Privacy and Ethics Recommendations for Computing Applications Developed to Mitigate COVID-19

White Paper Series on Pandemic Response and Preparedness, No. 1

NATIONAL
SECURITY
COMMISSION
ON ARTIFICIAL
INTELLIGENCE

Letter from the Executive Director

The COVID-19 pandemic confirms that health security is national security. No facet of American life has been spared. The nation's economic prosperity, the well-being of its citizens, the readiness of its armed forces, and its role in the world are all threatened by the virus. Americans have long understood that confronting biothreats--including pandemics--are important to national security, and today we can see that the pandemic response is a central pillar to that. A successful national security strategy must include the ways and means for detecting and containing biothreats at their source, supporting biomedical innovation, and improving emergency response at home and abroad.[1]

A successful strategy for confronting the pandemic depends in large part on harnessing and applying new technologies--many of which are underpinned by artificial intelligence. The mandate of the National Security Commission on Artificial Intelligence (NSCAI) is to consider how the United States should advance the development of artificial intelligence, machine learning, and associated technologies to comprehensively address its national security and defense needs. Harnessing AI and associated technologies now to confront the pandemic will help to ensure our global leadership going forward.

Understanding the virus, tracking its spread, and ultimately ending the pandemic will require data-driven solutions. Tested and novel AI technologies will help move us toward solutions more rapidly than was possible in previous pandemics. Several AI-related partnerships to combat the pandemic have already emerged between the U.S. government and commercial or academic organizations. Other research consortia are also focused on accelerating AI applications to mitigate the impacts of the pandemic. Such innovative and decentralized activities, if expanded and melded with additional investments and actions by the federal government, could produce a strategic impact that leaves the United States safer, healthier, and more resilient when a future pandemic strikes. The technological promise is real, even if it is not a panacea.

Given the magnitude of the crisis, the national security implications, and the promise offered by AI and related technologies to fight this virus, I believe the NSCAI has a responsibility to offer recommendations that can accelerate the response to COVID-19,

[1] The 2017 National Security Strategy identifies combating biothreats and pandemics as a central component of homeland defense and identifies detecting and containing biothreats at their source, supporting biomedical innovations, and improving emergency response as priority actions. See *2017 National Security Strategy of the United States*, The White House (Dec. 2017), https://www.whitehouse.gov/wp-content/uploads/2017/12/NSS-Final-12-18-2017-0905.pdf.

prepare the United States to thwart a future pandemic, and leave the nation in a stronger position after the crisis.

We have initiated temporary special projects to issue white papers that address AI-related aspects of pandemic response and implications of the crisis for America's security and strategic competitiveness. Each paper is a collaboration of participating Commissioners and select staff, and only reflects the views of the Commissioners and staff who have contributed to the special project. It does not reflect official action, deliberation, or decision by the NSCAI, all 15 Commissioners, a majority of Commissioners, or any Commissioner who was not part of the special project that produced the paper. These special projects are separate and distinct from the Commission's current lines of effort.

The first white paper, "Privacy and Ethics Recommendations for Computing Applications Developed to Mitigate COVID-19," offers recommendations to put civil liberties at the center of contact tracing methods, and to ensure that federally funded AI tools used in pandemic response account for potential bias and avoid introducing additional unfairness into healthcare delivery and outcomes.

As always, we welcome your comments and feedback submitted via congress@nscai.gov.

Ylli Bajraktari
Executive Director

Privacy and Ethics Recommendations for Computing Applications Developed to Mitigate COVID-19[2]

NSCAI Commissioners Dr. Eric Horvitz, Hon. Mignon Clyburn, Dr. José-Marie Griffiths, and Dr. Jason Matheny[3]

Great strides over the last twenty years in leveraging data and computing in science, technology, and policy fuel excitement about harnessing data and computing technologies to mitigate the COVID-19 pandemic. Proposed directions include uses of data and computing for modeling, inference, and decision making. Applications include high-performance computing for simulations and other analyses, in support of the design of therapeutics and vaccines, and computational modeling for tracking contagious diseases, monitoring the spread among individuals, predicting future outbreaks, and allocating healthcare resources.

Multiple opportunities to address pandemic challenges with data-centric computing methods include the use of machine learning. Pattern recognition, classification, and recommendations generated via machine learning can be employed in numerous ways. Yet in determining how data-centric computing technologies are developed and employed, who controls the applications and underlying data, and how the data is used, we must ensure that American values are preserved. Missteps could undermine core civil liberties, put inappropriate information and power in the hands of government or private corporations, and deepen inequalities in our healthcare and society. Beyond issues with privacy and security, there are reasonable concerns that potentially helpful

[2] This white paper reflects the views of the commissioners and staff who have participated in this special project on privacy and ethics recommendations for computing applications developed to mitigate COVID-19. It does not reflect official action, deliberation, or decision by the NSCAI, all 15 commissioners, a majority of commissioners, or any commissioner who was not part of this special project.

