DMI

SNAPSHOT

2022

TRANSFORMING DATA TO PROTECT HEALTH

CDC'S DATA MODERNIZATION INITIATIVE

DMI
@ CDC

"At CDC and throughout public health, we are in a pivotal moment for data and surveillance — one marked by opportunities, challenges, and the need for change."

— Rochelle Walensky, MD, MPH, CDC Director

CONTENTS

DMI 2022 SNAPSHOT

What does data modernization look like?

DATA MODERNIZATION LOOKS LIKE...

AN EPIDEMIOLOGIST

spending the day investigating COVID-19 cases instead of entering them into a spreadsheet

A COMMUNITY HEALTH WORKER

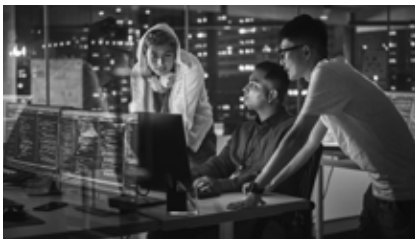
knowing precisely which neighborhood needs a pop-up vaccine clinic

A PUBLIC HEALTH ANALYST

producing a report that used to take two days in just two clicks of a button

A RESEARCH TEAM

analyzing more than 12 million health records to understand latent tuberculosis infection



A RESPONDER

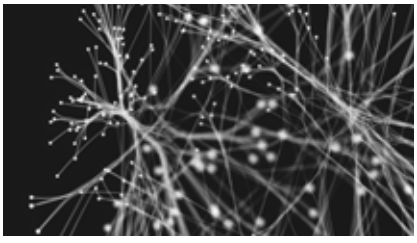
catching an uptick in overdose deaths and sending fentanyl test strips to the frontlines

A DATA SCIENTIST

visualizing social media feeds to help monitor suicide trends

A MULTI-STATE HEALTHCARE NETWORK

saving 160,000 staff hours that would have been spent on paperwork



A POLICYMAKER

using a real-time wastewater surveillance dashboard to send resources where they're needed most

Read the full stories in the 2022 DMI Snapshot

MADE POSSIBLE THROUGH TECHNOLOGIES LIKE...

Automated, **ELECTRONIC REPORTING** that replaces paper and reduces the burden on health department staff

A suite of **CLOUD-BASED** tools that saves time and costs while analyzing more kinds of data faster than ever before

A **DATA CATALOG** where CDC's experts can see which data exist in any part of the agency on any disease or condition

One **FRONT DOOR** for state and local partners to send data to all of CDC instead of sending to multiple programs in multiple ways

A **NORTH STAR ARCHITECTURE** for data that can be used by every level of public health to collect, transform, and share information

COMMON STANDARDS that help different data systems across public health and healthcare speak the same language

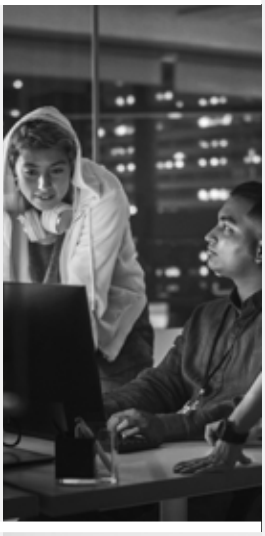
Dashboards that **VISUALIZE DATA** and make it available to the public, researchers, and policymakers in real time

Novel approaches like **ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING** that can discover relationships in the data that are hard for traditional methods to find

AND BY FINDING NEW WAYS TO...

Bring people from across public health, government, academia, and industry together to find common **SOLUTIONS** to common data problems

OPEN UP THE DATA so that more people have access to the information they need for decisions, both in emergencies and every day

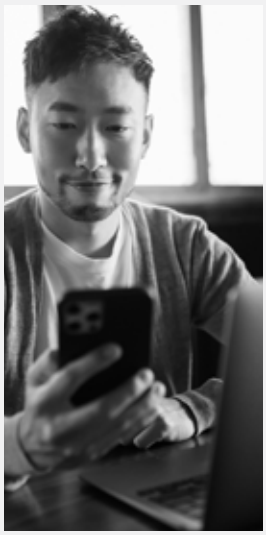


Make sure that **HEALTH EQUITY** is "baked in" from the start on every-thing we do

Offer public health staff opportunities to **LEARN BY DOING** and to access the latest data science training anytime, anywhere

Modernize the **POLICIES AND AGREEMENTS** that inform how data are exchanged so that we can get information where it needs to go to save lives

USE RESOURCES WISELY by making sure everything we invest in is connected, flexible, and sustainable for the future



Read the full stories in the 2022 DMI Snapshot

RESULTING IN...

Data increased

20,000+
healthcare facilities delivering automated, real-time case reports —up from just 187 before the pandemic

91%
complete race information in syndromic surveillance—up from 80% in 2019—which is critical important for understanding health disparities

~90%
of death records coded automatically using natural language processing—up from 75% with the previous system—with data to jurisdictions in minutes

Time saved

~80%
decrease in development time for CDC's response and surveillance programs using new Cloud components

1+
year reduction in the time it takes to understand suicide death trends thanks to new "nowcasting" capabilities

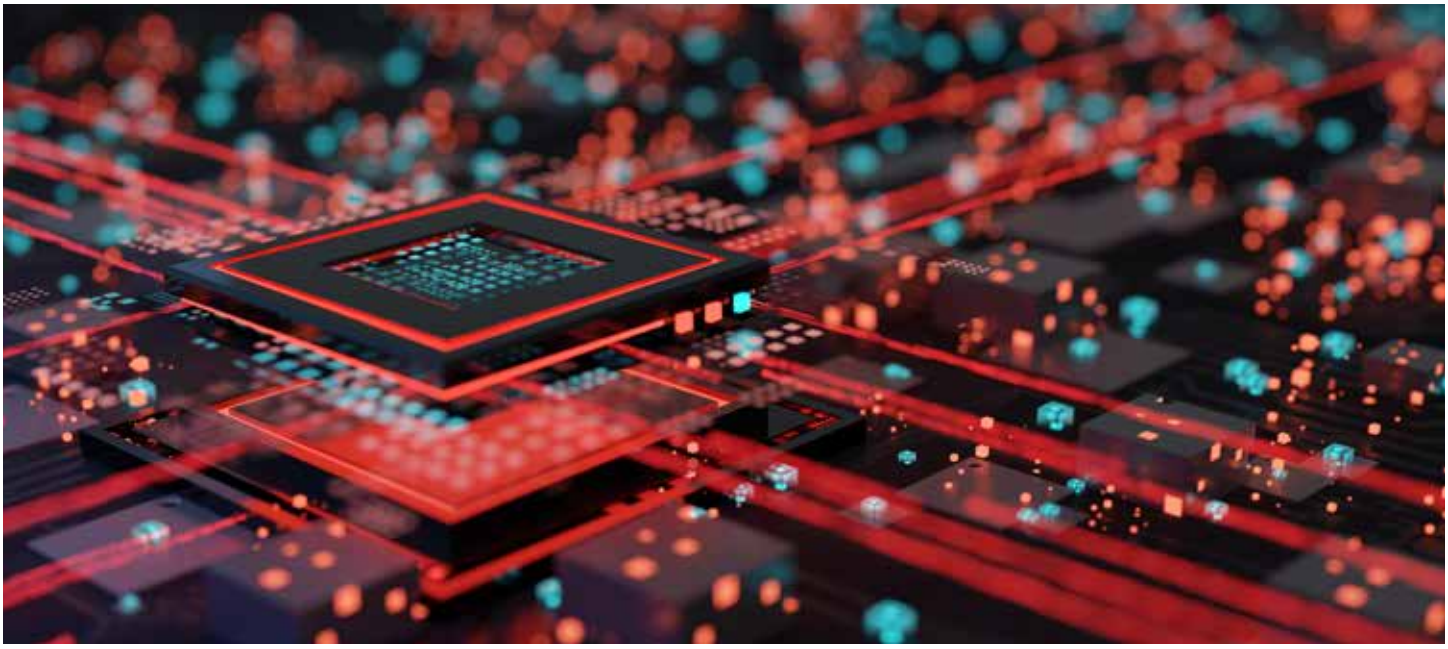
~5.5
days from sample collection to data available in CDC's wastewater surveillance system—down from ~4 weeks

People connected

3,700
training hours completed by CDC staff in the Data Academy—an 83% increase over the previous year

18
multi-sector meetings of the Consortium for Data Modernization, covering 24 high-priority topics

200+
CDC staff actively engaged in defining priorities and solving challenges through DMI Implementation Teams



Read the full stories in the 2022 DMI Snapshot

DATA

CORE DATA SOURCES

Finding and facing any threat

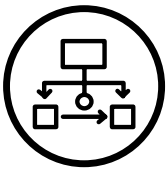
When America's health is threatened – no matter the cause – the nation's core data sources deliver the information needed for action.

