-April)

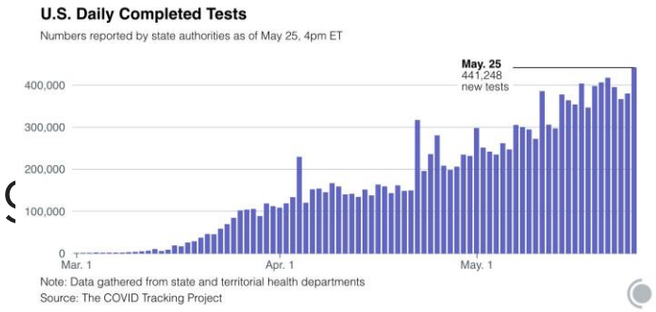
Technological response

- Contact tracing apps across countries and platforms
- Telework/Remote learning becomes the norm than exception

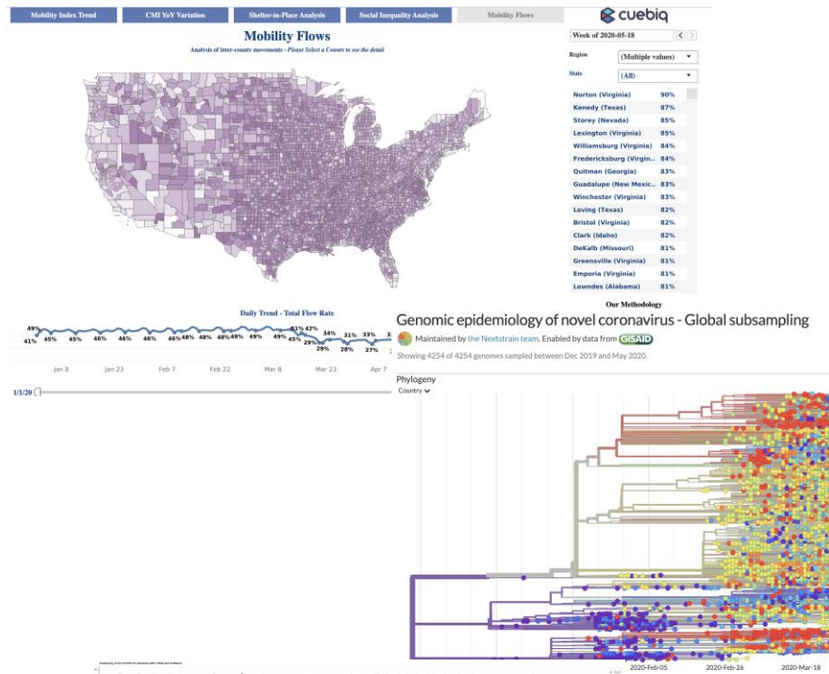
https://en.wikipedia.org/wiki/Timeline_of_the_COVID-19_pandemic
https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports

PLETHORA OF DATASETS

- **Case data, line lists**: Official sources, Crowdsourced repos, dashboards; Testing data; Racial/Ethnic data (COVIDTracking

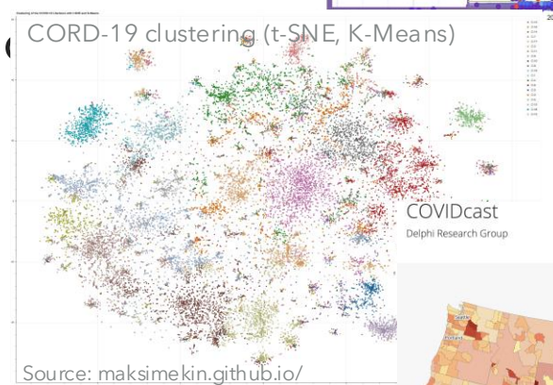


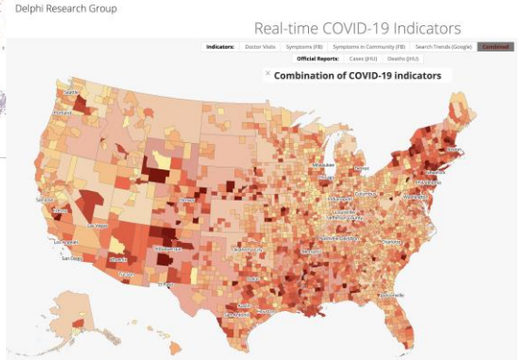
- **Mobility indices:** Google, Cuebiq, Facebook



- Genomic data: GISAID, NextStrain

- **COVID-19 literature:** Parameter estimates,CORD-1





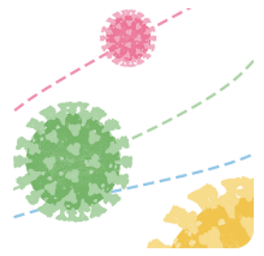
- **Social media:** Twitter, Search trends, Participatory surveillance (COVIDcast)

EMPHASIS ON MODELS

- Infectious disease models have been under scrutiny like never before

Where The Latest COVID-19 Models Think We're Headed — And Why They Disagree

FiveThirtyEight · 25 days ago



- Models guide social distancing measures and public policies

How COVID-19 models have changed behaviour and pandemic data

World Economic Forum · 14 days ago

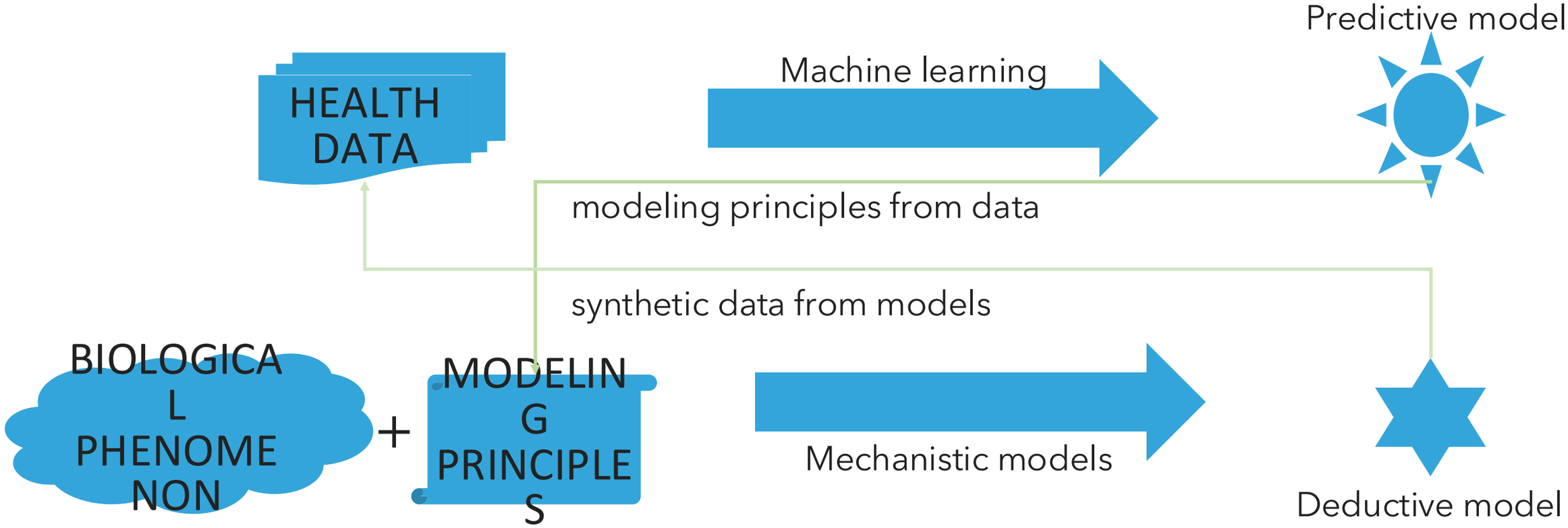
- Concepts like Herd immunity and ‘Flatten the Curve’ become part of public discourse.