[3] Commission staff who contributed to this special project white paper include Jessica Young, Rama G. Elluru, Chuck Howell, and Caroline Danauy.

technologies will not be applied in an inclusive and fair manner. For example, one or more promising solutions may rely on the ownership of and familiarity with mobile devices or, more essentially, on having access to broadband connectivity. Mobile devices, sophistication about device usage, and connectivity may not be readily available within populations in multiple regions of the country and with specific sets of socioeconomic attributes. On issues around fairness, systems and their models may rely on predictive or inferential models that rely on sampled data that is not representative of usage scenarios, leading to biases and unfair allocations of resources. The potential costs associated with losses of privacy, disparity of access to technologies, and unfair resource allocation linked to uses of data-centric computing technologies must be weighed against the value and urgency in adopting promising new data-centric computational tools to fight the current pandemic.

Adopting a series of best practices and standards will limit the risks that come with reliance on new technologies for responding to COVID-19, help defeat the current virus, and help provide insights and technology policy and infrastructure needed to identify and contain the rise of infectious disease threats needed to respond to future pandemics. We offer two recommendations.

Recommendation 1: Leverage technology, policy, and law to put civil liberties considerations at the center of contact tracing methods and tools

COVID-19 presents unique challenges to traditional contact tracing methods for controlling contagious disease. The scale of the pandemic threatens to overwhelm the labor-intensive practices of interviewing virus positive persons and those they have placed at risk, while the presymptomatic and asymptomatic transmission of the disease compounds the challenge of understanding and containing the disease's spread. [4]

[4] Recent studies provide evidence that significant numbers of infected people may remain asymptomatic, or be asymptomatic for prolonged periods of time, while spreading the illness. See Luca Ferretti, et al., *Quantifying SARS-CoV-2 Transmission Suggests Epidemic Control with Digital Contact Tracing*, Science (Mar. 31, 2020), https://science.sciencemag.org/content/early/2020/04/09/science.abb6936. A recent COVID-19 study in Iceland reported that a population-wide survey based on general outreach found that 43% of people who tested positive for SARS-CoV-2 (the virus that leads to the COVID-19 disease) had no symptoms at the time of the survey and 29% of people who tested negative for SARS-CoV-2 reported having symptoms. Daniel F. Gudbjartsson, et al., *Spread of SARS-CoV-2 in the Icelandic Population*, New England Journal of Medicine (Apr. 14, 2020), https://www.nejm.org/doi/full/10.1056/NEJMoa2006100. Another study of women admitted for labor and delivery in New York City found that 13.7% of these women tested positive for SARS-CoV-2 and that 87.9% of those who tested positive had no symptoms at presentation. Desmond Sutton, et al., *Universal Screening for SARS-CoV-2 in Women Admitted for Delivery*, New England Journal of Medicine (Apr. 13, 2020), https://www.nejm.org/doi/full/10.1056/NEJMc2009316.

AI-enabled technology could revolutionize contagious disease tracking by turning a labor intensive forensic process into a more manageable challenge that could dramatically supplement traditional methods.[5] By tracking geo-location and/or co-location data, mobile contact tracing applications offer tools ranging from helping a user improve the efficiency and completeness of manual contact tracing interviews to identifying and alerting individuals who may have been exposed through proximity to an infected person.[6] Contact tracing may be helpful in reopening the economy and in preventing or managing future waves of the pandemic in advance of the development of an effective vaccine.[7] Technologies that leverage location- or proximity-centric data from mobile phones may provide value in helping to control the spread of COVID-19 when they are coupled with effective policies for testing and with periods of isolation for those who test positive or are determined to be at high risk for having been exposed to the infection. Many entities have expressed enthusiasm with moving forward with efforts to deploy mobile based contact tracing as a part of larger pandemic mitigation

[5] Contact tracing has long been a fundamental public health tool for slowing the spread of contagious diseases. Conventional contact tracing involves interviewing a person who has tested positive for a disease to get information (voluntarily, or enforced via public health policy or by law depending on region) about all the locations he or she has been and the people with whom they have had contact. This information is then used to inform other citizens who have possibly been exposed to the infectious agent based on co-location with the infected person to take appropriate action. While mobile contact tracing methods are not necessarily enabled by machine learning or other AI technologies, they do rely on data to make inferences (such as inferring location proximities). Mobile contact tracing applications could prompt individuals for voluntary testing which would in turn provide greater quantities of data that can be used for further analyses and AI-assisted decision making (such as determining when and where to ease social distancing policies, predicting future "hotspots," and allocating resources). For more information on contact tracing, see *Contact Tracing*, U.S. Centers for Disease Control and Prevention, https://www.cdc.gov/coronavirus/2019-ncov/php/open-america/contact-tracing.html.