WHAT DO WE MEAN BY "CORE DATA"?

Core data are data that, together, provide a more complete picture of what's happening in our communities. These data can be used to understand any disease or condition, both in emergencies and every day.

CASE DATA

representing comprehensive disease and condition information



LABORATORY DATA

including test results and test type

EMERGENCY VISIT DATA

including clinical diagnoses, signs, and symptoms

VITAL STATISTICS DATA

including birth and death data

IMMUNIZATION DATA

that capture vaccine doses administered (both routinely recommended and response-related)

HEALTHCARE CAPACITY AND UTILIZATION DATA

that assess availability of healthcare resources, including staff, beds, and equipment

MODERNIZATION IS HOW WE SOLVE BIG CHALLENGES

50M

Since the start of the COVID-19 pandemic, laboratory reporting has skyrocketed — from ~30 million reports per year up to ~50 million every month

~70%

About 70% of healthcare organizations use fax to send or receive care records [1]

120+

CDC monitors ~120 diseases and conditions that — if left unchecked — could significantly harm the health of communities

[1] ONC Data Brief no. 54 (2021)

Read the full stories in the 2022 DMI Snapshot

MODERNIZING CORE DATA SOURCES DELIVERS FASTER, BETTER DATA FOR DECISIONS.

LESS MANUAL DATA ENTRY: More than 24.5 million COVID-19 case reports have been sent electronically from healthcare to public health agencies, each representing a report that a healthcare provider does not have to enter manually.

FASTER DATA FLOWS: A new workflow was created for over 55 state, territorial, and local jurisdictions, with laboratory data flowing to CDC within four weeks.

MORE EQUITABLE DATA: Recent innovations have increased the completeness of race information in emergency visit reports to 91% in 2022, up from 80% in 2019, which has been critically important for understanding health disparities.

DATA AVAILABLE SOONER: MedCoder, a new system for coding cause of death from death certificates, codes nearly 90% of records automatically, making data available for jurisdictions in minutes instead of weeks.

ROBUST DATA COLLECTION: Because we moved quickly to reinvent immunization data systems from start to finish, CDC had received data on 663,822,575 COVID-19 vaccines administered by the end of 2022.

MORE DATA FOR THE PUBLIC: New dashboards monitor how many people are in the hospital from any cause, which helped when the nation was struck by the "triple threat" of COVID-19, influenza, and respiratory syncytial virus (RSV).

"What we've been able to prove during the COVID-19 response is all of this is possible, we can see these capabilities in a not-too-far-off future, and—importantly—we know how to get there."

— DAN JERNIGAN, MD, MPH, DIRECTOR, CDC NATIONAL CENTER FOR EMERGING AND ZOONOTIC DISEASES

SPOTLIGHT

As electronic case reporting (eCR) is more fully adopted, it becomes possible for some facilities to finally "turn off the fax machines" for reporting cases.

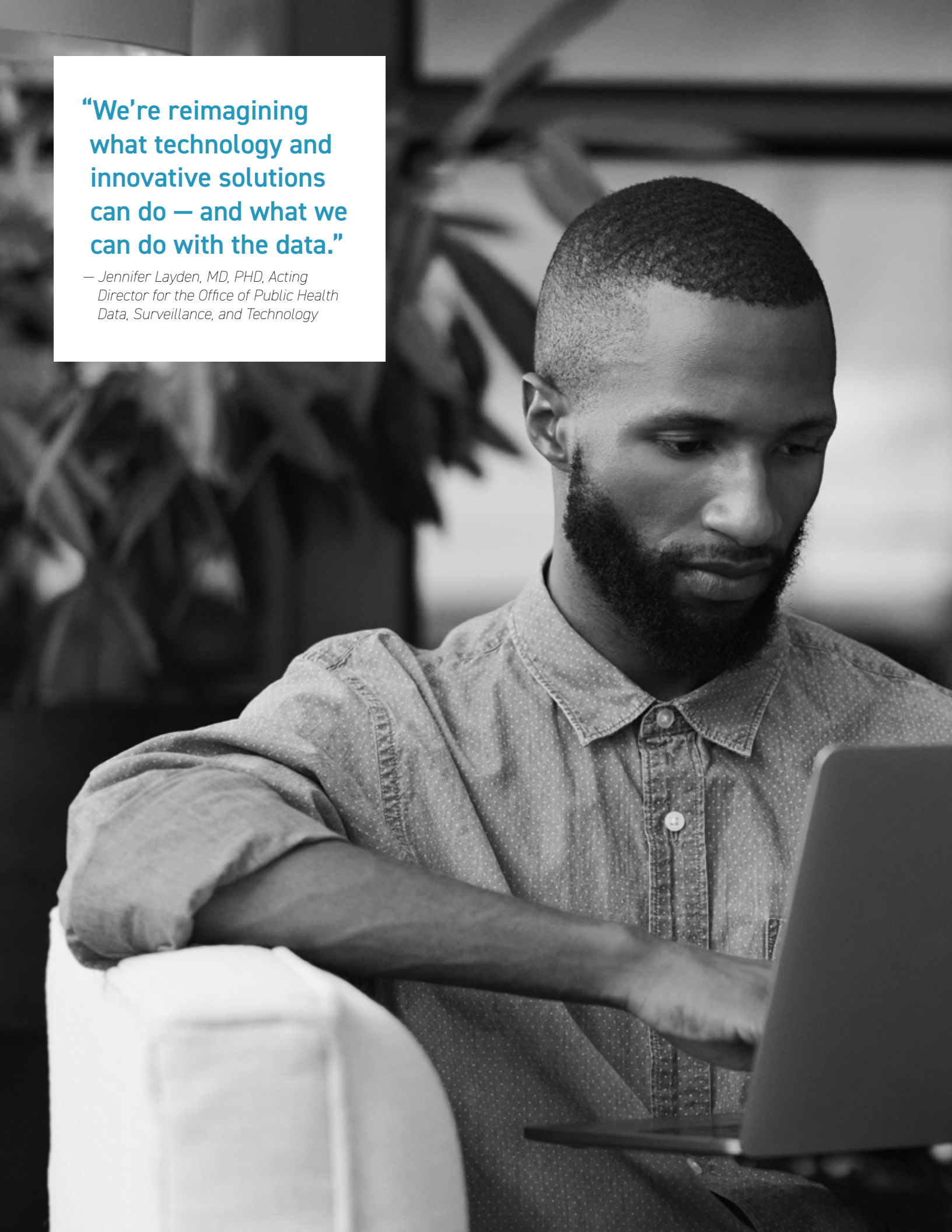
Healthcare's Use of eCR Widens

On January 20, 2020, **187 facilities** were using eCR for 5 pilot conditions.