How ‘Flatten The Curve’ Is On Its Way To Become A Historic Disease Image

Forbes · Apr 29

COMPUTATIONAL MODELS

Machine learning models	Mechanistic models
Establishing a statistical (correlation) relationship between inputs and outputs	Establishing a mechanistic (causal) relationship between inputs and outputs
Requires large datasets, often homogeneous linked records	Can work with small datasets coming from diverse sources
Can only make predictions that relate to patterns in the training data	Once validated, can be used for prediction where experiments are difficult to perform



Baker, Ruth E., et al. "Mechanistic models versus machine learning, a fight worth fighting for the biological community?." *Biology letters* 14.5 (2018): 20170660.

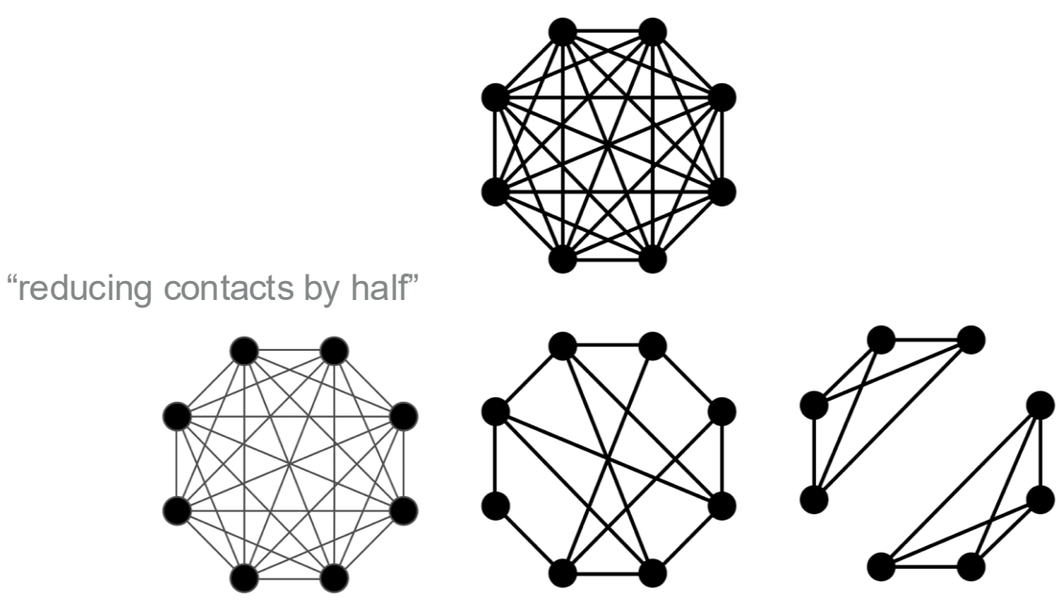
DISEASE MODELS PRIMER

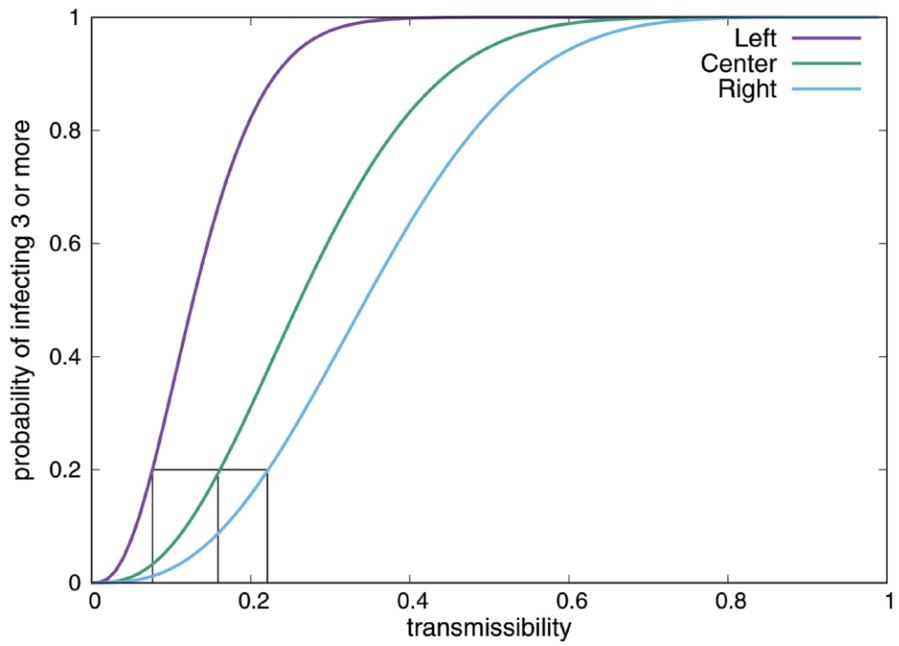
- SIR model – proposed in 1927 by Kermack and McKendrick

$$\frac{dS}{dt} = -\beta IS \qquad \frac{dI}{dt} = \beta IS - \gamma I \qquad \frac{dR}{dt} = \gamma I$$

β – transmission/mixing rate
$1/\gamma$ – mean infectious period

- Early insights into controlling the disease spread
  - Social distancing, Better protection = reduce mixing/transmission rate
  - Testing and Case isolation = lower infectious period
- Homogeneous assumption – misses network effects