[6] See Derek Thompson, *The Technology that Could Free America from Quarantine*, The Atlantic (Apr. 7, 2020), https://www.theatlantic.com/ideas/archive/2020/04/contact-tracing-could-free-america-from-its-quarantine-nightmare/609577/ [hereinafter Thompson, The Technology that Could Free America from Quarantine] ("If someone tests positive for COVID-19, health officials could obtain a record of that person's cell phone activity and compare it with the data emitted by other phone owners. If officials saw any GPS overlaps (e.g., data showing that I went to a McDonald's hot spot) or Bluetooth hits (e.g., data showing that I came within several feet of a new patient), they could contact me and urge me to self-isolate, or seek a test."). While some mobile tracing applications use geo-location data, the Commision does not endorse this approach as a best practice. See Best Practices Recommendation No. 8 below.

[7] See Selena Simmons-Duffin and Rob Stein, *CDC Director: 'Very Aggressive' Contact Tracing Needed for U.S. to Return to Normal*, NPR (Apr. 10, 2020), https://www.npr.org/sections/health-shots/2020/04/10/831200054/cdc-director-very-aggressive-contact-tracing-needed-for-u-s-to-return-to-normal. Contact tracing methods should be considered as synergistic and co-designed with strategies for widespread testing, including strategies involving ongoing, recurrent testing to quickly identify when individuals become COVID-19 positive, and limit transmission.

efforts. Concerns with privacy are paramount in applications that collect data or make inferences about individuals' locations, proximities, and activities. Privacy-sensitive policies and technologies, voluntary usage, and ethical practices around disclosure and consent are central considerations with the use of contact tracing technology, but there are no coherent national or international standards for contact tracing applications.[8]

Despite the promise, rapid adoption presents practical, philosophical, and legal obstacles and unknowns as contact tracing tools are employed without mandated and standardized safeguards in place.[9] We do not know how effective the technology will prove to be in real-world settings, with challenges including whether users will be motivated to opt in to voluntary contact-tracing digital platforms and the influence of

[8] For a detailed discussion of privacy-sensitive technical protocols, see Justin Chan, et al., *PACT: Privacy Sensitive Protocols and Mechanisms for Mobile Contact Tracing*, ArXiv (Apr. 17, 2020), https://arxiv.org/abs/2004.03544v3 [hereinafter Chan, PACT: Privacy Sensitive Protocols and Mechanisms for Mobile Contact Tracing]. For additional proposed solutions to protect privacy and civil liberties, see Karissa Bell, *ACLU Outlines Privacy Concerns for Contact Tracing Tech*, Engadget (April 16, 2020), https://www.engadget.com/aclu-privacy-principles-contact-tracing-220040940.html [hereinafter Bell, ACLU Outlines Privacy Concerns]; *Joint Statement on Contact Tracing*, (Apr. 19, 2020), https://drive.google.com/file/d/1OQg2dxPu-x-RZzETlpV3lFa259Nrpk1J/view) (reflecting the signatures of over 70 scientists and researchers at US institutions). See also eHealth Network, *Mobile Applications to Support Contact Tracing in the EU's Fight against COVID-19: Common EU Toolbox for Member States,* European Commission (Apr. 15, 2020), https://ec.europa.eu/health/sites/health/files/ehealth/docs/covid-19_apps_en.pdf. The European Commission set out a process towards the adoption of a toolbox of measures in response to the coronavirus pandemic. It focuses on: i) ensuring a pan-European approach for COVID-19 mobile applications which empower citizens to carry out social distancing measures and provide them with warning, prevention and contact tracing, and ii) developing a common approach for the use of anonymized and aggregated mobile data for modeling and predictions on the evolution of the virus. The strategy puts a strong focus on the need to ensure that the toolbox of actions integrates data protection and privacy-by-design principles, and it calls for "appropriate safeguards" by listing pseudonymization, aggregation, encryption, and decentralization as examples of best practice. *Id.*

[9] While industry has integrated certain safeguards into contact tracing proposals, a lack of consistency for protecting privacy and civil liberties nationwide remains as various applications are developed and adopted. See *Apple and Google Partner on COVID-19 Contact Tracing Technology*, Apple (Apr. 10, 2020), https://www.apple.com/newsroom/2020/04/apple-and-google-partner-on-covid-19-contact-tracing-technology/. At least three U.S. local government entities are in the process of adopting "location tracking app[lications] aimed at preventing new outbreaks of the novel coronavirus." Paresh Dave, *Three U.S. Local Governments to Adopt Coronavirus Contact Tracing App: MIT*, Reuters (Apr. 9, 2020), https://uk.reuters.com/article/us-health-coronavirus-app/three-us-local-governments-to-adopt-coronavirus-contact-tracing-app-mit-idUKKCN21R3PR.

