PREPARE
Pandemic Research for Preparedness & Resilience

Kick-off Workshop

15-16 December 2020

Summary Report

PREPARE Team
Madhav Marathe, UVA
Simon Levin, Princeton Univ
Anil Vullikanti, UVA
Li Xiong, Emory Univ
Erin Raymond, UVA
Golda Barrow, UVA

Workshop Program Committee
Srini Venkatramanan, UVA
Fei Wang, Weill Cornell Medicine

This material is based upon work supported by the National Science Foundation under Grant No. CNS-2041952.

Table of Contents

Executive Summary

Research Roadmap: Computing for preparedness and resilience for the next pandemic.

When we speak of the next pandemic, it's no longer a question of "if" but "when". The NSF research community is uniquely situated to ensure that the multidisciplinary and cutting-edge science completed through the CISE directorate can help society address the challenges of the next pandemic. From online education to social media misinformation to vaccination distribution, CISE scientists are poised to make a tremendous impact through the creation of a pandemic roadmap.

The Pandemic Research for Preparedness and Resilience (PREPARE) roadmap will serve as a blueprint for researchers, funding agencies, and policy makers on the role of computing in developing breakthrough solutions in this area.

This inaugural virtual workshop was open to PIs who have received NSF RAPID grants within the CISE directorate as well as researchers from academia, government, and industry who are interested in furthering efforts to identify and address preparedness and resilience for our nation and the world. The workshop included a venue for posters and lightning talks, remarks from NSF leaders and the PREPARE PIs, a keynote address, and breakout sessions to discuss challenges, gaps, and future directions of research and innovation.

Workshop Overview

The focus of the inaugural PREPARE workshop was to provide a platform for scientists to share their work, learn about cutting edge research, discuss the challenges and opportunities for future work, and develop potential collaborations. Given the geographic disparity and multidisciplinary nature of the participants, the program committee worked hard to find innovative ways to leverage available virtual platforms to maximize participation by attendees.

The two-day NSF-sponsored workshop was conducted virtually on 15 and 16 December using Zoom as the primary meeting platform and Gather.town as the poster session platform. The main program committee included Li Xiong (Emory University), Fei Wang (Weill-Cornell Medicine), Anil Vullikanti (UVA), Srinivasan Venkatramanan (UVA), Erin Raymond (UVA), and Golda Barrow (UVA). The first day of the workshop we had 126 unique attendees login to Zoom; the second day we had 210 unique attendees.

Prior to the workshop, we sought input from the CISE RAPID PIs to determine topics for the breakout sessions. In order to give all researchers an equal chance to present their work, we provided a website for lightning talks and posters that was accessible in the days before the workshop (PREPARE Dec 2020 Poster Session). There are 56 posters available to view. By keeping this site active, we hope to foster collaborations among all researchers interested in preparedness and resilience, even those who could not attend the workshop.

Opening remarks from Dr. Margaret Montonosi, NSF Assistant Director for CISE, Dr. Gurdip Singh, NSF Division Director of CNS, and our PREPARE PIs are available on our YouTube channel. The keynote address by Sir Roy Anderson FRS, Imperial College London is also available there. (Due to technical issues we do not have the recording of closing remarks from Dr. Erwin Gianchandani, Senior Advisor to NSF Director & Deputy Assistant Director for CISE). The full agenda can be found at the end of this report.

Breakout Sessions

A primary component of the workshop was the topical breakout sessions. The outcome of these sessions will help the PREPARE team to identify critical research paths that can become topical areas for future workshops. Using information from the NSF CISE RAPID PIs, the program committee identified the following breakout sessions:

I-A Computational Biology and Bioinformatics
I-B Epidemiology and Public Health
I-C Education, Training, and Workforce Development
I-D Computing and Data Infrastructure
II-A Privacy and Security
II-B Socioeconomic and Psychological Impact
II-C Surveillance, Contact Tracing, and Testing
II-D Infodemiology and Social Networks

Moderators were recruited for each session, and the program committee requested that they focus discussion on addressing the following questions:

- What are some of the data and computing challenges related to the pandemic?
- What CISE techniques have been used/deployed, and worked or helped the pandemic?
- What can be used/deployed to help? What are the challenges for using them?
- What needs to be done for building resilience for future pandemics?
- What are the new emerging CISE topics/challenges due to the broader economic and social disruption of pandemic and the adaptations?
- How does computing research complement clinical research?
- What are some coordination activities that can help (e.g., connecting with stakeholders, researchers in other areas)?
- What kind of support might be needed from NSF and other agencies?

Following the breakout sessions, one moderator from each group was asked to present the most salient points of their discussion. Subsequently, the groups created the following summary documents.

I-A Computational Biology and Bioinformatics

1. Data and computing challenges
   - More partners/stakeholders could be involved to hasten development of vaccines & tools for detection
     - Machine learning to identify best viral targets for vaccine
     - Stopgap vaccines
     - Assessing viral similarity or molecular mimicry to identify similar vaccine strategies
   - Data sharing to expedite research
     - COVID-ARC = platform for dissemination (https://covid-arc.loni.usc.edu/)
       - *Data*: Multimodal include imaging (CT, Xray), clinical
         - Range of file formats (.nifi, .dicom, .tiff, .png)
         - International datasets (China, Israel, Germany, Iran, USA, etc.)
       - *Privacy*: host can allow share publicly or to specific users
         - Centralized or federated models
       - *Features*: Quality Control, Analytics within site
     - Harmonization to address collection variability
     - No comprehensive dataset available with all modalities (different datasets for imaging, sequencing, etc.)
       - Difficult to conduct multimodal studies with non-linked datasets
     - Converted file formats (.png, etc.) are less useful for analysis (unclear whether noise is due to conversion or original acquisition)
     - Guohua Cao: interested in artificial intelligence to analyze poorly acquired images (collaboration with Duncan @ USC)