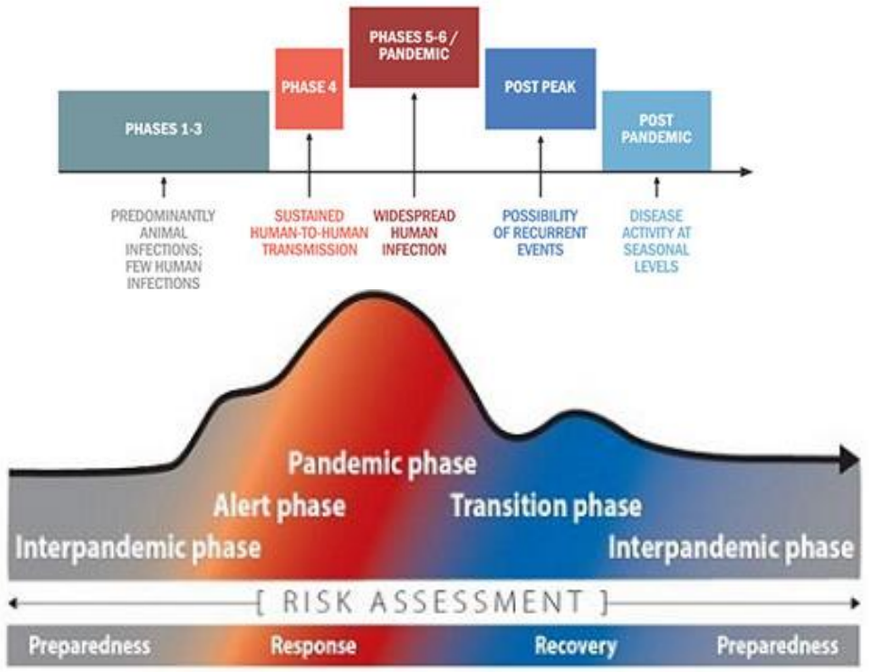


Source: Chen et al., Networked Epidemiology for COVID-19, to appear in SIAM News

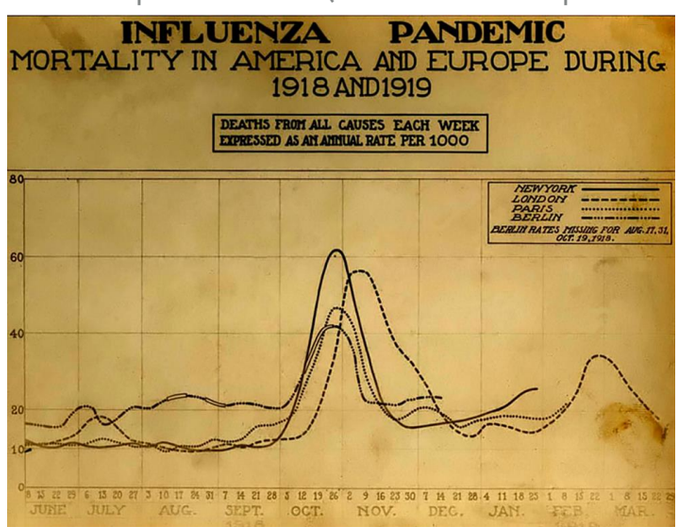
MODELING FOR PHASES OF A PANDEMIC

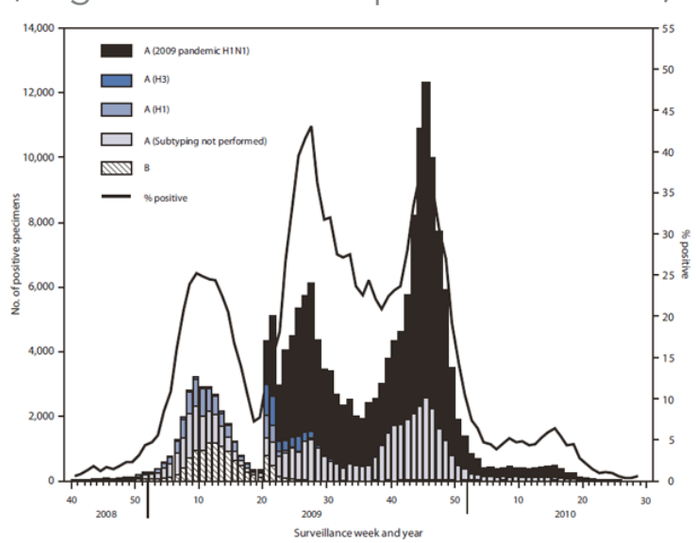
- Early phase
  - Risk of emergence (international travel)
  - Epi- and clinical parameter estimation
- Pre-peak phase
  - How long to peak? How many cases?
  - When will healthcare resources be exceeded?
- Interventions
  - What-If analyses on potential interventions
  - How have past interventions worked?
  - When to lift interventions? What next?

WHO Pandemic Influenza Risk Assessment

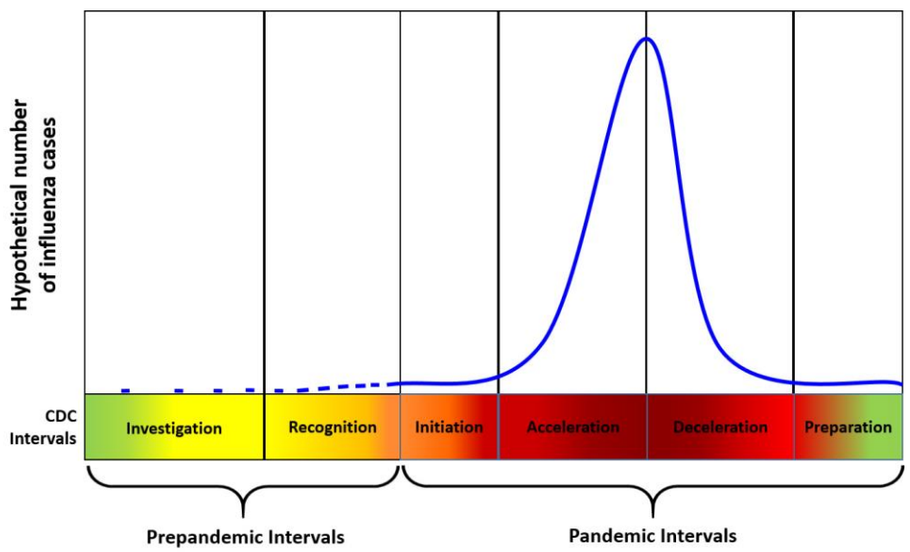


Multiple waves (Left: 1918 'Spanish flu'; Right: 2009 H1N1 pandemic in US)