rates of usage on tracing efficacy.[10] Additional unknowns are the legal and ethical implications of a private company, non-profit organization, or the government agency amassing and exploiting users' data, even if for a public good, and the ultimate vulnerabilities of such a system to adversarial attacks.[11] There is legitimate concern from civil liberties advocates that the government or corporate entity holding the database of the aggregated information might abuse the power that comes with having such information.[12] A government could use the data collected to track citizens' movements and identify people with whom they have had contact, leading to a coercive, involuntary approach to public health or the exploitation of the information for other

[10] The lack of desire to use contact tracing technology stems from distrust on two fronts. Users are wary of technology vulnerabilities, such as cybersecurity threats. Moreover, certain populations and communities also have pre-existing distrust of healthcare systems and surveillance programs more broadly. For those that are uncomfortable adopting automated contact tracing, traditional contact tracing applications can still provide benefits. For example, computing applications can provide general relevant information to users (e.g. information on testing resources) or augment manual tracing by helping people to remember their activities during interviews with public health workers.

[11] For a discussion of concerns with contact tracing technologies, see Ross Anderson, *Contact Tracing in the Real World*, Light Blue Touchpaper (Apr. 12, 2020), https://www.lightbluetouchpaper.org/2020/04/12/contact-tracing-in-the-real-world/; Ed Felton, *COVID-19, Technology, Privacy and Civil Liberties*, Princeton University Center for Information Technology Policy (Apr. 16, 2020), https://citp.princeton.edu/event/webinar-felten-covid/ [hereinafter Felton, COVID-19, Technology, Privacy and Civil Liberties]; and Bell, ACLU Outlines Privacy Concerns. Discussion on identifying vulnerabilities to adversarial attacks and best practices to mitigate them is included in Chan, PACT: Privacy Sensitive Protocols and Mechanisms for Mobile Contact Tracing. Our recommended best practices consider the integration of such methods and approaches.

[12] See, e.g., Angel Diaz, *Coronavirus, Location Tracking, and Civil Liberties*, Brennan Center for Justice (Apr. 7, 2020), https://www.brennancenter.org/our-work/analysis-opinion/coronavirus-location-tracking-and-civil-liberties. For a discussion of potential Fourth Amendment implications, see Alan Rozenshtein, *Disease Surveillance and the Fourth Amendment*, Lawfare (Apr. 7, 2020), https://www.lawfareblog.com/disease-surveillance-and-fourth-amendment.

purposes.[13] In the hands of a corporate actor, the same personal data could be exploited for narrow private gain.

Providing the technical and policy basis for public trust is important, as the effectiveness of envisioned contact tracing applications hinges on adoption.[14]

Recommended Actions:

(1) In order to ensure adequate protection of privacy and civil liberties, Congress should pass, and the President should sign legislation that mandates the consistent use of best practices and standards, identified below, for contact tracing applications deployed in the United States, whether by a government or non-governmental entity.

(2) Congress should require the Federal Trade Commission (FTC) to regulate the fielding of contact tracing applications, and enforce compliance with recommended best practices.[15] In coordination with other entities such as the National Institute of Standards and Technology and the Centers for Disease Control and Prevention (CDC), as appropriate, the FTC should approve an accepted list of applications that meet required best practices and standards. Critically, these administrators must engage diverse stakeholders from public, private, and civil society sectors, both when determining which mobile contact

[13] For example, Singapore offers a mobile contact tracing application, "which uses Bluetooth technology to keep a log of nearby devices. If somebody gets sick, that user can upload relevant data to the Ministry of Health, which notifies the owners of all the devices pinged by the infected person's phone." However, the civil liberty tradeoff with the Singapore approach is that a user has to register with his or her phone number. So, when a person is identified as "infected with the disease, the authorities can easily match the IDs with" other information and "impose restrictive measures directly on these people." See Thompson, The Technology that Could Free America from Quarantine. Similarly, experts and the WHO agree "that South Korea's extensive tracing, testing and isolation measures . . . have helped to reduce the virus's spread." However, South Korea's contact tracing applications can reveal highly sensitive information about an infected person, including "age, gender, and a detailed log of their movements down to the minute," raising concerns that an infected person could be easily identified leading some citizens to avoid getting tested. Mark Zastrow, *South Korea is Reporting Intimate Details of COVID-19 Cases: Has it Helped?*, Nature (Mar. 18, 2020), https://www.nature.com/articles/d41586-020-00740-y.

[14] Felton, COVID-19, Technology, Privacy and Civil Liberties.

[15] To make this recommendation easily and immediately implementable, Congress should pass legislation that expressly identifies the actions the FTC must oversee, as well as the actions that constitute unlawful practices requiring the FTC to institute an enforcement action.

tracing solutions should be approved and when conducting monitoring and oversight.[16]

We offer 12 best practices for contact tracing applications:

1. Make the use of mobile-based contact tracing applications **strictly voluntary.** The **government should not in any way compel or force** the adoption or use of these applications.