2. What CISE techniques have been used/deployed and have helped?
   - Efforts towards RAPID grants for quick deployment has been fruitful
   - Better software pipelines for finding vaccine targets to new pathogens are needed

3. What can be used/deployed to help? What are potential challenges?
   - Readily available data on infections (e.g. Johns Hopkins website)
   - Poor availability of data of different levels of quality to train ML programs
   - Phenotype challenge: Methods to connect infection strains and host genetics to symptoms and expression of the disease

4. What needs to be done for building resilience for future pandemics?
   - Infrastructure in place for multimodal data archive + analytic tools + hardware solutions
     - Readily available sequencing data
     - Rapid mobilization of resources for quick analysis
     - Develop hardware solutions for sample & sequence acquisition & analysis

- Leverage existing knowledge base (SARS-CoV-1) to identify similarities
- Faster funding for academia (COVID sequencing, etc.)
- Support for building pathogen testing infrastructure -- one that can detect the emergence of novel pathogens -- similar to early warning systems we have in place for detecting earthquakes.

5. What are the emerging CISE topics/challenges due to the broader economic and social disruption of pandemic and adaptations?

- Fatigue with social regulations (long-term isolation)
    - Identify relationships between business regulation with infection spread
- Identify groups at most risk
    - Epigenetic predispositions
    - Disparate access to healthcare resources

6. How does computing research complement clinical research?

- Public datasets (COVID-ARC)
    - Longitudinal data for outcome prediction and intervention assessment
- Necessary for vaccine development
- Should be considered for allocation of grants
- Cloud-based computing for in clinic, real-time analysis (clinically actionable results)
    - AWS, Google, Chameleon Cloud (https://www.chameleoncloud.org/)
    - On site infrastructure, ready for deploying with novel pathogens
    - Robust to diverse pathogens
    - Costly, need accessible platforms or specified funding requests
- Increasing specificity/sensitivity of detection
    - Higher false negative rate, low false positive rates
    - Modeling can help address these challenges

7. What are some coordination activities that can help?

- Added challenges in remote areas; coordination with urban centers
    - knowledge/data transfer (centralized or federated approaches)
- Coordinated dissemination of patient data, COVID sequences (strain information as well)
    - Greater infusion of funds at an early stage could ensure that data is available to a wider research community

Participants: Giri Narasimhan (FIU), Dominique Duncan (USC), Fei Wang (Weill-Cornell), Guohua Cao (VT), Deep Medhi (NSF), Satish Narayanasamy (U Michigan), Archna Bhatia (IHMC), Rachael Garner (USC), Wei Wang (UCLA)

I-B Epidemiology and Public Health

The first discussion revolved around data: "The Data Challenge". While we all agreed that there is more data available than for any pandemic before, there is still data that is not available to researchers. Some questions we asked:

- How to collect data?
- What do they mean exactly? For instance, in the daily count are all cases detected the same day or the ones that have been reported that day (then you need to trace back to when the infection happened...). Hospitalizations data is hard to get (and can be inconsistent or unclear). For testing, could we account for individuals with repeated tests? Spatial information on daily cases is usually not provided.
- How to share data? There are many dashboards, many of them are not consistent but still are helpful. Much improvement has been made over the past few months. Spatial data is claimed to breach confidentiality but it is unclear why.
- What kind of data? Modalities?
- How to make Public Health Authorities interested? Feedback? Showcase some scenarios and results that could be produced if we had access to data would create an incentive.
- Is it worth investing in "social" data? It requires a heavy machinery to analyze

Bottom-line: which data is most essential? Researchers doing modeling should come up with a list of what data are most important. We will not get access to everything so we need to express which data are critical and convey this information.

This discussion revolved around the role of models: "The Model Challenge":

- What are the questions to ask within the models? Why are we using the models for?
- What is the spatial and temporal resolution?
- What class of model? As of now Granted Agent Based Models (ABM) (mechanistic) is the ground truth (synthetic). There is no universal model, it depends on what question we are trying to answer. Some require a lot of preparation and development of a synthetic population while others are easier to manipulate but are more limited.
- Models are good at producing different scenarios and highlighting how different measures might impact the evolution of the spread. There are not good quantifiers because of the uncertainties and limitations of data.
- How general is the model? Will it apply to the next/other epidemics?

Models are currently unable to reproduce the evolution of a virtual pandemic for which we know everything. This is a major issue and highlights the complexity of the system.

The next discussion was about the complexity of models: "How simple should a model be?":

- Are ABM and Mean-Field models two complex? You need major investment in order to run them and understand the results.
- Do we need models with "in-between" levels of resolution (in time, space and also in epidemiological details).

- Simplicity vs nuances. What to include in the reduced, coarse-grained models?

The suggestion is that this could be a new direction for the MIDAS community.

The next discussion was triggered by a visit from an individual from the "Computational Biology & Bioinformatics" breakout room. The question was "Where is the overall bottleneck?":

- Is it due to limitation in data (not available)?
- Is the proper model missing?
- Are we asking the right questions?
- Accuracy? Quality?
- How to quantify uncertainty.
- Question/answer – not binary.

We did not come with a conclusive answer, the fact that models cannot predict even a synthetic pandemic is due to many factors.

The last part of the discussion revolved around a wish list: "What do we Wish?":

- Make results of the modeling + efforts tilted towards decisions(public health authority). How do we communicate results so that they will be meaningful and understandable to decision makers?
- Dashboard prediction of a kind. "Flattening the curve" is the only metaphor and the poster child on how a simple visual was able to influence mitigation measures as well as be understood by the population. We need more examples like that. What should they be and how should they look like?