On December 31, 2020, **22,000 facilities** were using eCR.

As of December 31, 2020, healthcare organizations have sent **more than 27 million COVID-19 reports** since the beginning of the pandemic.

Read the full stories in the 2022 DMI Snapshot



"We're reimagining what technology and innovative solutions can do — and what we can do with the data."

— Jennifer Layden, MD, PHD, Acting Director for the Office of Public Health Data, Surveillance, and Technology

SPRINTING TOWARD THE NEXT GENERATION OF CASE SURVEILLANCE

A discovery sprint offers a fresh and honest look at how we can better monitor all types of diseases

CDC and the U.S. Digital Service are taking an in-depth look at the current technologies, processes, and policies around disease surveillance data. The ambitious goal is to create the next generation of case surveillance for public health.

THE APPROACH: A DISCOVERY SPRINT

In May 2022, we launched a 12-week discovery sprint focused on examining the flow of case reporting and notification through the National Notifiable Diseases Surveillance System (NNDSS).

The sprint team looked at the full lifecycle of the data, starting at the point of care and ending in public health action.

THE RESULT: THE CURRENT APPROACH IS COMPLICATED

The team discovered that the current case surveillance approach is too complicated, and processes do not work smoothly. For example:

- **Too much time and too many resources are spent** on counting, cleaning, and collating case data that do not always meet the needs of users, such as jurisdictions and CDC programs.
- **State and local health departments are faced with competing priorities.** They also find it difficult to meet the many demands placed on them with the workforce, systems, and processes they have.

THE RECOMMENDATIONS: A NEW APPROACH IS NEEDED

The team made recommendations to develop a strategy and execution plan for next-generation case surveillance that would:

1. Support jurisdictions in **collecting and sending case data** in a quick, cost-effective manner.
2. Provide jurisdictions with **a single place to submit case data** to CDC.
3. Take advantage of opportunities to enhance national public health through **streamlined data sharing and situational awareness.**

ABOUT THE SPRINT

DISCOVERY SPRINTS ARE USEFUL WHEN YOU NEED TO QUICKLY BUILD A COMMON UNDERSTANDING OF A COMPLEX SYSTEM.

THE NATIONAL NOTIFIABLE DISEASES SURVEILLANCE SYSTEM MONITORS ABOUT **120 NOTIFIABLE DISEASES AND CONDITIONS** THAT – IF LEFT UNCHECKED – COULD SIGNIFICANTLY HARM THE HEALTH OF COMMUNITIES.

WE INTERVIEWED **148 PEOPLE** REPRESENTING **19 STATE AND LOCAL HEALTH DEPARTMENTS,** AS WELL AS PARTNERS AT THE COUNCIL OF STATE AND TERRITORIAL EPIDEMIOLOGISTS AND THE NATIONAL ASSOCIATION OF COUNTY AND CITY HEALTH OFFICIALS, TO DEVELOP RECOMMENDATIONS FOR THE SYSTEM.

REINVENTING TECHNOLOGIES TO MAKE DATA WORK BETTER

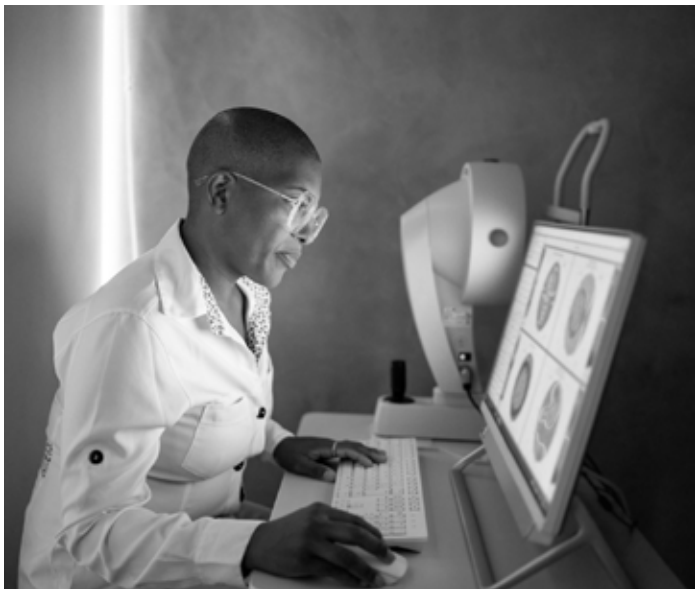
The North Star Architecture is how CDC and its partners are driving to make data available for decision-makers when they need it, not days or weeks later.

For decades, public health staff have struggled under the weight of too many different processes, hard-to-use tools that aren't scalable or reusable, and systems that can't "speak the same language." The North Star Architecture is a core component of CDC's approach for making public health data work better.

Alongside the Office of the National Coordinator for Health IT (ONC), we've been working with our partners to outline the vision for what this new framework will look like. **When we have common technologies in place, CDC and our state and local partners will spend less time managing data, and more time doing the lifesaving work of public health.**

"We need to know where we're going and how we're going to get there. Architecture and design matter. Context matters. Equity matters."

—JUDY MONROE, MD, PRESIDENT AND CEO, CDC FOUNDATION



Where we've been, where we're going...

AUGUST 2021

CDC and ONC brainstormed on early iterations during the process of establishing joint priorities.

SEPTEMBER 2021

CDC began meeting with the Council for State and Territorial Epidemiologists (CSTE) and member states to share ideas and further develop concepts.

DECEMBER 2021

CDC and ONC presented an early set of concepts to thought leaders in healthcare technology and public health and incorporated feedback.

FEBRUARY 2022

ONC presented an initial overview at the CDC Foundation Summit to share these ideas with our partners.

MARCH 2022

CDC presented an initial overview at the PHII DMI Learning Community to begin gathering broader state and local feedback.

SEPTEMBER 2022

CDC launched focus groups with state and local health departments to inform the design of future building blocks.

OCTOBER 2022

CDC completed a pilot project with Virginia Department of Health in which a prototype data processing pipeline aligned with the North Star Architecture blueprint was created and tested.

NOVEMBER 2022

CDC's IT Data and Governance Executive Committee voted to endorse the North Star Architecture.

ONGOING

We continue to gather input as we develop the architecture and improve our data pipelines.

Read the full stories in the 2022 DMI Snapshot

10 WAYS CDC IS CONNECTING PEOPLE WITH DATA FASTER

A key to CDC's Data Modernization Initiative is our investment in a cloud-based suite of Enterprise Data, Analytics, and Visualization (EDAV) tools that are available to the whole agency.

1

Travelers' Health

Helped CDC's infectious disease teams process large volumes of Custom and Border Protection Land and Air Data, reducing reporting time from days to minutes.

2

Pertussis

Decreased reporting time by 95% by streamlining a manual reporting process from a couple of days to just a couple of clicks.

3

Tuberculosis

Supported a new dashboard that allows researchers to better illustrate latent tuberculosis infection and more quickly understand disease patterns.

4

Suicide

Used machine learning to analyze social media and traditional data to "nowcast" suicide trends and show results on an interactive internal platform for CDC's scientists.

5

Flu Vaccines

Supported a new visualization dashboard to showcase influenza vaccination coverage.

6

Drug Overdose

Captured electronic health record information on overdoses from hospital emergency department visits, with a public visualization dashboard refreshed monthly.

7

Genomics

Helped CDC save time and costs to improve its analytic capabilities and present results of genomic analysis in easy-to-understand visuals.

8

Contact Tracing

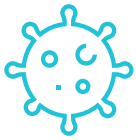
Quickly developed a contact tracing solution that allows travelers to submit information through any device, anywhere in the world, and provides data to jurisdictions for localized action.

9

Emerging Diseases

Streamlined data collection and reporting in surveillance systems that track diseases spread by mosquitoes and ticks.