2. Provide **disclosure** of how the collected data will be used, how long it will be kept, to whom it will be accessible and for what purpose, and the known risks and limitations of the application in advance of accepting usage.

3. Require that users must **explicitly consent** to the use of the application and data collected by it before any data is collected or analyzed. Users must be allowed to withdraw consent at any time. Additional consent should be requested in advance of sharing information about sensitive changes in status, such as new information on testing positive for illness.

4. Collect the minimum amount of data required for the task of contact tracing and hold the data for only as long as it is needed for the task before it is deleted. Automate **deletion of collected data after it has served its purpose**,[17] and within relatively short windows of time as needed for detecting risk of infection.[18]

5. Utilize **privacy-sensitive technology, architecture, and protocols**[19] including the use of data encryption with an effective and secure encryption scheme to minimize risks to privacy.

[16] Communities and demographics that are especially integral to a multi-stakeholder process are those disproportionately falling victim to COVID-19, those that are historically less inclined to trust programs that track movement, and those that are less digitally connected.

[17] The purpose should be narrowly tailored to address the pandemic.

[18] For example, with the PACT Protocol, this is a limited two-week period given the incubation period of the SARS-CoV-2 virus. See Chan, PACT: Privacy Sensitive Protocols and Mechanisms for Mobile Contact Tracing.

[19] For examples of privacy-sensitive architecture and protocols, see Chan, PACT: Privacy Sensitive Protocols and Mechanisms for Mobile Contact Tracing.

6. When performing data analytics that is consented to by end users, include **aggregation of the data of multiple users** that adhere to strict rules of **anonymization**, such as the use of principles of *k*-anonymity where values of k are large enough to minimize threats of identification of individuals.

7. Store user data on the user's own device(s), unless express permission is granted to share. To obviate the concerns with having a central database held either by a governmental or a private entity, a **third-party free approach** is strongly preferred, as the risk of third party access is reduced if the data is not shared, but instead stored on local, user devices. [20]

8. Store encrypted location data on a user's own device for the sole purpose of use by individuals for memory jogging in support of manual contact tracing. However, for the sharing of data in larger automated solutions, **limit shared data to encrypted information about *proximities* among users,** rather than absolute locations, to minimize risks to privacy.

9. Carefully consider **challenges with inclusiveness and potential discrimination** based on systematic differences in familiarity, abilities, and disparity of access to mobile devices for contact tracing technologies, as well as differences in COVID-19 testing among populations of different demographics and socio-

[20] Under a third-party free approach, all user data is stored on the user's mobile device, unless a user voluntarily chooses to report that they have tested positive and share encrypted data. In this case, the positively tested user sends a server his or her encrypted ID that can then be used by other voluntary participants to determine if they have been in close contact with that person--without identifying any PII ("Personal Identifying Information") of the person self-identifying as positive. This approach mitigates the security and privacy risks of requiring a trusted third-party. A "typical [trusted third-party] model . . . involves a centralized registration process wherein users subscribe to a service" that "aggregates personally sensitive information." *Privacy and the Pandemic: UW and Microsoft Researchers Present a "PACT" for Using Technology to Fight the Spread of COVID-19*, Univ. of Washington (Apr. 8, 2020), https://news.cs.washington.edu/2020/04/08/privacy-and-the-pandemic-uw-researchers-present-a-pact-for-using-tech

economics, including race, ethnicity, gender, age, education, and income levels.[21] To mitigate differences in familiarity with contact tracing applications, **informational resources should be provided** both online[22] and offline.[23] In areas where mobile contact tracing will not be feasible and areas that are disparately impacted by testing access and resources,[24] **gaps must be filled with more intensive manual contact tracing efforts.**[25] The Federal government should invest in understanding and addressing disparate effectiveness of contact tracing because of differences in access to technologies.