Participants: Michael Chertkov (U Arizona), Monique Chyba (U Hawaii), Chi-Ren Shyu (U Missouri-Columbia), Eyal Oren (SDSU), Mac Hyman (Tulane), Madhav Marathe (UVA), Gautam Dasarthy (ASU), Wendy Ju (Cornell Tech), Simon Levin (Princeton), Alon Efrat (U Arizona), Anthony Morciglio (Georgia St Univ), Chaowei Yang (GMU), Ilan Mandel (Cornell Tech), Krishna Narayanan (Texas A&M), Li Xiong (Emory), Philip Pare (Purdue), Sanjiv Kapoor (IIT), Yaling Yang (VT)

I-C Education, Training, and Workforce Development

Highlights of discussion

- High-quality virtual social spaces that allow for meaningful interaction are in short supply, and this heavily impacts online learning as well as gatherings of grown-ups.
    - One RAPID supported development of an alternative
- Online K-12 education needs more attention than the more heavily-studied college / adult online learning space.
    - There is much more still to learn, and much more need to address issues of equity.
        - (See notes from breakout II-B: Socioeconomic & Psychological Impact)
    - Lots of data / experiences, there is risk that things will be lost / forgotten without concerted efforts to gather best practices and challenges
    - Education about the pandemic: misinformation spreads so easily, addressing this could be done through integrating with cybersecurity / digital citizenship curriculum in K-12
    - What can higher ed learn from K-12 online teaching practices, and vice versa?
- Idea of "strategic computing reserve" could apply to distributing computing resources to educational systems (devices, online spaces, etc.)
- Ways NSF / other agencies can support
    - Need to connect across divisions with an education research focus. CISE is somewhat unique in providing opportunities to develop technologies to address educational challenges studied in other divisions
    - Long-term studies will be needed to validate both what was most effective in education during this pandemic (retroactive), but also for ongoing study of online spaces that will sustain after this pandemic (proactive)
    - How can we incentivize junior faculty to focus in this area?
        - Grad / undergraduate students could be a tremendous resource when engaging with K-12 schools.

Guiding questions / notes

- What are the adult and youth education challenges posed by a pandemic?
    - Education about the pandemic
    - Education during the pandemic
        - Zoom is not ideal! What works well for getting people online meaningfully?
            - Most platforms designed for single audiovisual events / streams. Tools are not integrated
        - We have tidbits of knowledge, but not long-term designed to learn about this kind of learning
        - Youth education is so hard! Having students stay engaged in this format is very challenging.

        - Default is to replicate physical rooms in an online environment. Teachers could move between classes rather than having students move
        - Could explore different models for interaction / engagement other than A/V streaming
    - Providing policy / best-practices guidance in a federalist country. Contexts change from state to state, district to district, school to school.
    - Educators have different levels of flexibility
        - Are teachers incentivized to do extra work / teach differently?
- What needs to be done for building resilience for future pandemics?
    - We need connectivity everywhere. Students without devices had to do asynchronous learning, go to physical buildings, etc. Compounding inequities.
        - High-quality local networks may be an alternative
    - Ongoing research on online learning that will be left over from pandemic
- How does computing (education) research complement clinical research?
    - Most professors don't relish exams in fully online learning
        - Need different models of assessment for online learning (tests are also not ideal for assessment during in-person instruction)
- What are some coordination activities that can help (e.g., connecting with stakeholders, researchers in other areas)?
    - Connect across education-focused programs / subdivisions within NSF
    - Guidance from educational and racial equity scholars
    - Can we transfer learning about online learning from higher ed to K-12?
- What kind of support might be needed from NSF and other agencies?
    - Continue to fund large-scale implementations of virtual environments.
    - K-12 has priority over better-studied college space. We have more to learn, and there is more need in K-12
    - Are universities acknowledging that this is a priority? Will there be space for junior faculty to pursue this kind of educational research?
    - Getting grad / undergrad students to help in K-12 will likely be more fruitful than directly engaging faculty
    - What can we learn from K-12 students and teachers that applies more generally?
        - What are agencies outside of CISE studying?
    - Strategic computing reserve
        - Computing infrastructure / IT support for open-source conference platform
        - Ad hoc virtual spaces for emergencies
        - Distributed infrastructure getting where it is needed
            - Framework for requesting resources
- What are some of the data and computing challenges related to the pandemic?

    - Lots of data / experiences are happening in K-12, and it's equally challenging and important to access this information. A few companies are getting very data-rich and it's not connected to research
        - CL: creating open-source conference platform to address this challenge
- What CISE techniques have been used/deployed, and worked or helped the pandemic?
    - The internet exists :)
    - Clowdr.org
- What can be used/deployed to help? What are the challenges for using them?
    -
- What are the new emerging CISE topics/challenges due to the broader economic and social disruption of pandemic and the adaptations?
    - Social virtual spaces are very lacking (in education and in general)
    - Spread of misinformation / how to be a good online citizen

Participants: Ashutosh Dhekne (GT), Wendy Guan (Harvard), Kyle Johnsen (U Georgia), James Joshi (NSF), Crista Lopes (UCI), Andy Rasmussen (Chicago Public Schools), S. S. Ravi (UVA)

I-D Computing and Data Infrastructure

What are some of the data and computing challenges related to the pandemic?

- Scalability - infrastructure for storage and computation
- Delay in delivery of components
- Direct methods of sensing: illness statistics, contact tracing, etc.
- Alternative means of sensing pandemic-related information: population mixing, work-from-home compliance, network status/overload
- Reliability and uniformity of the data
  - Adversaries trying to intentionally corrupt the data
  - Coherent and transparent sharing (along with some info)
- Kinds of data and use: Epidemiological data, contact tracing, social networking, surveys (say mask usage compliance), clinical data (CT scans, …)
- Business and educational continuity
- Economic, socio-demographic, supply chain
- Access to sensitive data - Establish protocol for future
  - existing models by Unacast etc. for data sharing, access control
  - Environment for running experiments
- Where to perform computing? Where will data reside post-pandemic?