10



Nursing Home COVID Tracking

Received data from 15,000 facilities to dynamically update reports and ensure the public has the most current information.

Read the full stories in the 2022 DMI Snapshot

DMI IN ACTION

Modernized tools help expand tuberculosis studies

The third phase of the Tuberculosis Epidemiologic Studies Consortium (TBESC-III) is exploring ways to improve testing and treatment for latent tuberculosis (TB) infection.

For TBESC-III, the consortium formed new partnerships with four major primary care sites. With that came a need to receive and analyze data from the sites' electronic medical record (EMR) systems.

They were able to successfully tap into CDC's agencywide resources created through the Data Modernization Initiative (DMI) to make this new phase of work happen.

WHAT THE CONSORTIUM NEEDED:

- A pipeline capable of processing a huge amount of data – more than 500K patients and 12 million records at baseline.
- A process to streamline the different steps of data collection, quality assurance and control (QA/QC), data cleaning, and analytics.
- The ability to provide a fast response back to the participating sites.

WHAT DMI DELIVERED:

CDC's new enterprise data, analytics, and visualization (EDAV) capabilities allowed them to process the massive amount of EMR data quickly and accurately with:

- A **data dictionary** that contains more than 200 variables
- A **real-time QA/QC process** that was able to handle ~13 million records in 1 hour
- A **cleaning step** that automatically addresses minor QA/QC issues
- An **analytic code** to generate a latent tuberculosis infection cascade of care
- **Tools** to visualize the results

THE RESULTS:

Better data automation and data readiness allow CDC staff to respond more quickly, review data in real time, and ultimately report information back to sites. Because of this, they can more efficiently monitor the impact of interventions to determine whether they improve patient testing and treatment outcomes.

The data will help answer:

How is latent TB infection screening, testing, and treatment offered to patients in primary care settings?

What is the impact of primary care interventions that are designed to increase the number of eligible patients who are tested for TB infection and who complete treatment for latent TB infection?

OUTBREAK RESPONSE

Real-time data for emergencies

When lives are on the line, modernization saves staff time and delivers faster data for decisions.

In 2022, we made huge strides toward modernized, cloud-based solutions to meet the needs of current and future emergency responses — no matter how big or small.

RAISING EMERGENCY OPERATIONS TO THE NEXT LEVEL

WHAT WE DID: In 2022, we completed the first phase of CDCReady, a state-of-the-art infrastructure that will streamline where CDC staff go for response-related information before, during, and after participating in a response and support the agency's 24/7/365 emergency operations.

WHY IT MATTERS: CDCReady will save staff time and enable CDC's Emergency Operations Center to function with:

- **Flexibility** to scale systems according to unique response needs
- **Automation** for data sharing and formulating new insights across the agency
- **Security** to set user permissions based on data and their need to know
- **Reduced costs** for operations and maintenance

A COMMON PICTURE FOR REAL-TIME DECISION-MAKING

WHAT WE DID: During the COVID-19 response, the Department of Health and Human Services (HHS) established HHS Protect to collect, analyze, and share COVID-19 data for the response. On February 18, 2022, management and operational ownership of the platform transitioned to CDC's National Center for Emerging and Zoonotic Infectious Diseases.

WHY IT MATTERS: This transition continues the mission to securely process, use, and share more types of data – and more details – than ever before.

SUPPORTING BROADER PANDEMIC DETECTION AND RESPONSE

WHAT WE DID: CDC developed the System for Outbreak Response, Coordination, and Surveillance (SOURCE) to improve surveillance for multiple respiratory illnesses. This includes diseases like influenza, measles, mumps, and Legionnaires' disease.

WHY IT MATTERS: SOURCE can scale up to meet all sizes of emergencies, while making it easier to share data in real-time with state and local partners.

With the surveillance platform of the future, we will be able to use the same system for any size outbreak — from 300 cases to 3 million.

We will be better prepared for the next pandemic or big outbreak.

Read the full stories in the 2022 DMI Snapshot

SIX WAYS WE'VE INNOVATED FOR BETTER PUBLIC HEALTH

"We're reimagining what technology and innovative solutions can do — and what we can do with the data."

—JENNIFER LAYDEN, MD, PHD, ACTING DIRECTOR, CDC OFFICE OF PUBLIC HEALTH DATA, SURVEILLANCE, AND TECHNOLOGY

1 WE DISPLAYED CANCER DATA FOR AREAS SMALLER THAN A COUNTY.

WHAT WE DID: We used modern approaches to provide a close-up view of six types of cancer that impact communities too often.

HOW WE DID IT: Following the first federal release of cancer incidence data at the sub-county level in March 2022, CDC's National Center for Environmental Health, and the National Program of Cancer Registries piloted new ways to display cancer data for areas smaller than a county. The data came from 27 cancer registries and covered six cancer types: lung, female breast, prostate, colorectal, kidney, and melanoma.

HOW IT'S HELPING PEOPLE: Finer spatial resolution data can be used to guide efforts to lower the chance that people in a community will get cancer.

2 WE AUTOMATED THE PROCESS TO CLASSIFY HEALTH RECORDS TO BETTER UNDERSTAND THE RISKS MOTHERS AND BABIES FACE FROM COVID-19.

WHAT WE DID: We found a solution that allowed medical records to be interpreted and recorded by computers rather than by hand.

HOW WE DID IT: By applying innovative natural language processing methods, CDC and Georgia Tech Research Institute tapped into the power of the Surveillance for Emerging Threats to Mothers and Babies Network (SET-NET) to rapidly analyze text from thousands of patient health records and classify each reported COVID-19 case as asymptomatic, mild, moderate-to-severe, or critical.

HOW IT'S HELPING PEOPLE: This work helps us better understand the increased risks and severity of COVID-19 infection as well as other diseases.

Read the full stories in the 2022 DMI Snapshot

3 AMID THE RAPID GROWTH OF WASTEWATER SURVEILLANCE, WE DISCOVERED BETTER WAYS TO DISPLAY DATA FOR THE PUBLIC.

WHAT WE DID: CDC epidemiologists teamed up with data visualization experts to reinvent the public-facing wastewater surveillance dashboard.

HOW WE DID IT: Experts developed an easy-to-understand interface to help the data-savvy public understand percentage changes and relative levels of SARS-CoV-2 concentrations in each community, and simplified how trends in cases were displayed.

HOW IT'S HELPING PEOPLE: People can use the dashboard to understand what's going on in their community, while more in-depth pages support the needs of health departments, academics, and other technical viewers.

4 WE'VE LINKED DATASETS TO DELIVER NEW INSIGHTS THAT CAN ANSWER AMERICA'S COMPLEX HEALTH QUESTIONS.

WHAT WE DID: Data linkage projects are enhancing algorithms through data science tools like machine learning.

HOW WE DID IT: The National Center for Health Statistics (NCHS) data linkage program used enhanced algorithms to conduct the first-ever linkage between NCHS' survey data and Veterans Administration (VA) administrative data.

HOW IT'S HELPING PEOPLE: The resulting data provide a unique opportunity to examine the factors that influence disability, chronic disease, healthcare utilization, and expenditures among veterans enrolled in the VA healthcare system.

5 FROM INJURIES TO DISABILITIES TO INFECTIOUS THREATS, WE'VE MADE MORE DATA THAN EVER AVAILABLE FOR RESEARCHERS, POLICYMAKERS, AND THE PUBLIC.

WHAT WE DID: We've created open datasets that are publicly accessible and easy-to-use.

HOW WE DID IT: In 2022, we added 150 readily accessible datasets to data.cdc.gov. We also continued improving the CDC COVID Data Tracker, where 64 data feeds provide data on important topics like vaccines and community case levels. 5 million users of CDC WONDER had access to more than 100 unique datasets.