[21] See David Waldstein, *C.D.C. Releases Early Demographic Snapshot of Worst Coronavirus Cases*, New York Times (Apr. 8, 2020), https://www.nytimes.com/2020/04/08/health/coronavirus-cdc-demographic-study-hospitalizations.html [hereinafter Waldstein, C.D.C. Releases Early Demographic Snapshot]; Robert Samuels, *Covid-19 is Ravaging Black Communities. A Milwaukee Neighborhood is Figuring Out How to Fight Back*, Washington Post (Apr. 6, 2020), https://www.washingtonpost.com/politics/covid-19-is-ravaging-black-communities-a-milwaukee-neighborhood-is-figuring-out-how-to-fight-back/2020/04/06/1ae56730-7714-11ea-ab25-4042e0259c6d_story.html [hereinafter Samuels, Covid-19 is Ravaging Black Communities]; Erin Durkin, *Hispanic and Black New Yorkers are Dying at Highest Rates from Coronavirus*, Politico (Apr. 8, 2020), https://www.politico.com/states/new-york/albany/story/2020/04/08/hispanic-and-black-new-yorkers-are-dying-at-highest-rates-from-coronavirus-1273789 [hereinafter Durkin, Hispanic and Black New Yorkers are Dying at Highest Rates]; John Blake, *Native Americans were Already Decimated by a Virus. They're Scared it Could Happen Again*, CNN (Apr. 14, 2020), https://www.cnn.com/2020/04/14/us/native-americans-coronavirus-blake/index.html [hereinafter Blake, Native Americans were Already Decimated by a Virus]; Maya King, *Black Doctors Blast 'Woefully Anemic' Data on Minority Coronavirus Cases*, Politico (Apr. 20, 2020), https://www.politico.com/news/2020/04/20/minority-cases-coronavirus-197203 [hereinafter King, Black Doctors Blast 'Woefully Anemic' Data].

[22] Users can be motivated to adopt contract tracing by building trust around them through disclosure of informational resources and education on contact tracing technology, including its value and risks. See Jelle Prins, *This is What a Contact Tracing App Could Look Like*, OneZero (Apr. 27, 2020), https://onezero.medium.com/openui-a6b9c3d741de.

[23] Those distributing offline resources about mobile contact tracing should, like manual contact tracers, be trusted partners from within the communities themselves.

[24] For example, this may be due to lack of smartphone penetration, broadband access, or phone plans with adequate data allotments.

[25] Very few states have enough manual contact tracers per capita to meet expected needs. See Selena Simmons-Duffin, *We Asked All 50 States About Their Contact Tracing Capacity. Here's What We Learned*, NPR (Apr. 28, 2020), https://www.npr.org/sections/health-shots/2020/04/28/846736937/we-asked-all-50-states-about-their-contact-tracing-capacity-heres-what-we-learne. Beyond a state's baseline needs for contact tracers, many communities with lower internet and/or tech device access will need more intense resource allocation. See *Broadband Availability, Digital Equity and COVID-19 App*, Esri, (last accessed May 4, 2020), https://broadbandusa.maps.arcgis.com/apps/webappviewer/index.html?id=39dc6d41b49d420e94aef77c441f8af2.

10. Review of proposed solutions and applications should be undertaken by **expert panels of privacy and security experts,** including panelists with expertise in cryptographic methods and adversarial attacks on security, **and representatives from civil liberties organizations.**[26]

11. **Publish proposed systems and invite red-teaming, including adversarial attacks** as part of the design process, including study of malevolent attacks that can flood systems with false information about infection and well-being.

12. Adopt proposed systems that are **co-designed in the context of larger strategies and considerations,** including disease testing, and modeled on a larger scale aimed at understanding the effectiveness of technology given real-world considerations, such as sensitivity to different levels of usage in communities and different coverage or availabilities of tests for COVID-19.[27] Proposed systems should be used as a tool to complement, not replace, human efforts such as manual contact tracing.

Recommendation 2: Ensure that federally funded computing tools created and fielded to mitigate the COVID-19 pandemic are developed with a sensitivity to and account for potential bias and, at a minimum, do not introduce additional unfairness into healthcare delivery and outcomes.

AI methods, including machine learning, are being applied to understand molecular interactions for designing therapeutics and vaccines; to perform triage at multiple phases of care, including identifying at-risk patients for guiding hospital admission and predicting risk of physiologic decline and of mortality to guide therapy; and, more generally, to conduct modeling and decision support for the allocation of scarce resources.[28]

[26] This should include representation from communities most likely to be adversely impacted by any potential shortcomings of contact tracing applications and their fielding. See also Footnote 16 for critical voices to include.

[27] See Chris Stokel-Walker, *Can Mobile Contact-Tracing Apps Help Lift Lockdown?*, BBC (Apr. 15, 2020), https://www.bbc.com/future/article/20200415-covid-19-could-bluetooth-contact-tracing-end-lockdown-early, (discussing role of contract tracing in larger strategies to address COVID-19).

[28] See e.g., Kelly A. Wittbold, et al., *How Hospitals are Using AI to Battle Covid-19*, Harvard Business Review (Apr. 3, 2020), https://hbr.org/2020/04/how-hospitals-are-using-ai-to-battle-covid-19.