What CISE techniques have been used/deployed, and worked or helped the pandemic?

- Remote work and virtual social life (Zoom, …)
- HPC infrastructure
- Sensing and mobile computing
- Networking and data management
- Mobile contact tracing
- Data Visualization
- Reaction from social media on different strategies (such as wearing masks)
- Educate people

What can be used/deployed to help? What are the challenges for using them?

- Recommendation systems (nudges to encourage desirable behavior such as practising masking, social distancing)
- Data visualization based on these measures
- Reuse of existing infrastructure to apply to pandemic-related challenges
- Bridge gap between data analysts and health officials
- Access to high-speed internet especially in under-priviledged societies
  - Infrastructure is limited
  - Place for isolation or working is also a challenge
  - Minimizing infrastructure costs

- Domestic violence and sensors to monitor things like mental stability and stress to minimize it
- Adapting by service industries

What needs to be done for building resilience for future pandemics?

- A ‘centralized’ and individual emergency response for workplace, schools, kids both
- A framework for effective response with lessons learned and best practices

What are the new emerging CISE topics/challenges due to the broader economic and social disruption of pandemic and the adaptations?

- Low-cost computing and network infrastructure that is affordable
- Educational infrastructure supporting students at all levels

How does computing research complement clinical research?

What are some coordination activities that can help (e.g., connecting with stakeholders, researchers in other areas)?

- Access to industry/health data
- Protocols to efficient and secure data sharing

What kind of support might be needed from NSF and other agencies?

- Transition of successful results from RAPID for future
- Workshops from transition teams
- Phase 2 funding for continuity

Questions that should have been asked:

- What are the kinds of data and what is it used for?
- (other questions that were missed?)

Kinds of data:

- illness counts over time
- government recommendations over time
- social networking
- individual contact
- direct Internet measurements

Uses of data:

- epidemiology modeling
- economic activity modeling
- Internet resilience
- Educate people

Sharing of data:

- Security
- Rapid sharing

Participants: Abhijit Suprem (GT), Alice Koniges (U Hawaii), Anil Vullikanti (UVA), Archanaa Krishnan (VT), Arif Sadri (FIU), Aryya Gangopasdhyay (UMBC), Calton Pu (GT), Chen Li (UCI), Chris Barrett (UVA), HB Acharya (RIT), Hafiz Asif (Rutgers), Indrajit Ray (NSF), Indrakshi Ray (Colorado St Univ), Jaideep Vaidya (Rutgers), John Heidemann (USC-ISI), Mimi McClure (NSF), Murat Kantarcioglu (U Texas Dallas), Raphael Stern (U Minnesota), Sai Manoj Pudukotai Dinakarrao (GMU), Shantanu Sharma (UCI), Sharad Mehrotra (UCI), Srini Venkatramanan (UVA), Vivek Singh (Rutgers), Xiao Song (USC), Yu-Ru Lin (U Pittsburgh)

II-A Privacy and Security

What are some of the data and computing challenges related to the pandemic related to security and privacy?

- Data collection and sharing challenges due to privacy concerns.
- Privacy risk assessments for publicly shared data
- The quality of the data and the sensitivity of the data
- Fine grained data availability for research and operational purposes
- Sensitivity and Robustness of analysis, especially with respect to incomplete and inaccurate data while protecting privacy.
- Vulnerability of analysis to misinformation and targeted noise/inaccuracies within the data
- Cyber attacks against vaccine research and deployment (e.g. supply chain).

What CISE techniques have been used/deployed, and worked or helped the pandemic?

- Cryptography
  - Encrypted Data processing
  - Secure Multi-party Computation for Contract Tracing
- Secure Data Processing using Trusted Execution Environments (TEEs) and enclaves
- Anonymization and Data Redaction
- Differential Privacy
- Privacy-risk analysis models
- Privacy-preserving machine learning techniques
- Data Visualization for Situational Awareness
- Policy based Data Sharing
- Blockchains
- Malware and Intrusion Detection Tools
- Vulnerability analysis for critical infrastructure
- Phishing and Spam Detection

What can be used/deployed to help? What are the challenges for using them?

- Usability
- Trust and Concerns regarding misuse
- Unacceptable Utility vs Privacy Trade-offs for some existing tools
- Availability
- Compatibility with systems in place
- Cost of adoption
- Difficulty in mapping legal and regulatory requirements to technological solutions.

What needs to be done for building resilience for future pandemics?

- Start building the privacy-preserving and secure pandemic data collection, sharing and analysis infrastructure
- Develop pilot testbeds for evaluation of security and privacy tools for pandemic data in the past and the future
- Collection and appropriate centralization of pandemic data (an effort similar to the UCI ML repository)

What are the new emerging CISE topics/challenges due to the broader economic and social disruption of pandemic and the adaptations?

- The impact of security and privacy tools on different socioeconomic groups.
- Targeted attacks that negatively impact certain subpopulations to a greater extent.
- Variability in the rate of adoption of tools due to environmental and cultural factors

How does computing research complement clinical research?

- Privacy and Security tools are a critical enabler to allow data collection, data sharing, data analysis, data management, and data protection
- Privacy-preserving analytics can enable large scale collaborative research
- Maintain data integrity, provenance, authenticity, and reliability of data

What are some coordination activities that can help?

- Having closer coordination across agencies
- Better understanding of data utility, privacy and security needs of the stakeholders (e.g., medical researchers)

What kind of support might be needed from NSF and other agencies?

- Create a cross-cutting consortium that coordinates the needs across agencies (NIH, NSF, etc.) to understand the needs from the perspective of research, and find both collaborative teams who can possibly collaborate together on the pandemic specific solutions and find ways to provide additional funding to the existing project to extend to solving the problem. An example was the KDD consortium
- Follow-up funding for promising RAPID projects.
- Building larger teams that can leverage work across projects.