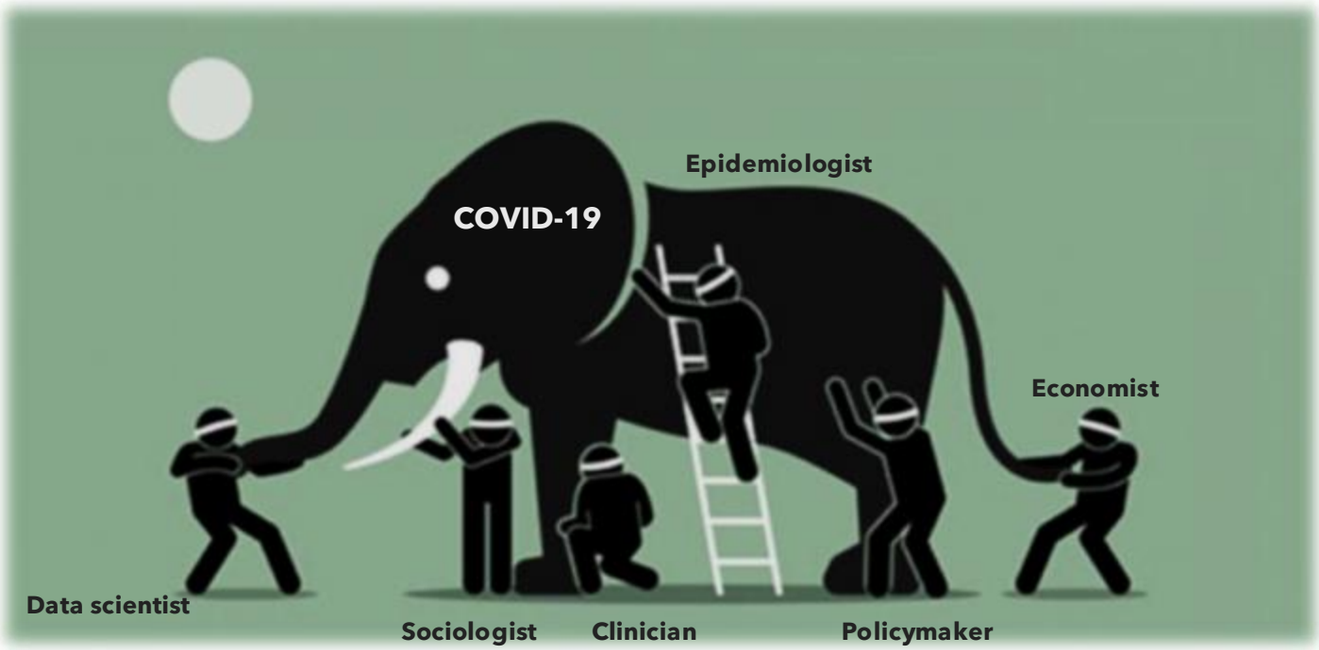
CDC Pandemic Intervals Framework

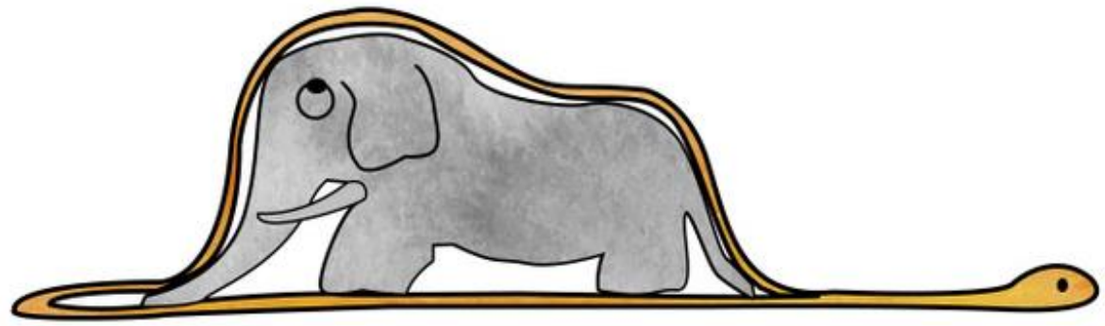


DATA AND MODELING CHALLENGES

5V's: Volume, Velocity, Veracity, Variety, Value

- How to handle the amount of data being generated, both in terms of magnitude and rate of change?
- How to verify and track the provenance of different datasets?
- How to integrate the varied datasets into a coherent view of the system?
- How to use the data meaningfully to guide decision-making?



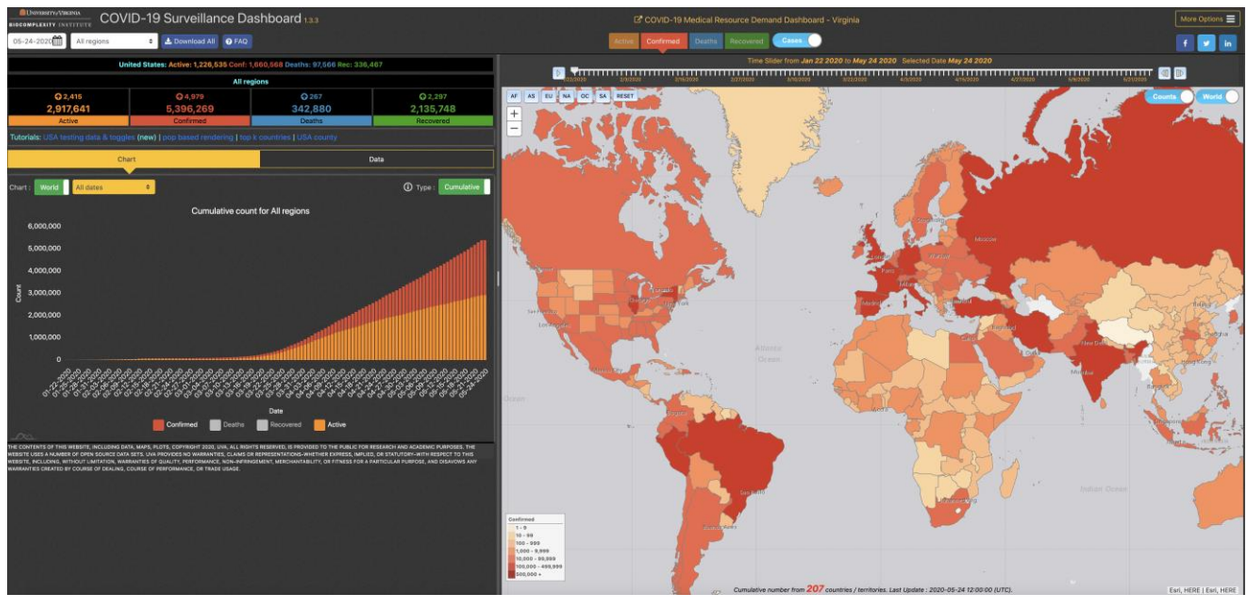


Source: The Little Prince (1943)

With four parameters I can fit an elephant, and with five I can make him wiggle his trunk. – John Von Neumann

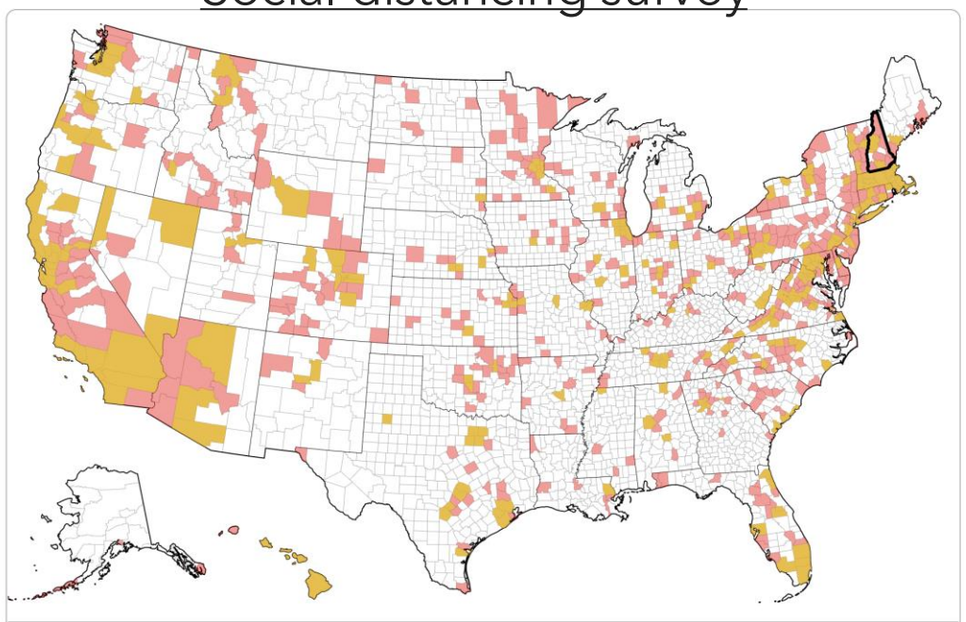
OUR INVOLVEMENT - DATA

Surveillance dashboard



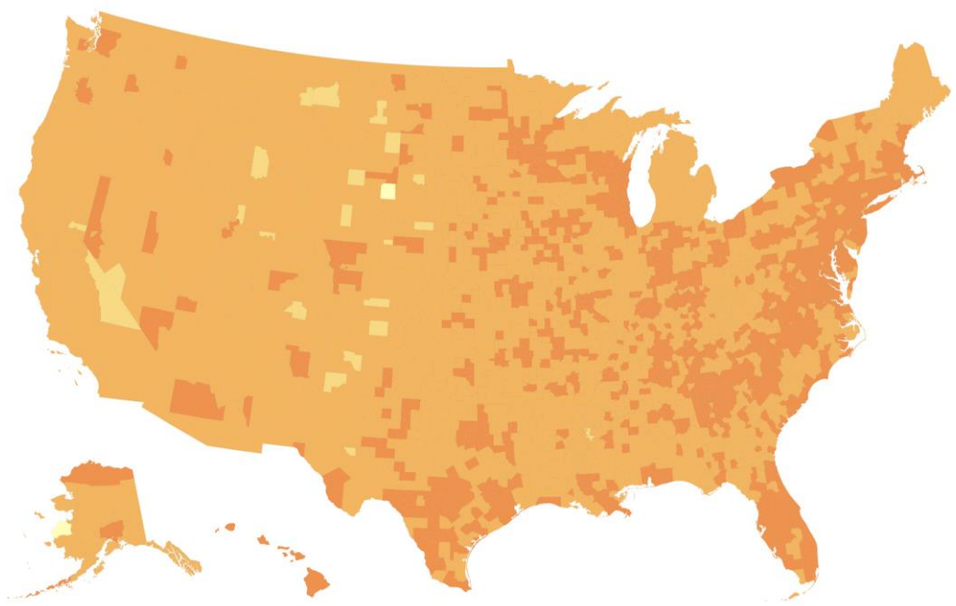
https://nssac.bii.virginia.edu/covid-19/dashboard/