HOW IT'S HELPING PEOPLE: Open data is one of the most significant innovations we've seen in public health data modernization. It transforms how we gather data and how we release timely information to the public, making important data available often in weeks, not months or years.

6 WE TAPPED INTO NEW METHODS AND DATA SOURCES TO "NOWCAST" WEEKLY SUICIDE TRENDS.

WHAT WE DID: We brought together data from multiple sources to help CDC's experts understand suicide trends in near real-time.

HOW WE DID IT: The first step was to compile new sources, such as social media data and online forum data. Next, we developed and applied new methods for analyzing and sharing the data, which resulted in an automated process and an internal-facing prediction and visualization platform. The process that was developed is now being adapted for other injury outcomes.

HOW IT'S HELPING PEOPLE: Suicide is a leading cause of death in the United States. CDC's scientists can use this data to model national suicide death trends on a week-to-week basis, allowing evidence-based public health decisions to be made. For example, in a recent suicide cluster investigation in Iowa, the dashboard was used to provide real-time national context.

PUTTING EQUITY AT THE CENTER OF DATA

Research indicates that as much as **80% to 90%** of a person's health is determined by health-related behaviors, socioeconomic, and physical environment factors that typically are outside of medical care. These factors are captured in what are called **social determinants of health (SDOH)** data.

CDC's Data Modernization Initiative is tapping into both health data and SDOH data to better identify and address the drivers of health inequities.

"Health equity should not be an afterthought; it should be integrated from concept to finish. Only then will we be able to get the answers our communities need and deserve."

—KRISTIE CLARKE, EPIDEMIOLOGIST, CDC PUBLIC HEALTH INFORMATICS OFFICE

WHEN IT COMES TO CHRONIC DISEASES, WE'RE MAKING SURE HEALTH EQUITY IS "BAKED IN" TO EVERYTHING WE DO.

CDC's National Center for Chronic Disease Prevention has been at the forefront of applying new technologies and approaches that support better data for all.

HIGHLIGHTS

We added modules that track topics like food and housing insecurity, transportation, social support, well-being, and economic stability into our surveillance systems for behavioral risk factors and pregnancy risk, as well as into the PLACES system.

We launched a collaborative effort to expand the exchange of electronic health record data on housing, food access, transportation, and other social needs.

We're enhancing how health information is exchanged between health systems and community service organizations who address chronic health conditions—for example, lifestyle change programs and diabetes self-management programs.

Foundational to our work is understanding the causes of disease — and the causes of health.

Read the full stories in the 2022 DMI Snapshot

TEAMWORK IS RESULTING IN NEW INSIGHTS THAT HELP PEOPLE EXPERIENCING HOMELESSNESS, INCARCERATION, OR LIVING WITH DISABILITIES.

Experts from CDC's Public Health Informatics Office, Office of Readiness and Response, and Center for Birth Defects and Developmental Disabilities came together to improve the standards and systems that are used to collect information on social determinants of health.

HIGHLIGHTS

We partnered with the National Healthcare for the Homeless Council to crowdsource data from clinics and create a co-branded, web-based dashboard to help fill in some of the existing gaps in COVID-19 data.

We worked with UCLA law school to use "web scraping" to find data on how people experiencing incarceration were impacted by COVID-19.

We developed ways to understand the impact of COVID-19 on people living with disabilities by using medical diagnosis codes as a proxy for missing information.

We're ensuring the data we need are accessible and can fuel action more easily — both for emergencies like COVID-19 and for everyday threats like diabetes, pneumonia, heart disease, or injuries.

AS NEW TECHNOLOGY REVEALS WHAT COULDN'T BE SEEN BEFORE, WE'RE ALSO LOOKING AT ITS FUTURE IMPACTS ON PUBLIC HEALTH.

CDC's Office of Science teamed up with experts from Georgia Tech Research Institute to help uncover and prevent potential biases that may arise as we incorporate more artificial intelligence into our data workflows.

HIGHLIGHTS

We're working to develop guidance and training resources that will help public health researchers effectively navigate many of the health equity-related challenges they encounter when conducting data science using health data.

We're looking at every stage of the workflow—the data sources themselves, how data are prepared, and how they are analyzed and reported.

We want to find the right tools and techniques to consider and assess health equity from end-to-end, so that we can uncover which improvements will give us the greatest potential to increase health equity.

If we apply new technologies like artificial intelligence without considering equity issues, the risk is that we could make the problem larger instead of solving it.

PEOPLE

IN THEIR OWN WORDS

DMI supports state and local partners

Modernization means less paperwork…and more help for people in communities.

"We had all these ideas, but we didn't have the resources or funding. With DMI funding, DC was able to rapidly implement many changes — some in just a few weeks."

— MATT MCCARROLL, CHIEF OF LABORATORY OPERATIONS, WASHINGTON, DC, PUBLIC HEALTH LABORATORY

WASHINGTON, DC, LAB PUSHES DOWN PAPERWORK

Thanks to electronic lab reporting, epidemiologists in the Washington, DC lab no longer waste time manually standardizing information on big piles of spreadsheets that have come in by fax. They also saved time by using DMI funds to hire dedicated informaticians and automating test ordering systems, reducing a 48-hour process to 12 hours.

LINKED DATA LEADS DALLAS COUNTY, TX, TO THOSE MOST IN NEED

More granular vaccination data — down to the census block level — has allowed Dallas County's health workers to know which doors to knock on, focusing on 17 zip codes where their most vulnerable residents live and bringing pop-up vaccination clinics right into neighborhoods.

"DMI is a tremendous opportunity to make a big leap forward for public health. And COVID has definitely highlighted the importance of doing this."

— PHILIP HUANG, MD, MPH, DIRECTOR, DALLAS COUNTY HEALTH AND HUMAN SERVICES AND CHAIR-ELECT, BIG CITIES HEALTH COALITION

"Not only does DMI help with better, more detailed, quicker data, it provides options so that we are not reliant on a single source of data during a major crisis like a pandemic."

— ANGELA DUNN, MD, MPH, EXECUTIVE DIRECTOR AT SALT LAKE COUNTY HEALTH DEPARTMENT AND CSTE PRESIDENT

SALT LAKE COUNTY, UT, CONNECTS PEOPLE WITH PREVENTION

Combining data from criminal justice, homelessness, health data, and vital records data has allowed Salt Lake County to understand the opioid abuse problem more holistically. When the data identified a cluster of fentanyl-related deaths, they sent more fentanyl test strips out through syringe providers, saving more lives.

Read the full stories in the 2022 DMI Snapshot

IDAHO RAPIDLY CATCHES CASES OF MIS-C IN CHILDREN

In Idaho, electronic case reporting (eCR) is being used to catch cases of multisystem inflammatory syndrome (MIS-C), a dangerous condition that can strike children weeks after COVID-19 infection. An evaluation of medical records in 2022 indicated that, among Idaho facilities that had implemented eCR, no cases were missed.

"We are so thankful for the DMI funding and are enthusiastically enhancing and supplementing our small, but mighty, workforce to tackle the data surveillance and infrastructure improvement projects that were on our 'when we have the money' list."

— KATHRYN TURNER, PHD, MPH, DEPUTY STATE EPIDEMIOLOGIST, IDAHO DIVISION OF PUBLIC HEALTH

"This is a solvable problem. We may not know everything about a pandemic, but we can always find ways for data to connect."

— ANNE ZINK, MD, CHIEF MEDICAL OFFICER FOR THE STATE OF ALASKA AND ASTHO PRESIDENT

SIMPLEREPORT APP MAKES ALASKA CASES COUNT

The free SimpleReport app that CDC developed with the US Digital Service filled critical gaps in COVID-19 case reporting in remote rural areas, in schools, and in the fishing and tourism industries Alaska's economy relies on. SimpleReport was a "huge game changer," enabling the state to go from calls and faxes to an automated tool.