Participants: Archanaa Krishnan (VT), James Joshi (NSF), Jaideep Vaidya (Rutgers), John Heidemann (USC-ISI), Li Xiong (Emory), HB Acharya (RIT), Shantanu Sharma (UCI), Murat Kantarcioglu (U Texas Dallas), Sharad Mehrotra (UCI), Sharad Sharma (Bowie St), Aryya Gangopasdhyay (UMBC), Wendy Guan (Harvard), Xiao Song (USC), Hafiz Asif (Rutgers), Yaling Yang (VT)

II-B Socioeconomic and Psychological Impact

Existing technical topics/challenges

- Using computational models for socioeconomic impact analysis
- Brain research as it pertains to mental health
- Sentiment analysis on social and news media
- Domestic violence and crime
- HCI technologies for improving workplace/school/social interactions
- Agent-based modeling and uncertainty and accuracy (understanding behavior, mobility data, social distance measure, etc.)

Emerging topics due to the disruption

- Gaming platforms role in mental health support
- Fundamental research in the area of fairness (resource access, high-speed internet, accessibility, vaccination)
- Community engagement (vulnerable communities, influencing decision making)
- Science Communication to nudge behavioral changes
- Understanding and integrating cultural/political dependencies (e.g., South East Asian, Nordic countries, North America)
- Healthcare workers fatigue; General COVID fatigue resulting in lower compliance

Data and Computing challenges

- Attitude, social media data, covid-arc (https://covid-arc.loni.usc.edu/), brain data
- Crime data, Economic impact across sectors (e.g., small businesses) https://github.com/stccenter/COVID-19-Data/tree/master/Socioeconomic%20Data
- Incentives/Infrastructure for data sharing; privacy implications;
- Covid-19 impact survey: https://www.nlm.nih.gov/dr2/Coronavirus_Impact_Scale.pdf
- https://covidstates.org/ 50 state COVID surveys
- Any focused surveys on different communities? (schoolchildren, workplace, essential workers, researchers)
- Integrating these datasets to get a more complete picture in a spatiotemporal framework

Researcher and Stakeholder Outreach

- Domestic violence (Reduction in recording and not in incidence)
- Identifying stakeholders, especially the vulnerable communities
- Misinformation, science communication, information usage, which channel or techniques can be used or are used.

Support needed from NSF and other agencies

- Understand the disparities of virtual working, stay at home, and how that impact the communities, psychologically, social interactions, so to inform/leverage future technology development
- Better integration with 'Future of Work' like initiatives
- Data and publication archive, discovery, and analyses capabilities

- Advancing science information consumption, e.g., how social media are used? How information is understood?
- Continuation and aggregation of what we learned from the RAPIDS to better prepare
- Understand the long term impact of covid to families, young generation, in a 5-10 years or even longer time frame.

Participants: Chaowei Yang (GMU), Srini Venkatramanan (UVA), Fei Wang (Weill-Cornell), Vivek Singh (Rutgers), Simon Levin (Princeton), Gloria Mark (UCI), Dominique Duncan (USC), Deep Medhi (NSF), Mitra Basu (NSF)

II-C Surveillance, Contact Tracing, and Testing

What are some of the data and computing challenges related to the pandemic?

- The interplay between privacy and utility of data (location is useful and good, but presents privacy challenges; health symptoms)
- Several data quality issues (noise, sampling, reporting, missing)
    - Direct vs indirect transmissions (can be detected with location information but not with exposure-notification apps based on bluetooth proximity)
    - How do we know if people wore masks? (human contact tracers needed in the loop)
    - Behavioral aspects of reporting symptomatic vs asymptomatic
- How to improve data collection?
    - Lack of trust is a major challenge (what type of messaging works?)
    - Lack of adequate testing infrastructure (i.e. too slow). There is opportunity to reduce the latency of testing infrastructure) in different places (i.e. airport, drive thoughts, hospitals, different sites).
    - Another aspect of testing is how to quickly capture anomalous symptoms (especially when the viruses are constantly evolving)?
- Modeling issues pertaining to integration of actual data on spread of virus with physics-based models of aerosol spread to understand hot spots

What CISE techniques have been used/deployed, and worked or helped the pandemic?

- There is technology already in place (Google Maps / Apple Notification API) using BLE; 20 states have adopted this; Privacy-preserving contact tracing and exposure notifications
- Spatiotemporal spread and risk analysis (Several dashboards)
- Group / pool testing has worked (implemented in many countries such as India and Israel)
- Optimized bioinformatics tools for analyzing genome sequence data

What can be used/deployed to help? What are the challenges for using them?

- Develop new sequencing technologies (hand held portables) that would do on-board processing of sequencing that could be used anywhere (like diabetes machines)
- Monitoring sewer systems to find strains of new viruses (has worked very well in spain)
- Social distancing which reduces the possibility of contact
- Collect advance information about mobility patterns (e.g., type of mobility, number of co-locations) in an aggregate fashion for preparedness for the next potential pandemic to have a better understanding of economic impacts for closing certain businesses in certain neighborhoods
- New location and co-location privacy research to enable better data collection

What needs to be done for building resilience for future pandemics?

- Better surveillance infrastructure (e.g., continuous privacy-aware occupancy monitoring of rooms/floors/hallways in buildings)
- New surveillance apps (monitoring parking lots, restaurant reservations)
- Technology and policy for efficient ways to do contact tracing & quarantine infrastructure
- How to scale testing infrastructure quickly?
- Unlike the current pandemic, where we have been very reactive to the events in terms of developing technology, can we be more proactive for surveillance and contact tracing i.e. build an app for future disasters
- Rapid hotspot detection

What are the new emerging CISE topics/challenges due to the broader economic and social disruption of pandemic and the adaptations?