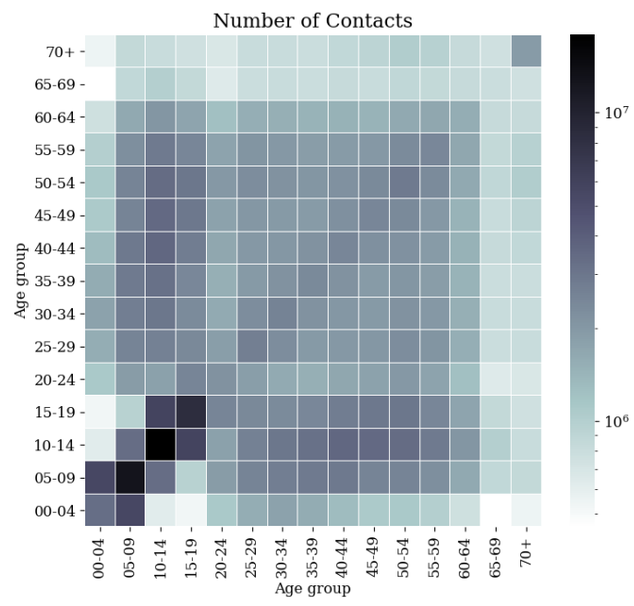
Social distancing survey



https://socialdistancing.stanford.edu/

Population contact rates by age and county

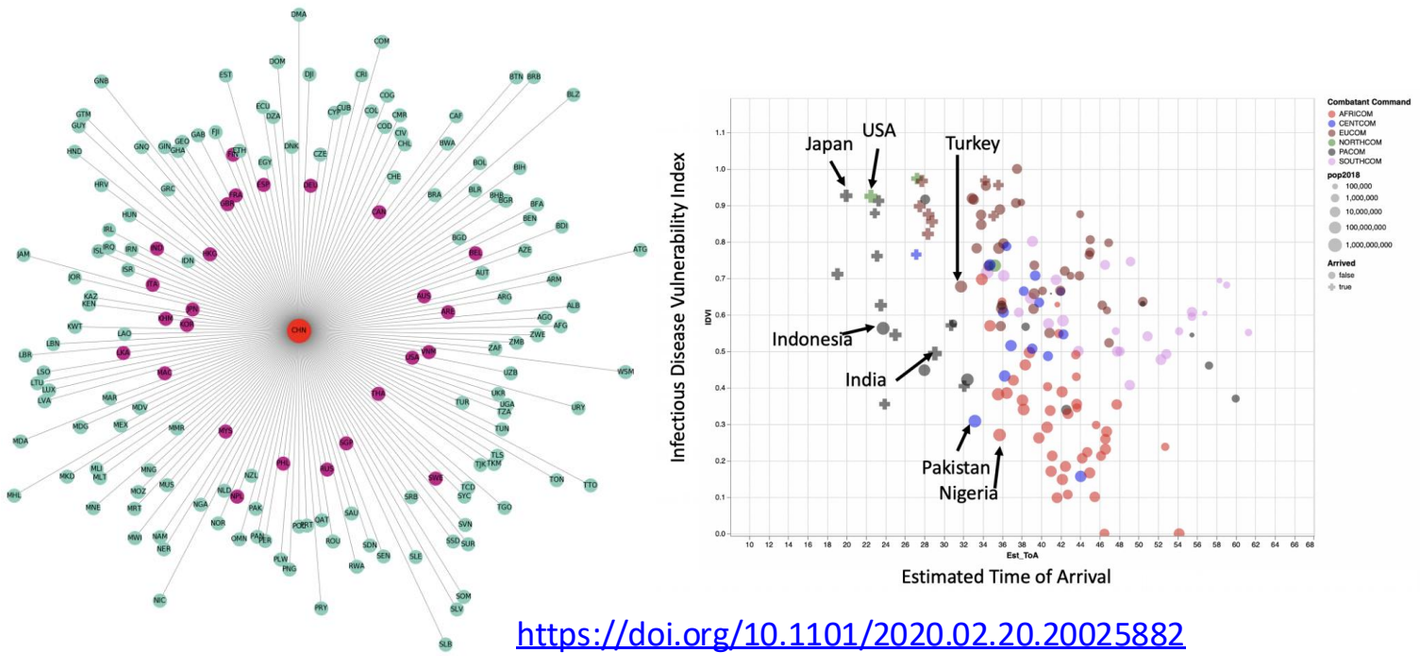




https://doi.org/10.18130/V3/OMEPOT

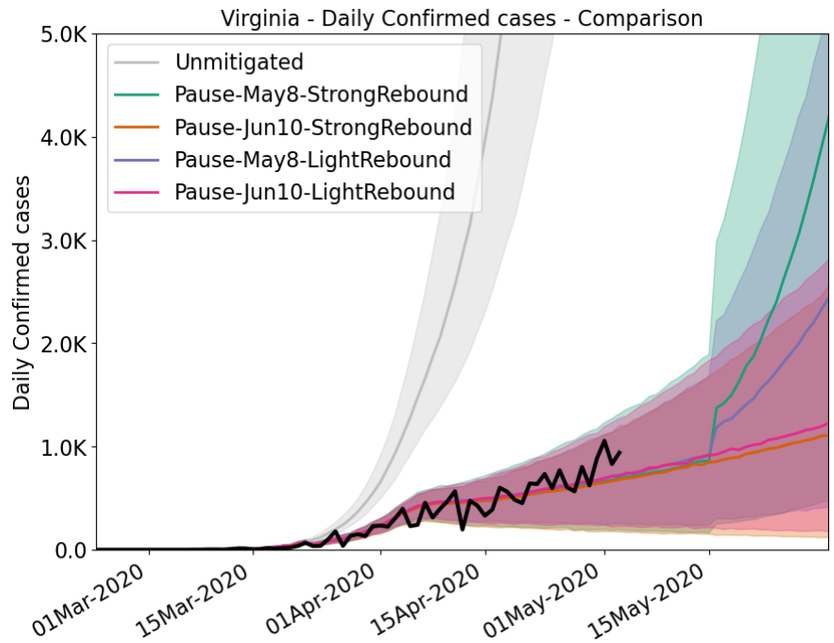
OUR INVOLVEMENT - MODELS

Early international risk assessment