Read the full stories in the 2022 DMI Snapshot

ACCELERATING PROGRESS THROUGH PARTNERSHIP

Our partners are at the heart of data modernization — past, present, and future.

A NEW CONSORTIUM GATHERS INSIGHT AND PERSPECTIVES.

WHAT WE DID:

In Spring 2022, CDC and the Office of the National Coordinator for Public Health Information Technology (ONC) launched the Consortium for Data Modernization to bring together a diverse group of voices from across government, public health, industry, and academia.

WHY IT MATTERS:

The Consortium offers CDC and its partners a place to discuss the common challenges we share. For example, the group has been at the center of gathering input as we develop the North Star Architecture and improve public health data pipelines.

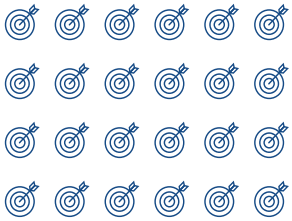
WHAT'S NEXT:

In 2023, the Consortium will continue to meet on critical issues - from data exchange, core data sources, and immunization systems, to non-infectious diseases, emergency response data, governance, workforce development, and the future of public health surveillance.

BY THE NUMBERS

18

CONSORTIUM MEETINGS IN 2022



24

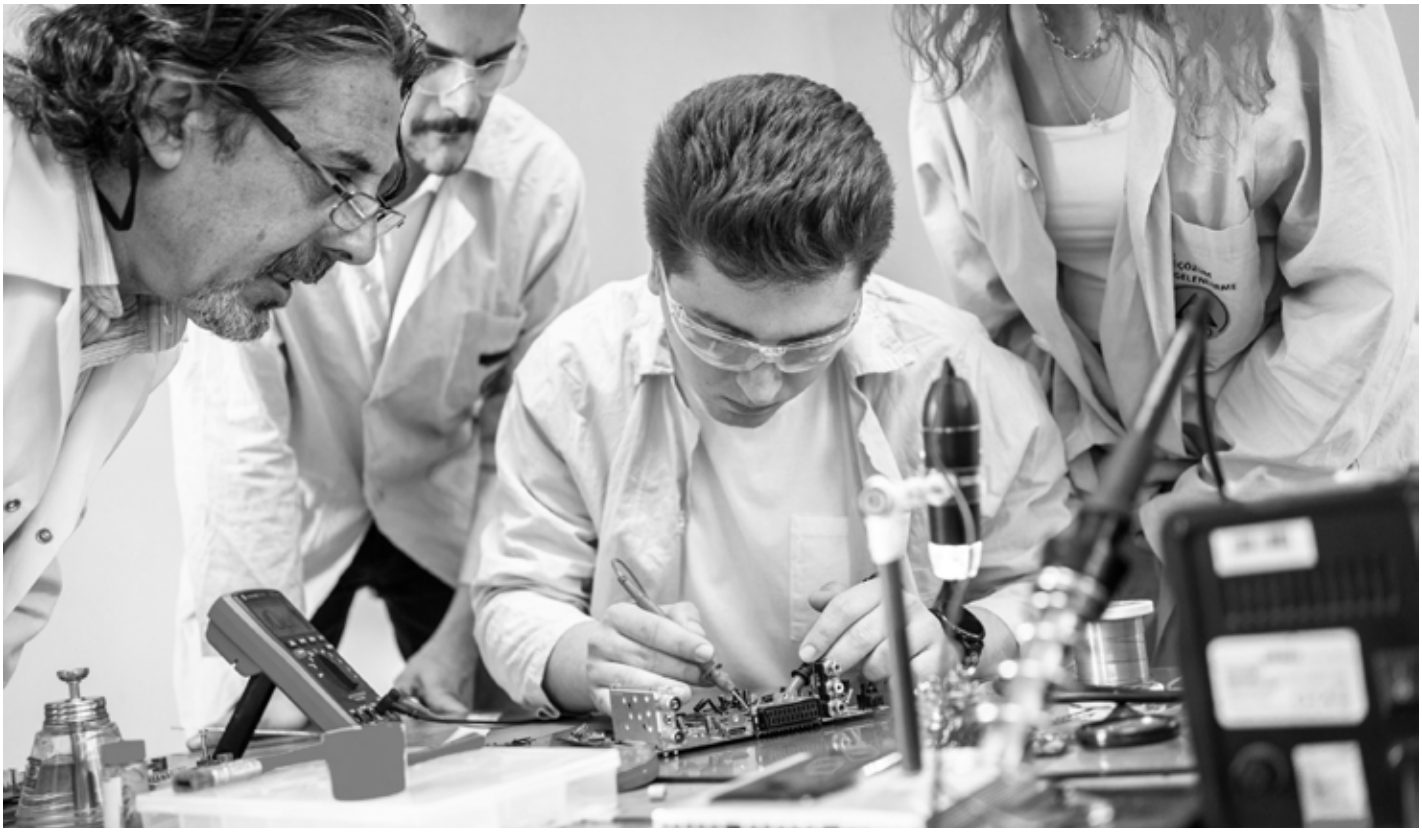
HIGH-IMPACT TOPICS FOR DATA MODERNIZATION DISCUSSED

47

MEMBERS REACHED ACROSS 13 ORGANIZATIONS

The purpose of the Consortium is to only seek individual perspectives and experiences, not group consensus advice from partners. Consortium participation varies based on discussion topic.

Read the full stories in the 2022 DMI Snapshot



2022: PARTNERSHIP IN ACTION

A PUSH FOR CONTINUED PROGRESS

The Data: Elemental to Health campaign stressed the need for annual resources for state, local, and territorial health departments, stating that "DMI is the single most important investment to ensure continued progress in modernizing our public health data infrastructure."

ADVANCING WORLD-CLASS DATA AND ANALYTICS AT CDC

A new Data and Surveillance Workgroup was formed to support the CDC Advisory Committee to the Director by looking at agency-wide modernization activities from a holistic perspective and recommending solutions.

EXTENDING THE MODERNIZATION CONVERSATION

In Spring 2022, CDC Foundation launched the "Lights, Camera, Action: The Future of Public Health" summit series to discuss the data and technology challenges faced by public health. Altogether, CDC Foundation summits and listening sessions reached 73 partners across 64 organizations.

INNOVATION PARTNERSHIP

DMI continued working in partnership with the U.S. Digital Service to create pandemic-ready tools and systems for health departments, modernize the hiring process at CDC, and conduct "discovery sprints" to quickly learn about some of the biggest challenges facing public health data and surveillance.

PEER-TO-PEER LEARNING

The Public Health Informatics Institute brought together a DMI Learning Community where modernization leads from 64 jurisdictions share practical knowledge, build partnerships, and collaborate around common challenges.

A NEW ALLIANCE FOR HEALTHCARE DATA EXCHANGE

The Helios FHIR Accelerator began actively working with healthcare providers and insurers on ways to access immunization history information to help our partners on the frontlines improve vaccine coverage rates.

Read the full stories in the 2022 DMI Snapshot

BRIDGING THE GAPS IN THE DATA SCIENCE WORKFORCE

We're increasing the ability to use next-generation skills for public health action.

CDC's data modernization strategy includes recruitment, upskilling (improving the knowledge and abilities of existing staff), reskilling (learning new skills within a role), and retention of staff with data science and other data-oriented skills.

BY THE NUMBERS

25
DATA SCIENCE UPSKILLING TEAMS TACKLING COMPLEX PUBLIC HEALTH CHALLENGES

600
JOB APPLICATIONS RECEIVED IN RECORD TIME THROUGH A NEW HIRING PROCESS

3,700+
HOURS OF DATA SCIENCE TRAINING ACCESSED IN CDC'S DATA ACADEMY

4
PRESIDENTIAL INNOVATION FELLOWS HIRED TO SUPPORT MODERNIZATION PRIORITIES

DEVELOPING STATE-OF-THE-ART SKILLS

THE CHALLENGE: With the development of new systems comes the need to help people learn how to use them.