- Bioinformatics tools aided by new algorithms, software and hardware systems.
- Disaster management and preparedness specifically for pandemic and contagious disease
- Pandemic data collection and infrastructure for quick sharing of data

How does computing research complement clinical research?

- Data infrastructure to facilitate clinical research
- Clinical data (anonymized) to be available for integration into ML spread models
- Use wearable data to inform clinical decision/diagnosis
- How can we reduce the number of people who have to go to the clinic? (Develop devices that can be shipped to people to collect physiological data so they do not have to visit clinics)

What are some coordination activities that can help (e.g., connecting with stakeholders, researchers in other areas)?

- Develop applications that will influence behavior of individuals to minimize physical distancing (i.e. crowdsourcing shopping, etc.)

What kind of support might be needed from NSF and other agencies?

- Work with other government organizations (i.e. FDA, NIH/CDC) to improve scalability of testing infrastructure given the regulatory and compliance challenges
- NSF can create a marketplace of ongoing research funded by NSF for other agencies to co-fund
- Support for translating CISE research into practice. Unlike typical computer science ventures, translating technology focused on medical science is more challenging, due to time and cost of clearing regulatory hurdles.

Participants: Cyrus Shahabi (USC), Raju Gottumukkala (U LA Lafayette), Ashutosh Dhekne (GT), Giri Narasimhan (FIU), Alon Efrat (U Arizona), Alice Koniges (U Hawaii), Anil

Vullikanti (UVA), Guohua Cao (VT), Krishna Narayanan (Texas A&M), Satish Narayanasamy (U Michigan), Sanjiv Kapoor (IIT), Alon Efrat (U Arizona), Michael Chertkov (U Arizona)

II-D Infodemiology and Social Networks

Challenges, research problems, outcome

1. What are some of the data and computing challenges related to the pandemic?

The group discussed dealing with misinformation such as hoax, rumor etc and social media as a potential platform to collect such information. Social interactions should incorporate social, political and cultural factors. However, challenges associated with the highly dynamic nature of such social media networks were also discussed to ensure meaningful insights. In addition to social media, publicly available discussion forums such as UberPeople can also be considered. Geotagged data on Twitter is sparse with representation bias, hard to make inferences about attitudes in different areas, perhaps partner with Twitter to expand meta data with IP addresses; hard to reason about targeted messaging for non-pharmaceutical interventions. Annotation data takes time to gather which hinders the development of ML/AI techniques.

2. What CISE techniques have been used/deployed, and worked or helped the pandemic?

Google mobility data, social media interactions from Facebook and Twitter. Zoom has been extensively used for collaborations and meetings. Several efforts came up with creating COVID information commons. NLP/ML techniques have been primarily used in addition to advanced visualization tools.

3. What can be used/deployed to help? What are the challenges for using them?

More support needed from sponsoring agencies, especially to promote and sustain cross-disciplinary collaborations. While agencies encourage forming such teams, however funding is limited and making longer lasting impacts through such collaborations requires funding continuation.

6. How does computing research complement clinical research?

Use of agent-based modeling techniques, deep learning approaches for forecasting and identifying solutions, mining large-scale heterogeneous data etc. helped to complement clinical research. However, more emphasis should be provided to make such solutions more human-centric.

Emerging topics

4. What needs to be done for building resilience for future pandemics?

Two directions for the potential strategies for building resilience for future pandemics: (1) We need to consolidate the knowledge and research outcomes into "lessons learned" via means such as workshops, forums, and joint reporting. And we need to communicate these lessons to different types of audiences by using languages/presentations suitable for researchers, practitioners, and the general public. (2) We need to evaluate and better understand what different types of non-pharmaceutical interventions (NPIs) we can do to increase the resilience of different communities, e.g., considering the social-psychological aspect of risk-taking, challenges of behavioral adaptation by different socio-demographic groups, etc.

5. What are the new emerging CISE topics/challenges due to the broader economic and social disruption of pandemic and the adaptations?

Three emerging challenges were highlighted in the discussion: (1) The wide-range of impact: The pandemic has created differential impacts on different socio-technical systems. For example, e-commerce, online shopping, online education, and the sharing economy have been differentially affected by the pandemic. This offers an opportunity to understand these systems and how they impact different household conditions. (2) The heterogeneous societal conditions: People react to this pandemic differently. The studies that investigate how people and government take and react to the NPIs may generate insights for making future policy and effective public messaging. (3) The dynamic nature: The situations have evolved and there is a need to keep tracking and understanding public perceptions over different phases and events: e.g., vaccination side effects, acceptance, etc.

Actionable recommendations/next steps
7. What are some coordination activities that can help (e.g., connecting with stakeholders, researchers in other areas)?
Three important coordination activities could be potentially helpful: (1) Expediting the collaborations between industry and academia: current state requires data-sharing agreements that can take a long time to establish if at all allowed. (2) Building connections or research sharing networks with international communities. (3) Lowering the barrier of collaboration across different disciplines by providing an appropriate incentive structure (see the next summary).

8. What kind of support might be needed from NSF and other agencies?
Two kinds of support can be particularly helpful. (1) The ongoing RAPID finding has made lots of active research possible. It would be crucial to continue the support for the active research development (e.g., addressing misinformation, understanding shared economy, etc.) with a proper evaluating mechanism. (2) To lower the barrier of cross-disciplinary collaboration, it would be needed to introduce new structural support (e.g., more distributive incentives as opposed to CONVERGENCE) to support cross-disciplinary research and encourage the research collaboration between industry and academic.