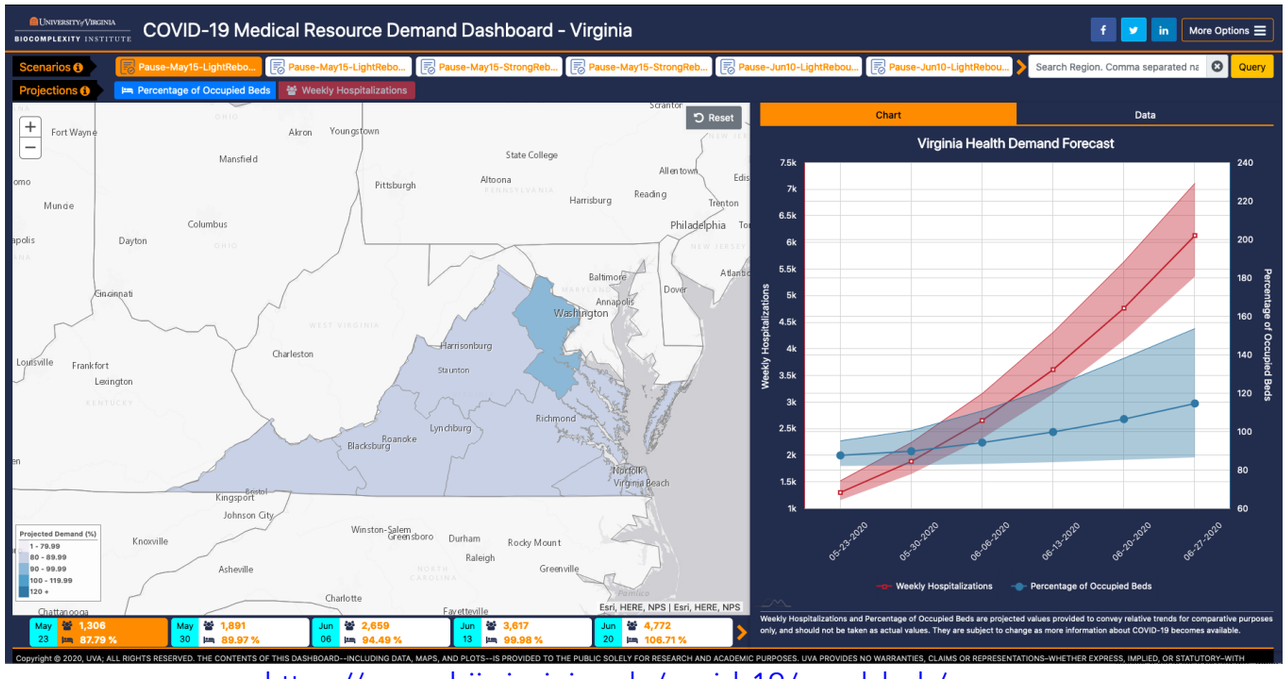
https://doi.org/10.1101/2020.02.20.20025882

Official modeling support for Virginia, DoD



https://www.vdh.virginia.gov/coronavirus/covid-19-data-insights/

Medical resource demand projections

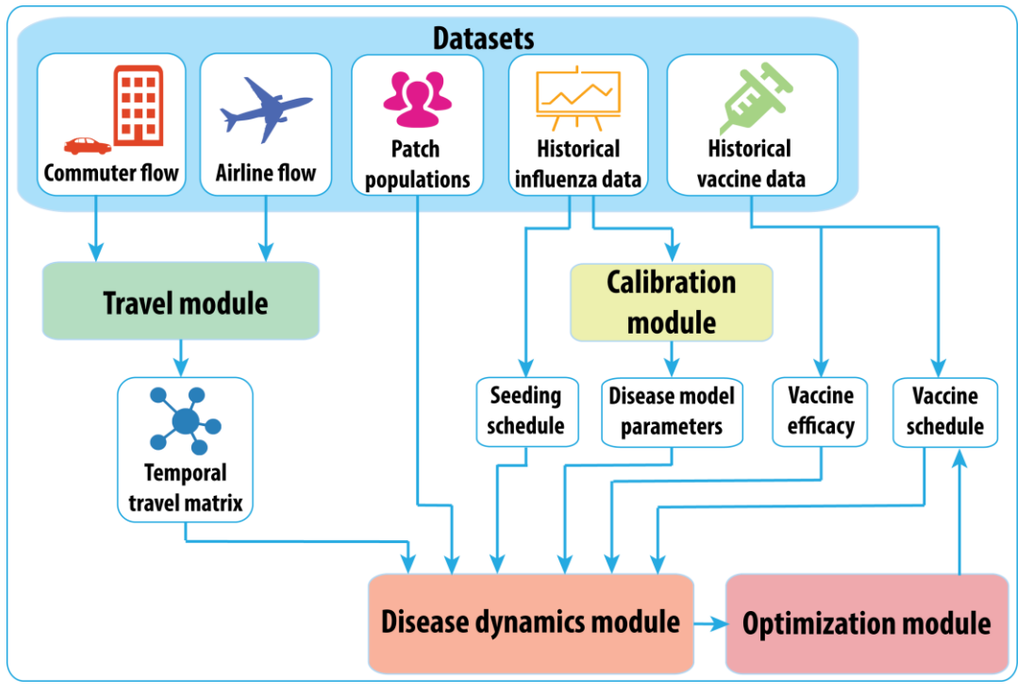


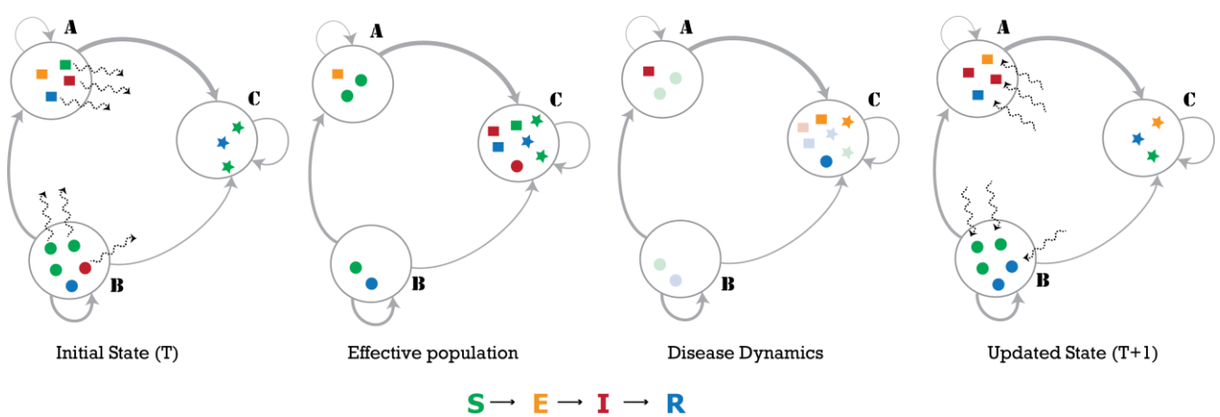
https://nssac.bii.virginia.edu/covid-19/vmrddash/