WHAT WE'RE DOING: The Data Science Upskilling program offers a team-based, "learning by doing" approach that enables CDC staff and fellows to develop and use their data science skills while working to address the nation's most pressing health priorities. In July 2022, a total of 25 teams demonstrated how they used data science skills in predictive analytics, data reporting, artificial intelligence, and machine learning to solve complex public health challenges.

WHY IT MATTERS: This year's teams tackled a broad variety of urgent public health challenges, such as evaluating rural health and access to care, providing local health departments with new visualizations of viral hepatitis data, evaluating mental health and stress levels in West Virginia's 3rd graders, and looking at the intersection of COVID-19 and health equity.

Did you know?

BUILT FROM THE SAME IDEAS AS DSU, CSTE'S DATA SCIENCE TEAM TRAINING PROGRAM IS A TEAM-BASED, ON-THE-JOB TRAINING PROGRAM TO PROMOTE DATA SCIENCE UPSKILLING AT STATE, TERRITORIAL, LOCAL, AND TRIBAL PUBLIC HEALTH AGENCIES.

A one-stop shop for modernizing skills

CDC's Data Academy delivers a one-stop source to help users of all experience levels learn about the agency's enterprise data analytics and visualization (EDAV) tools, including self-paced courses in Databricks, Power BI, R, Socrata, Tableau, and Azure Data Factory. The number of training hours completed by CDC staff increased 83% in 2022, to a total of more than 3,700 hours by the end of the year.

Read the full stories in the 2022 DMI Snapshot

RECRUITING DIVERSE, WORLD-CLASS TALENT

THE CHALLENGE: Across the entire government, about half of all open and competitive job announcements close with no job selection.

WHAT WE'RE DOING: In 2022, CDC partnered with the U.S. Digital Service to pilot Subject Matter Expert-Quality Assessments (SME-QA), an innovative strategy used by federal agencies to recruit world class and diverse talent as efficiently as possible. In the first test of the new hiring process at CDC, 600 applications were received in record time, and qualified candidates were shared with hiring managers across the agency.

WHY IT MATTERS: Attracting and retaining the best talent will accelerate the work currently underway and maximize CDC's modernization success. We will be able to better identify, recruit, and retain a critical workforce in Health IT, data science, and cybersecurity who can help public health work better and faster to generate meaningful public health insights.

Seeking solutions....

OVER THE PAST DECADE, STATE AND LOCAL GOVERNMENT HEALTH AGENCIES HAVE LOST AT LEAST 40,000 POSITIONS - MORE THAN ONE-FIFTH OF THE TOTAL WORKFORCE.[1]

Bringing industry innovators to CDC

In 2022, we welcomed four new Presidential Innovation Fellows who bring private-sector technical experience to support modernization goals for the whole agency. As industry experts, they are already applying their experience as changemakers to help us all think differently and come together across CDC's many projects, programs, and centers.

[1] When We Need Them Most, the Number of Public Health Workers Continues to Decline - de Beaumont Foundation

Read the full stories in the 2022 DMI Snapshot

MAKING SPACE FOR BOLD IDEAS

Aligning change at CDC will ensure modernization is executed well.

With CDC's Data Modernization Initiative (DMI), we're taking a team approach that enhances collaboration within CDC, encourages adaptability, and creates a space where bold ideas and creativity can thrive.

"It has been enlightening to have conversations with others who I would not normally interact with — to know that we aren't alone in the data and system problems we're facing, and then to work together to find the common ground."

— LYNDSAY BOTTICHIO, PRIORITY TEAM CO-LEAD, AUGUST 2022

BY THE NUMBERS

200+

CDC STAFF ACTIVELY ENGAGED

We stood up cross-agency Implementation Teams to work together on specific DMI priorities and objectives.

Staff who work at many different jobs – and on all different diseases and conditions – are now in one place. These diverse teams are helping to unlock the answers to large public health challenges.

10

PRIORITY TEAMS ALIGNED TO 5 DMI PRIORITIES

Teams worked to develop the next phase of Objectives and Key Results (OKRs) and align them to the data modernization priorities.

OKRs are a collaborative goal-setting model for achieving challenging, ambitious goals. Teams reviewed the objectives and updated key results for the next phase of DMI.

15

ESSENTIAL ACTIVITIES IDENTIFIED

Teams identified essential activities to prioritize the work and resources needed to achieve DMI goals.

As DMI embarks on the next phase, teams are transitioning from planning to actively learning, piloting, and testing so that we can modernize effectively together.

Did you know?

OVER TIME, MUCH OF THE WORK PUBLIC HEALTH DOES HAS BECOME SEPARATED, OR "SILOED," BECAUSE PROGRAMS ARE USUALLY FUNDED AND SET UP TO ADDRESS ONLY A SINGLE DISEASE OR CONDITION. AN IMPORTANT JOB FOR DATA MODERNIZATION IS TO TAKE THE DISCONNECTED DATA STREAMS AND AREAS OF EXPERTISE AND BRING THEM TOGETHER — WHAT WE CALL "BREAKING DOWN SILOS."

Read the full stories in the 2022 DMI Snapshot

"We really are working together differently, and it's intentionally designed that way. Data modernization is complex, and all perspectives and brains are needed to determine next steps."

— Katie Fullerton, MPH, Senior Advisor for Surveillance and Data Modernization (Office of Readiness and Response)

POLICY

INVESTING RESOURCES WISELY

We're charting a unified course for every new data and technology project.

As CDC embarks on a wide variety of data modernization projects, we want to make sure that we're solving problems in ways that line up with the future vision of data modernization for the whole agency.

Unified IT and Data Governance (ITDG) continues to play an important role in making sure our investments are smart, efficient, and drive modernization ahead in the most effective ways.

IN 2022, ITDG CONTINUED MOVING CDC FORWARD TOGETHER.

120	new investments reviewed
98	decision memos provided

Shared governance has already saved millions for CDC by:

We recently added an Architecture Review Team to our governance process to ensure technical alignment across the agency.

Read the full stories in the 2022 DMI Snapshot

KEY INVESTMENTS FOR DATA MODERNIZATION

ITDG has been key to the adoption of modern technologies and processes at CDC that will help us realize our goals.

ENTERPRISE DATA, ANALYTICS, AND VISUALIZATION

CDC CLOUD STRATEGY

CLOUD DATA HUB

ENERPRISE DATA CATALOG

NORTH STAR ARCHITECTURE

CDC “FRONT DOOR” FOR DATA EXCHANGE

PROGRAM AGNOSTIC DATA ECOSYSTEM

CASE REPORTING DATA FLOW

INTEROPERABILITY

Solving the puzzle of data access and exchange.

For systems to "speak the same language," we need all the right pieces in place.

Modernization relies on getting data where it needs to go to protect health. Newer policies and modern standards can help us access and exchange important health information that is not readily available now.

In 2022, we continued working with the Office of the National Coordinator of Health IT (ONC) and other partners from across healthcare, government, and the private sector on shared priorities to advance public health.