Participants: Arif Sadri (FIU), SS Ravi (UVA), Wei Wang (UCLA), Chris Barrett (UVA), Abhijit Suprem (GT), Raphael Stern (U Minnesota), Chen Li (UCI), Archna Bhatia (IHMC), Yu-Ru Lin (U Pittsburgh), Elena Zheleva (U Illinois Chicago), Calton Pu (GT)

Future Topical Workshops

Based on the summary reports submitted from each breakout session, the PREPARE team believes that future virtual workshops should focus on the areas listed below. Please note there are a number of subtopics that are common to each of the broader topics; as one small example, issues relating to education might be considered within both the context of the social aspects of a pandemic as well as the use of computing technology in lessening potential negative impacts of inequity in delivery mechanisms. There are multiple areas within these topics that can and should overlap for thorough examination.

- Social, behavioral, economic, and political aspects of pandemics. Subtopics may include:
    - vaccine hesitancy, challenges wrought by social media, role of interdisciplinary research, role of technology to sustain workforce and education productivity
- Role of computing technology in anticipating, controlling, responding, and assessing pandemics. Subtopics may include:
    - contact tracing, surveillance, creating contingency infrastructure to deal with unexpected outbreaks, security, testing, network issues with increased work-from-home and online education, infodemiology
- Access, creation, and maintenance of data and computing resources. Subtopics may include:
    - rapid access to resources, privacy considerations, barriers to access

The PREPARE team expects these future workshops to have a similar structure as the inaugural workshop, and will consider adding panels and further discussion opportunities as necessary.

Useful Administrative Lessons

Each time one organizes an event of this magnitude, there are areas for improvement and lessons learned that can be applied to the next event.

Pre-event

Qualtrics/registration

- Ask for email address
- Include co-PIs Add skip logic to avoid confusion among non-presenting attendees
- Uploading files to Qualtrics worked just fine
- Ask for permission to distribute emails and have posters/videos public
- If possible, use an "all in one" platform to avoid multiple registration and log-in steps (Whova if they add poster session functionality?)

Zoom registration

- Worked OK to have registration link at end of Qualtrics

Communication

- Send more frequent/thorough emails
- Meet with moderators in advance to go over expectations for breakout sessions Add all links to prepare-vo.org website - page for the event

During event

General sessions/speaker sessions

- Have closed captioning (CC) capability turned on

Poster Session

- If using platform like gather.town, have a live demo before poster session
- Whatever platforms we use, make sure we check restrictions (like Gather.town doesn't work on Safari; need image files (PNG or JPEG) for poster uploads) and let people know
- Need something to entertain people in between sessions (there were a few times with awkward gaps - maybe trivia questions in the poll?)
- Set aside place for "helpdesk" especially if virtual platform

Post-event

- If we have a list of Kaltura posters somewhere, indicate which ones have mp4s or other attachments

Workshop Agenda

Day 1 // December 15

12:00pm Welcome and Overview of NSF and CISE's Response to COVID-19 and its Impacts on the Research Community - Dr. Margaret Martonosi, NSF Assistant Director for CISE

Dr. Madhav Marathe, PREPARE PI, Biocomplexity Institute, UVA

1:00pm Poster Session I

2:00pm Poster Session II

3:00pm Poster Session III

4:00pm Poster Session IV

Day 2 // December 16

12:00pm Opening Remarks - Dr. Gurdip Singh, NSF Division Director of CNS

12:15pm Keynote Speaker - Professor Sir Roy Anderson FRS FMedSci "Where do people acquire SARS-CoV-2 infection and the challenges in creating herd immunity by mass vaccination"

1:30pm Breakout Session I

[I-A Computational Biology & Biomedical Informatics; I-B Epidemiology & Public Health; I-C Education, Training, & Workforce Development; I-D Computing & Data Infrastructure]

2:45pm Breakout Session II

[II-A Privacy & Security; II-B Socioeconomic & Psychological Impact; II-C Surveillance, Contact Tracing, & Testing; II-D Infodemiology & Social Networks]

4:00pm Breakout Summary Session

5:00pm Closing Remarks - Dr. Erwin Gianchandani, Senior Advisor to NSF Director & NSF Deputy Assistant Director for CISE

Complete Poster Listing

	Lead author	Institution	Poster Title
101	Acharya, H.B.	Rochester Institute of Technology	Censors doing little against COVID misinformation.
102	Cao, Yifeng	Georgia Institute of Technology	6Fit-A-Part: A Device for Physical Distancing
103	Hopfer, Suellen	University of California, Irvine	RAPID: Leveraging Twitter Data for Real-time Public Health Responses to Coronavirus: Identifying Affective Desensitization, Loneliness and Depression
104	Kapoor, Sanjiv	Illinois Institute of Technology	Modeling the resurgence of Covid-19
105	Mehrotra, Sharad	University of California, Irvine	T-Cove: A TIPPERS enabled Organizational Tool for COVID-19
106	Narasimhan, Giri	Florida International University	RAPID: Bioinformatic Search for Epitope-based Molecular Mimicry in the SARS-CoV-2 Virus
107	Pirolli, Peter	IHMC	Improving Computational Epidemiology with Higher Fidelity Models of Human Behavior
108	Sadri, Arif	Florida International University	#COVID-19: Understanding Community Response in the Emergence and Spread of Novel Coronavirus through Health Risk Communications in Socio-Technical Systems
109	Sharma, Sharad	Bowie State University	Visualization, Analysis and Prediction of COVID-19
110	Shyu, Chi-Ren	University of Missouri-Columbia	Geospatially-Enabled Deep Analytics for Real-time Mitigation and Response to COVID-19 Outbreak for American Rural Populations
111	Srinivasan, Ashok	University of West Florida	Leveraging New Data Sources to Analyze the Risk of COVID-19 in Crowded Locations
112	Vaidya, Jaideep	Rutgers University	Privacy-Preserving Crowdsensing of COVID-19
113	Williams, Dmitri	University of Southern California	Managing pandemic well-being through gaming
201	Belding, Elizabeth	UC Santa Barbara	Characterizing Inequities in Internet Access and Quality in California
202	Cheng, Xiang	Virginia Tech	KHOVID: Interoperable Privacy Preserving Digital Contact Tracing
203	Gangopadhyay, Aryya	UMBC	Deep Learning Models for Early Screening of COVID-19 Using CT Images
204	Furht, Borko	Florida Atlantic University	NSF RAPID: Modeling and Tracking Covid-19 Using Big Data Analytics
205	Stern, Raphael (w/Philip Pare)	University of Minnesota	Capturing the Effects of Transportation on the Spread of COVID-19 with a Networked SEIR Model
206	Wang, Wei	UCLA	Dynamic Graph Neural Networks for Modeling and Monitoring COVID-19 Pandemic
207	Kantarcioglu, Murat	UT Dallas	Privacy Risk Assessment Framework for Person-Level Data Sharing During Pandemics
208	Resnik, Philip	University of Maryland	Developing a Curated Topic Model for COVID-19 Medical Research Literature

