BACKGROUND – FOR RESPONSE A3 – DATA MODERNIZATION

The COVID-19 pandemic[[1]](#footnote-2) revealed important limitations to public health infrastructure and exposed significant challenges and equity concerns. One of the most significant lessons learned is that health information’s ability to help address the pandemic is dependent on a workforce trained and available to use a coherent and accessible data infrastructure. Despite previous investments in health information technology, various critical staff and health care data sources were simply not available.

In the early days of the COVID-19 pandemic, we sought to assess, surveil, analyze, and communicate public health information. We needed answers to questions such as:

* Who are the most vulnerable people, and where is the infection spreading?
* How many COVID-19 patients does each health system serve, and what is the system’s capacity for treating them?
* Does each health system have an adequate supply of personal protective equipment (PPE), intensive care units (ICUs), and ventilators—and most importantly, do they have the appropriately trained and adequately rested staff that are required to deliver care and monitor complex equipment?
* For each COVID-19 patient, what are the key data elements of treatment and outcome, and what does a population-scale analysis of these data elements tell us about best practices?

*At first glance, the task seemed simple enough*: the petabytes of digital data that the healthcare system generates daily only needed to be extracted, integrated, and disseminated in useful forms with the use of a wide range of digital tools such as telemedicine, biosensors, easy-to-use digital apps, machine learning (ML), and artificial intelligence (AI). *Instead:* we discovered that we lacked critical staff knowledgeable in data science. Information gathering was slow, sources were extremely siloed, and communication was difficult. Systems that were effective for their original, focused purpose were not flexible and integrated enough to meet this new, pandemic challenge. The healthcare interoperability that we did achieve required a significant architectural mobilization of largely ad hoc collaborations and new system deployments. Current solutions are not sustainable for foundational public health operations nor prepared to respond to the next public health emergency.

Approach

The State of XX is requesting funds to improve our capacities to support basic public health protections, programs, and activities that are key to ensuring community health, responding to emergencies, and achieving equitable outcomes. We are focusing on the integration, connectivity, and analytical capabilities of existing systems’ data and partnerships.

Previous investments in our public health system were not flexible enough to meet all our needs in the COVID-19 pandemic. In this request, we have identified areas of data modernization which we need to achieve the activities described in the NOFO and focused on flexibility to be able to adjust to address future needs.

An underlying theme in our data modernization efforts is using technology that allows the maximum flexibility for staff enablement, including functionality for our data platform. By choosing a platform which provides a single environment for many different analytic tools and cloud platforms, we will widen our applicant pool; applicants like college students who have been trained on open source can seamlessly work with significantly more experienced data scientists who are experts with more traditional analytic tools such as SAS. Further, ease and efficiency of training on a single platform with a variety of tools allows us to recruit workers from underserved groups who may not have had previous exposure to multiple systems, thus increasing the diversity of our workforce.

We are focusing on ensuring a single data model through the environment so that all our staff can be reassigned as emergency needs arise and be comfortable with the tools, environment, and data schema. This will allow us to respond much faster with a more comprehensive view to future needs.

Flexibility will be considered in many areas for our data modernization: flexibility of environment, flexibility to manage variable workloads whatever the data source, flexibility of analytic and BI tools, flexibility to reuse analytic models and data sets in ways that were not expected, and data storage flexibility. All these options are integral to our public health modernization efforts, with components which are reusable, secure, transparent, extendable, adaptable, and **flexible**.

Figure 1 describes the way in which we expect to answer critical public health questions:



Figure 1: Public Health Analytic Ecosystem with Shared Data

We believe that the most efficient way to augment our existing systems is to implement a connected data store and analytic platform which can reach out to legacy systems while also providing a space for new data sources to be analyzed in conjunction with what we already collect, as depicted in Figure 2.



Figure 2: Operationalized Analytics to Improve Public Health

Activities

We will strengthen the capabilities of our public health department to assess and surveil the health of our communities, prepare for emergencies, and coordinate with other parts of the healthcare system.

## Activity A3.1

One of the most critical parts of a data infrastructure modernization project is the people who make it happen. We will quickly identify and hire a data modernization director and identify/hire a supporting team to expand departmental policies, procedures, and practices to orient towards an agile procurement, planning and implementation approach for DMI implementation. The director will ensure that the data infrastructure is inclusive to the needs of counties and cities and is technically enforceable.

This team will also engage contractors with technology, assessment, planning, system design and implementation expertise to support activities, if needed.

## Activity 3.2

We will assess and report the current capacity, gaps, and opportunities to modernize the public health data infrastructure and workforce, framed by the requirement for sustainability in design and development. This will include completing the required assessment outlining current capacity, gaps, and opportunities using the assessment tool available for use by the CDC.