NETWORK BASED MODELS

Spatial heterogeneity

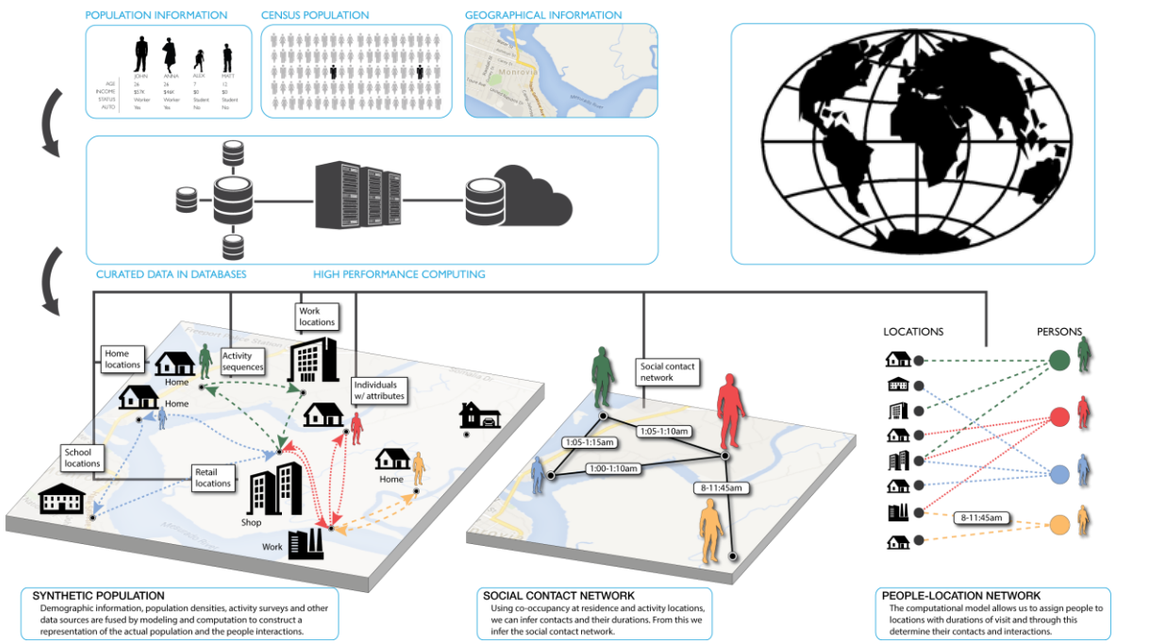
- County level metapopulation model for US
- Captures aggregate connectivity and mixing

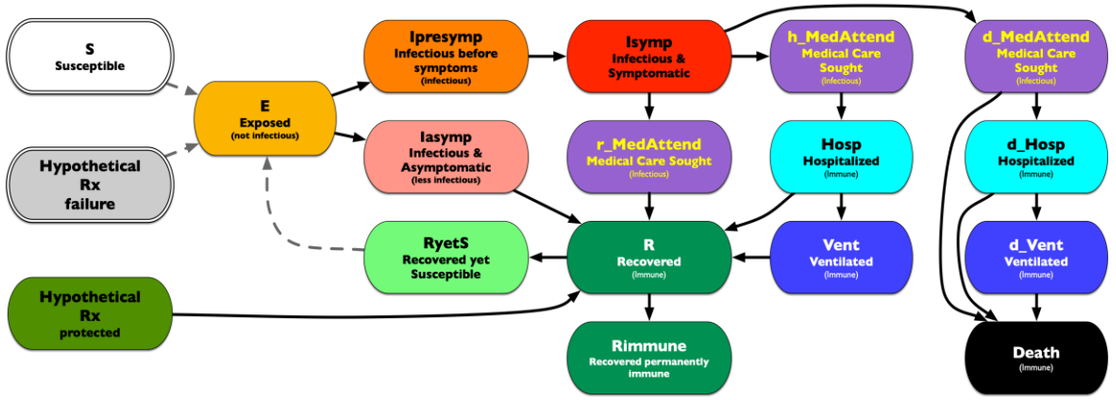




Social heterogeneity

- Individual level agent-based model
- Realistic activity schedules and social contact

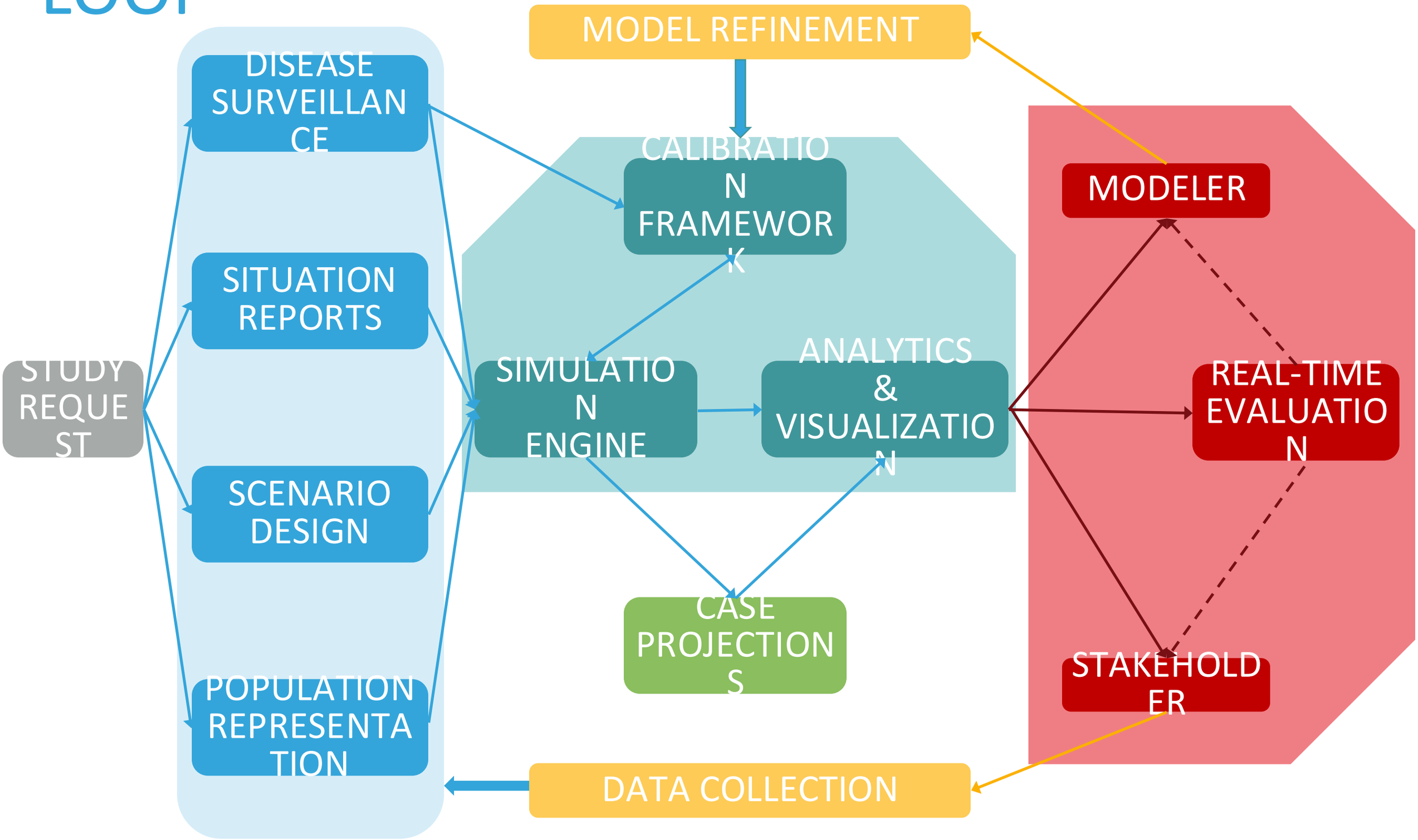




Venkatramanan, Srinivasan, et al. "Optimizing spatial allocation of seasonal influenza vaccine under temporal constraints." PLoS computational biology 15.9 (2019): e1007111.

Venkatramanan, Srinivasan, et al. "Using data-driven agent-based models for forecasting emerging infectious diseases." Epidemics 22 (2018): 43-49.