	Lead author	Institution	Poster Title
209	Nguyen, Kien	University of Southern California	PREP: Pandemic Risk Evaluation Platform Beyond Contact Tracing for COVID-19
210	Cai, Han	MIT	TinyTL: Reduce Memory, Not Parameters for Efficient On-Device Learning
211	Thomas, Sylvia	University of South Florida	RAPID: Early Detection of Disease Outbreaks using Self-Organizing Patterns - COVID-19
212	Marathe, Madhav	Biocomplexity Institute, UVA	COVID-19 Response Support: Building Synthetic Multi-Scale Networks
213	Chen, Jiangzhuo	Biocomplexity Institute, UVA	Transfer Learning Techniques for Better Response to COVID-19 in the US
301	Chyba, Monique	University of Hawaii	RAPID: Modeling COVID-19 transmission and mitigation using smaller contained population
302	Clark, David	MIT	The Internet during the COVID Pandemic: NSF-funded RAPID studies of performance and use
303	Johnsen, Kyle	University of Georgia	New Models for Online Social Experiences Supporting Immediate Needs of NSF REU students, Conferences and the Advancement of Remote Research
304	Ju, Wendy	Cornell Tech	Profiling Public Mobility and Social Distancing Behaviors in COVID-19 New York City
305	Ray, Indrakshi	Colorado State University	Ensuring Integrity of COVID-19 Data and News Across Regions
306	Wang, Fei	Cornell University	Understanding the Transmission and Prevention of COVID-19 with Biomedical Knowledge Engineering 1
307	Squicciarini, Anna	Penn State	Social Un-distancing: Understanding Self-Privacy Violations in Online Communities during the COVID-19 Pandemic
308	Narayanan, Krishna	Texas A&M University	Make Every Test Count - How Group Testing can Increase COVID Testing Throughput
309	Cui, Jiaming	Georgia Institute of Technology	New Data Driven Approaches for Adaptive Surveillance
310	Song, Xiao	University of Southern California	Measuring the Internet during Covid-19 to Evaluate Work-from-Home
311	Lopes, Cristina	University of California, Irvine	Virtual Conferences for the Pandemic and Climate Change
312	Yang, Chaowei	George Mason University	Spatiotemporal Analytics and Simulation of COVID-19
313	Dasarathy, Gautam	Arizona State University	Learning and Leveraging Graphs for Understanding Policy, Improved Predictions, and Combatting Misinformation
314	Atashzar, S. Farokh	New York University	RAPID: SCH: Smart Wearable COVID19 BioTracker: Remote Assessment and Monitoring of Symptoms for Early Diagnosis, Continual Monitoring, and Prediction of Adverse Event
401	Chertkov, Michael	University of Arizona	Graphical Models of Pandemic

	Lead author	Institution	Poster Title
402	DeChoudhury, Munmun	Georgia Tech	Does Sharing of Misinformation on Social Media Worsen Anxiety? A Case-Study in the Context of the COVID-19 Pandemic
403	Duncan, Dominique	University of Southern California	COVID-19 Data Archive (COVID-ARC)
404	Guan, Wendy	Harvard University	Building a Spatiotemporal Platform for Rapid Response to COVID-19
405	Kasarapu, Sreenitha	George Mason University	Developing Pandemics and Healing Models for COVID-19 to Assist in Policy Making
406	Lin, Ji	MIT	MCUNet: Tiny Deep Learning on IoT Devices
407	Oladunni, Timothy	University of the District of Columbia	The Trajectory of COVID-19 in African American Community
408	Singh, Vivek	Rutgers University	Countering Language Biases in COVID-19 Search Auto-Completes
409	Vajdi, Aram	Kansas State University	Mitigation of COVID-19 transmission in college towns of Kansas
410	Xiong, Li	Emory University	REACT: Real-time Contact Tracing and Risk Monitoring via Privacy-enhanced Mobile Tracking
411	Xu, Lanyu	Wayne State University	CORPUS: An Edge Intelligence-Assisted Multi-Granularity COVID-19 Risk Prediction and Update System
412	Nayak, Kartik	Duke University	Poirot: Private Contact Summary Aggregation
413	Zheleva, Elena	University of Illinois at Chicago	Stay-at-home attitudes and their impact on the COVID-19 pandemic
414	Pu, Calton	Georgia Institute of Technology	Superspreader Event Detection with Knowledge Acquisition
415	Feng, Wuchun	Virginia Tech	Higher Accuracy and Availability of COVID-19 Testing and Monitoring via Post-CT Image Boosting and Analysis
416	Cao, Guohua	Virginia Tech	A Computational Deep-Learning Approach for Fast, Accurate CT Diagnosis of COVID-19