While great value can be provided, statistical analyses, including uses of machine learning procedures, can introduce unintended biases. The consequence of bias is pronounced in healthcare where social determinants of health significantly impact one's access to information and technology, vulnerability to disease, and where many underlying health disparities are correlated to race and poverty.[29]

If properly designed, statistical analyses, including machine learning procedures, can assist governmental and non-governmental agencies to build insights about underserved populations and to help ensure that resources and attention are fairly distributed based on risk and need. Such tools can provide analyses and visualizations to advise Congress on how best to distribute pandemic federal relief funds. Data aggregation can support analyses of existing disparities exposed by COVID-19, yielding insights about who is most at-risk of negative health outcomes, and, therefore, should be prioritized in receiving resources/interventions.[30] An essential starting point is developing a baseline understanding of how socioeconomic factors and social determinants of health are shaping COVID-19 health outcomes, which requires important surveying. Despite inconsistent data reporting policies, growing evidence indicates that African American, Hispanic, and Native American populations in particular are disproportionately impacted by COVID-19.[31] The tools being developed to help visualize and analyze the

[29] See *Social Determinants of Health: Know What Affects Health*, CDC (Jan. 29, 2018), https://www.cdc.gov/socialdeterminants/; *Social Determinants of Health*, Office of Disease Prevention and Health Promotion, U.S. Department of Health and Human Services (last accessed May 4, 2020), https://www.healthypeople.gov/2020/topics-objectives/topic/social-determinants-of-health; Vann R. Newkirk II, *The Coronoavirus's Unique Threat to the South*, The Atlantic (Apr. 2, 2020), https://www.theatlantic.com/politics/archive/2020/04/coronavirus-unique-threat-south-young-people/609241/ [hereinafter Newkirk, The Coronoavirus's Unique Threat].

[30] Cheyenne Haslett, *CDC Releases New Data as Debate Grows Over Racial Disparities in Coronavirus Deaths*, ABC News (Apr. 8, 2020), https://abcnews.go.com/Politics/cdc-releases-data-debate-grows-racial-disparities-coronavirus/story?id=70041803.

[31] See Waldstein, C.D.C. Releases Early Demographic Snapshot; Samuels, Covid-19 is Ravaging Black Communities; Durkin, Hispanic and Black New Yorkers are Dying at Highest Rates; Blake, Native Americans were Already Decimated by a Virus; King, Black Doctors Blast 'Woefully Anemic' Data.

interaction of factors that shape COVID-19 vulnerability and outcomes will require systematic data collection and reporting along factors such as race, gender, and region.[32]

Failing to consider these factors in the testing and evaluation of computing solutions developed to fight the pandemic, foresee resource allocation needs, and prioritize treatment could result in unwanted biases and blindspots. Such failures can skew predictions, diagnoses, risk scores, and decisions about where, or to whom, finite resources and care should be prioritized. The outcomes could exacerbate existing disparities, which would run contrary to our nation's values of equity and fairness, undermine confidence in our public institutions, violate the law, and delay, suspend, or halt the use of otherwise valuable computing tools.

Recommended Actions:

(1) Collect and analyze data[33] with a goal of identifying and mitigating disparate impacts during the pandemic. The CDC, Centers for Medicare and Medicaid Services, and other relevant federal organizations should collect and analyze data about COVID-19 interventions and outcomes[34] with segmentation by population factors such as race, ethnicity, gender, employment status, and census tract.[35] The federal government should also encourage, and where legally permissible, require healthcare systems across the country to collect and publish this kind of data.

[32] Le Wen Chiu, *Population Health and Health Equity is Excited to Announce the Launch of a New Website*, Univ. of Cal. San Francisco (Apr. 3, 2020), https://pophealth.ucsf.edu/news/population-health-and-health-equity-excited-announce-launch-new-website; Newkirk, The Coronoavirus's Unique Threat; Kate Kelland, *Explainer: Do Men Fare Worse with COVID-19?*, Reuters (Apr. 8, 2020), https://www.reuters.com/article/us-health-coronavirus-men-explainer-idUSKBN21Q1IE; Eric Levenson, *Why Black Americans are at Higher Risk for Coronavirus*, CNN (Apr. 7, 2020), https://www.cnn.com/2020/04/07/us/coronavirus-black-americans-race/index.html; Sandra Chapman, *Indiana Urged to Release Data on Racial Disparities for COVID-19 Deaths*, WTHR (Apr. 8, 2020), https://www.wthr.com/article/indiana-urged-release-data-racial-disparities-covid-19-deaths.

[33] This should include data to answer inequality issues raised in this memo, particularly inclusiveness and potential discrimination challenges. See Best Practices Recommendation No. 9.

[34] Examples of interventions and outcomes include who receives access to testing, is admitted to hospitals, is admitted to ICUs, is provided pharmaceutical intervention with drugs in limited supply, and survival/death rates.

[35] These are examples of relevant population factors though a non-exhaustive list. Sensitive data collection and analyses should use data analytics, including aggregation of the data of multiple users that adhere to strict rules of anonymization, such as the use of principles of *k*-anonymity where values of k are large enough to minimize threats of identification of individuals.

(2) Congress should require federal entities to routinely review the collected data and outcomes, look for signals suggestive of disparate impact, and report those results to Congress.[36] Reporting on the above data and analyses will provide congressional committees overseeing the distribution of COVID-19 relief funds with visibility into factors needed to ensure federally funded interventions mitigate, rather than exacerbate, disparate impact.