ABDUCTIVE APPROACH – HUMAN IN THE LOOP



CALIBRATING MECHANISTIC MODELS

Calibrating to real world data is challenging

- Data is noisy and inconsistent
- Stochasticity – Observing one instance
- Models can't explain everything

  nursing home outbreaks, racial disparities

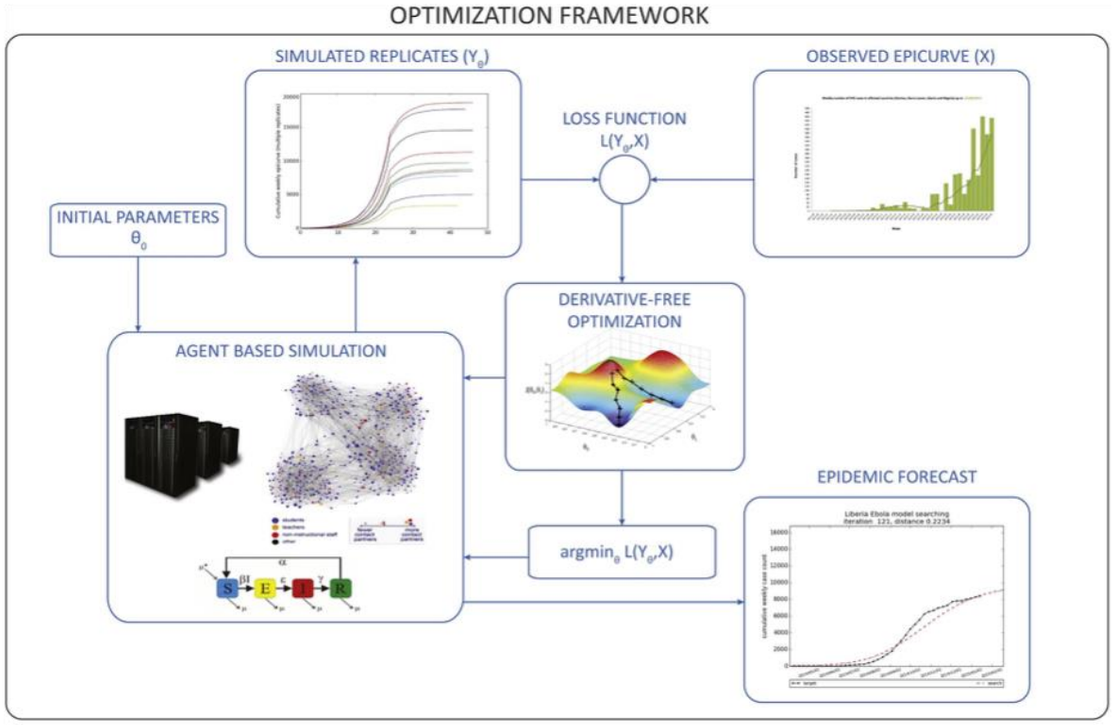
- Known unknowns (asymptomatic infections)
- Parameter identifiability (case fatality rate, reproduction number)

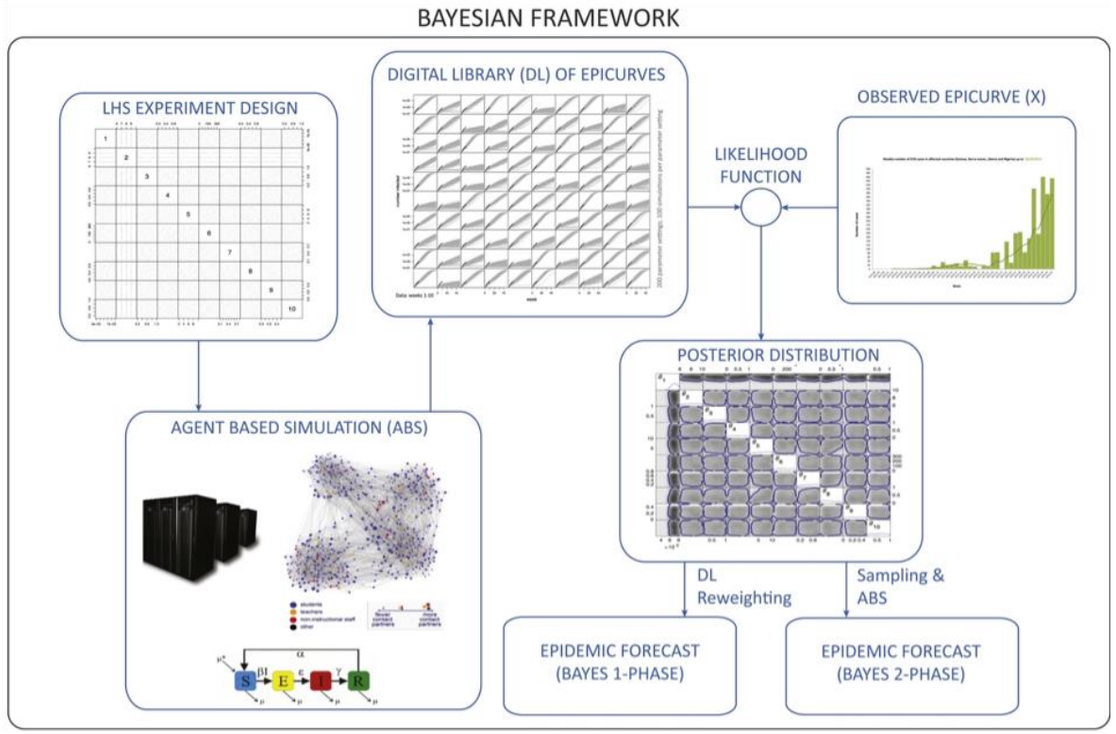
➢ **Leverage allied data streams**

➢ **Choose models with interpretability**

➢ **Quantifying uncertainty in projections**

➢ **Compare and contrast multiple models**





Venkatramanan, Srinivasan, et al. "Using data-driven agent-based models for forecasting emerging infectious diseases." Epidemics 22 (2018): 43-49.

FRONTIERS

STATISTICS

- Latent relationships and anomaly detection
- Missing data; Discrepancy modeling
- Uncertainty quantification; Model ensembles

MODELING

- Modeling coupled co-evolving complex systems

  Behavior → Information → Epidemics; Info-demics.

  Socio-economic cascades; Socio-political epidemiology

- Control theoretic view; Feedback systems

COMPUTING

- Seamless flow from data ingestion to simulation analytics
- Robust pipelines using cloud infrastructure
- Privacy preserving computations on-device
- Common collaborative platforms; AI for assisting teams;

LESSONS (BEING) LEARNT

- Ensembling approaches prevail over others
- Theory guided Data science (TGDS) bridges the domain experts and data scientists
- Human in the loop is unavoidable. Wisdom of (expert) crowds.
- Data summarization (visualizations) is key to make decisions (or even make sense of the data)
- Team Science is essential to address real-world problems. Rethinking credit sharing practices.
- Open source environment (papers, code, data) accelerates research and ensures reproducibility. Incentive mechanisms!
- Automated pipelines must be built with scalability, robustness, development and testing in mind.
- Translational research: Start from the problem. Not from the data. From papers to practice.

WAY FORWARD

- Pandemic far from over; Multiple waves likely
- Impact of weather unclear; May not help reduce overall burden
- Test, trace and isolate still the best way to avoid lockdowns until pharmaceutical interventions

- Global epidemics are becoming increasingly common;
- There are no individual experts; Emphasize on evidence-based decision making;
- Sustained efforts needed now and beyond (during peacetime);
- Centers for Forecasting and Responding to Epidemics (similar to hurricane center);

QUESTIONS?

srini@virginia.edu
@sriniv_venkat

https://biocomplexity.virginia.edu/