## Activity 3.3

In this activity we will create and fine tune our implementation plan. We have already identified the high-level themes of connected data, flexibility, and sustainability for our data modernization infrastructure. We have included a strawman conceptual architecture which aligns with the “North Star” vision descriptions that we have found. As we move forward in the project, we intend to identify needs, solidify our architecture, and define the technical standards for this modernization vision. We will incorporate services and resources from the CDC and other jurisdictions wherever possible. This will include working with partners and stakeholders to establish the most critical information sharing/analytic needs and develop a pathway to quickly see value and public benefits in our endeavor. We expect that this agile, fast time-to-value approach will encourage future partnerships in our community, both inside our jurisdiction, but also from non-governmental sources as well. With this in mind, our goal is to utilize an agile focused methodology to ensure our implementation plan is flexible and adaptable as we hear of lessons learned from other jurisdictions.

## Activity 3.4

Once we have firmly established the goals, use cases, and work plan during Activity 3.3, we will implement an agile work plan to enhance and improve our data infrastructure.

These things will include:

* Data Exchange and Systems Interoperability
* Data and IT Governance
* Data and Analytics, Visualization and Dissemination
* Data Standardization and Linkage
* System Security and Data Protection
* Process Automation
* Achieving Efficiencies

## Additional Content for A2.2F

In this task, we seek to strengthen our organizational competencies addressing information technology. Specifically, we need to advance our information technology services by procuring the hardware and software required to access electronic health information and to improve the department's operations and analysis of health data.

We need timely and sufficient foundational data to guide public health planning and decision making at the state and local level. We will improve our ability to collect data and access data from other parts of the state government.

To leverage new and enhanced data sources, we will increase the interoperability of our data systems—so that public health can share data with other agencies such as education or justice. We are seeking an analytics platform that will allow users within the State to retain control of their data and perform analytics in conjunction with the broader State data. The platform will enable data exploration and analytics of both curated and raw data in analytic sandboxes across the State agencies. These sandboxes will be easy to create and accessible by anyone to facilitate assessment.

Previous tempos of annual community health assessment were not sufficient to prepare our decision-makers for the challenges of COVID. We will use this investment to conduct more frequent updates to community or statewide health assessments, including analysis of root causes of health disparities and inequities.

Our current data views are neither broad nor detailed enough. With this investment, we will add capabilities to analyze, interpret, and use data from a variety of sources including granular data and data disaggregated by geography (e.g., census tract, zip code), sub-populations, race, ethnicity, and other variables that fully describe the health and well-being of a community and the factors that influence health. We will also widen our view by collaborating with other states in our region and between agencies within our State.

The improvements we make with this investment will give us the ability to access 24/7 laboratory resources capable of providing rapid detection. We will develop the ability to participate in or support surveillance systems to rapidly detect emerging health issues and threats.

Improved surveillance data will be substantially larger and faster than the data we have now. To efficiently ingest, store, and analyze this data, we will need systems that can handle complex, heavy workloads. We need to accelerate our response to future emergencies.

We are seeking investments to improve our preparedness and response strategies to address a range of events including natural or other disasters, communicable disease outbreaks, environmental emergencies, or other events. With improved data and systems, we will integrate social determinants of health, and actions to address inequities, including ensuring the protection of high-risk populations, into our plans.

To become better prepared for future emergencies, we need an infrastructure which can run on, or easily migrate to, multiple cloud and on-prem environments, ensuring that public health operations can continue to function when most needed. A resilient public health data system must provide minimal downtime, automatic recovery, and seamless upgrades. In this request, we seek to enhance our existing systems with a data analytics capability.

## Activity A2.3

To help enhance communications and to implement strong communications capabilities and products, our public health department must be better able to respond to data requests and translate data into information and reports that are valid, complete, statistically accurate, and accessible to a wide range of audiences. In this proposal we are seeking support for a system that is better integrated with federal systems, such as the CDC’s national statistical reporting systems. Tight integration with a wide range of third-party tools is critical for productivity and responding to requests from the CDC timely and efficiently.

We will obtain analytic solutions that will help us report and use public health data in ways that are relevant to communities experiencing health inequities or ability to support community-led data processes. We will ensure that our systems are not only capable of creating data models, but also can manage and deploy those models in a timely manner so that we can get data-informed decisions out to communities and to the public.

We want to establish a dialogue with our communities and stronger partnerships within our State and across our region. With this funding, we will develop new capabilities to transmit and receive routine communications to and from the public in an appropriate, timely, and accurate manner, on a 24/7 basis. We will get in front of misinformation with a proactive health education/health communication strategy.

To improve our coordination, we need to be able access data wherever it resides. The data analytics platform we are requesting will allow our analysts to collaborate outside of the agency with any number of data sources and different kinds of data—population health data, census data, weather data, other States’ data, etc. Often, it will not be practical or possible to copy external data. Having a platform that can access data whether it exists on-premises, in a cloud environment, or in multiple cloud environments can address this challenge.

1. The US Secretary of Health and Human Services declared the Coronavirus 2019 (COVID-19) to be a public health emergency (PHE) as of January 27, 2020 (Azar 2020). As of the writing of this proposal, the United States is still under a COVID-19 PHE (Becerra 2021). For the purposes of this proposal, the “COVID-19 pandemic” refers to the two-year period from January 27, 2020, to January 27, 2022. [↑](#footnote-ref-2)