(3) Congress and the relevant executive branch organizations should require that any organization developing and using federally funded computing tools with machine learning components to fight COVID-19 adopt the following practices to identify and address potential disparities in healthcare delivery and outcomes:

- When using classification and prediction technologies, challenges with representativeness of data used in analyses, and fairness of inferences and recommendations made with systems leveraging that data when applied in different populations, must be considered explicitly and documented. Techniques include asking key questions about disparate impact starting early in the development process.[37] They also include documenting deliberations, actions, and approaches used to ensure fairness and lack of unwanted bias in the machine learning application to allow for adequate oversight from Congress or other entities.

[36] One potential signal might be a significant difference in the progression of the disease between races (e.g. from the percentage that test positive that go to hospital, the percentage in hospital that go to ICU, the percentage in the ICU that are intubated, and the percentage of deaths). This signal could be used to flag further review. An example of a potential, subtle signal that machine learning tools could help detect would be if a distinctly different progression (as described above) occurred for women of the same race in some census tracts or zip codes compared to other census tracts or zip codes that appear similar. This would be a flag to investigate what was driving the divergence.

[37] Some examples of potential questions include: Did you determine what harm would be caused if the AI system makes inaccurate predictions or recommendations? Did you put in place ways to measure whether your system is making an unacceptable amount of inaccurate predictions or recommendations? Did you clearly communicate characteristics, limitations and potential shortcomings of the AI system? In case of the system's development: to whoever is deploying it into a product or service? In case of the system's deployment: to the (end-)user or consumer? Did you ensure that you have agreed to an adequate working definition of "fairness" that you apply in designing the AI system? Did you establish a quantitative analysis or metrics to measure and test the applied definition of fairness? Did you carry out a risk or impact assessment of the AI system, which takes into account different stakeholders that are (in)directly affected? *Ethics Guidelines for Trustworthy AI*, High-Level Expert Group on Artificial Intelligence, European Union at 26-31 (Apr. 8, 2019), https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai. See also Michael Madaio, et al., *Co-Designing Checklists to Understand Organizational Challenges and Opportunities Around Fairness in AI* (Apr. 25-30, 2020), http://www.jennwv.com/papers/checklists.pdf.

- There is no single definition of fairness. System developers and organizations fielding applications must work with stakeholders to define fairness, and provide transparency via disclosure of assumed definitions of fairness. Definitions or assumptions about fairness and metrics for identifying fair inferences and allocations should be explicitly documented. This should be accompanied by a discussion of alternate definitions and rationales for the current choice. These elements should be documented internally as machine-learning components and larger systems are developed, so they can be shared with oversight entities as needed.

- Given assumed goals of fairness in inferences and recommendations, tools and techniques should be developed to identify and mitigate biases in systems relying on machine learning. Examples of such tools include Aequitas by the University of Chicago,[38] Fairlearn by Microsoft,[39] and AI Fairness 360 by IBM.[40] Such tools and techniques should be used to assess system accuracy and errors relative to one or more fairness metrics; this would help to determine if decisions or recommendations are made in a fair and appropriate manner when it comes to key demographic attributes such as gender, age, and socioeconomic status.

[38] See Aequitas, *Bias and Fairness Toolkit*, http://aequitas.dssg.io/. Aequitas "can be used to audit the predictions of machine learning based risk assessment tools to understand different types of biases, and make informed decisions about developing and deploying such systems." *Aequitas*, Center for Data Science and Public Policy, The University of Chicago, http://www.datasciencepublicpolicy.org/projects/aequitas/.

[39] See Microsoft, *Fairlearn*, https://github.com/fairlearn/fairlearn. As Microsoft explains: "Fairlearn is a Python package that empowers developers of artificial intelligence (AI) systems to assess their system's fairness and mitigate any observed unfairness issues. . . . Fairness is fundamentally a sociotechnical challenge. Many aspects of fairness, such as justice and due process, are not captured by quantitative fairness metrics. Furthermore, there are many quantitative fairness metrics which cannot all be satisfied simultaneously. Our goal is to enable humans to assess different mitigation strategies and then make trade-offs appropriate to their scenario." *Id.*

[40] See IBM, *AI Fairness 360 Toolkit*, http://aif360.mybluemix.net/. The toolkit can "help you examine, report, and mitigate discrimination and bias in machine learning models throughout the AI application lifecycle. Containing over 70 fairness metrics and 10 state-of-the-art bias mitigation algorithms developed by the research community, it is designed to translate algorithmic research from the lab into the actual practice of domains as wide-ranging as finance, human capital management, healthcare, and education." *Id.*

NATIONAL
SECURITY
COMMISSION
ON ARTIFICIAL
INTELLIGENCE