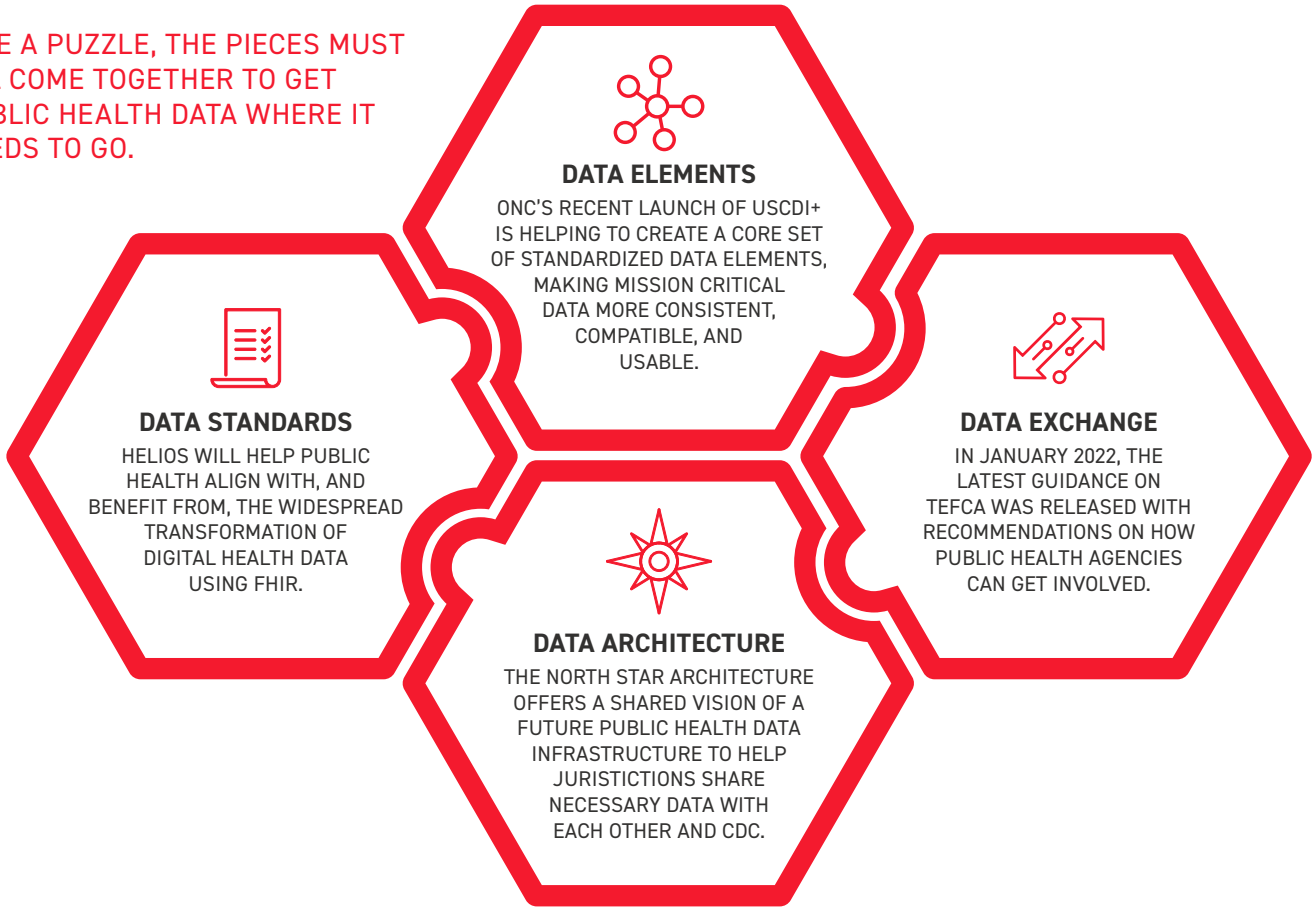
Together, we're looking at interoperability from different angles, including the agreements we use to exchange data, the architecture it flows through, the elements that are captured, and the standards that are used to share them. For example:

The **Trusted Exchange Framework and Common Agreement** (TEFCA) guides how different, individual systems connect to share information consistently – without having to come up with their own approaches and rules for doing so.

Fast Healthcare Interoperability Resources, or FHIR (pronounced "fire"), offers specifications that give data structure and make information available to support the needs of public health officials and their many partners.

And the **North Star Architecture** shows where everyone is going together. You can think of it like a diagram for planning a city that includes things like a library or post office that will provide benefits for many people in the community.

LIKE A PUZZLE, THE PIECES MUST ALL COME TOGETHER TO GET PUBLIC HEALTH DATA WHERE IT NEEDS TO GO.



Read the full stories in the 2022 DMI Snapshot

MAKING MODERNIZATION METRICS MATTER

Monitoring and evaluation changes are delivering a more unified view of progress

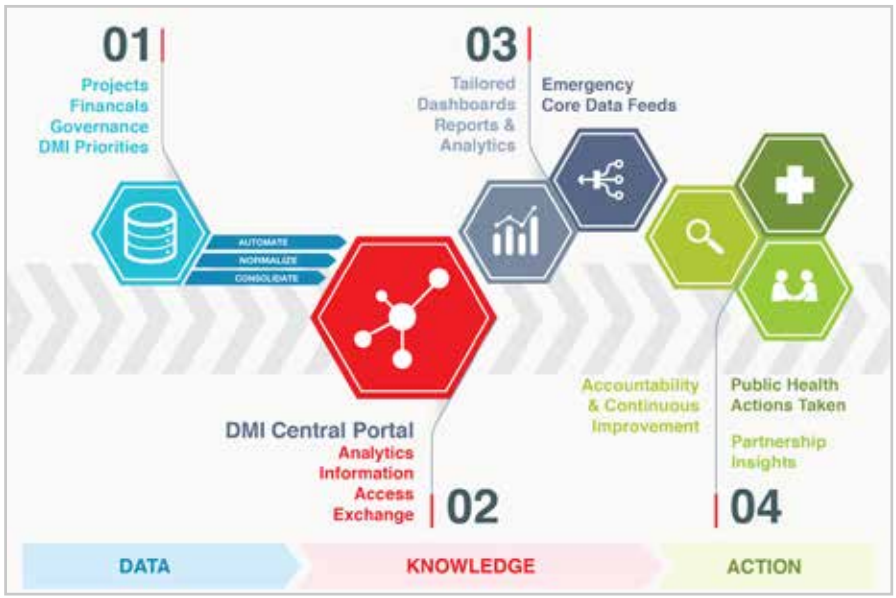
CDC's Data Modernization Initiative involves many different systems – large and small – with data on all kinds of diseases and conditions. It tackles everything from emergency response to everyday disease surveillance to partnerships, policies, innovation, governance, staffing, and training.

MONITORING AND EVALUATION ARE HOW WE MAKE SURE WE ARE DELIVERING ON THE PROMISE OF DATA TO PROTECT AMERICA'S HEALTH.

Measuring progress is not just about what we think we need to know, but also about understanding the value of all our work and ensuring we can answer timely and relevant questions in a rapidly changing environment. By building a strong monitoring and evaluation framework, we can support success every step of the way.

AS DMI CONTINUES TO GROW, SO DO THE WAYS IN WHICH WE TRACK OUR PROGRESS AND SUCCESS.

By tracking common metrics in critical areas across all projects and programs involved in DMI, we can begin to get a birds-eye view of what's happening. The goal is to bring together CDC data and information in ways that will fuel our ability to understand progress in near real-time. We're now building a new Info Hub that will include all the metrics and information we gather through monitoring and evaluation, as well as data from other DMI-related systems.



How do we take 150+ activities and measure progress in a way that's consistent and usable?

How do we compare where we are now to where we were before?

How do we make sure we're all heading in the right direction, and in the right ways?

Read the full stories in the 2022 DMI Snapshot

DATA READY

Modern policies and processes

A big part of our mission is to ensure the right policies, authorities, data use agreements (DUAs), and relationships are in place to support modern data exchange.

WHEN IT COMES TO DATA MODERNIZATION, TECHNOLOGY AND POLICY MUST EVOLVE HAND IN HAND.

THE CHALLENGE:
Currently, CDC receives data from a wide variety of public health and healthcare sources, often in slow or inconsistent ways, creating major gaps and blind spots in our public health surveillance systems.

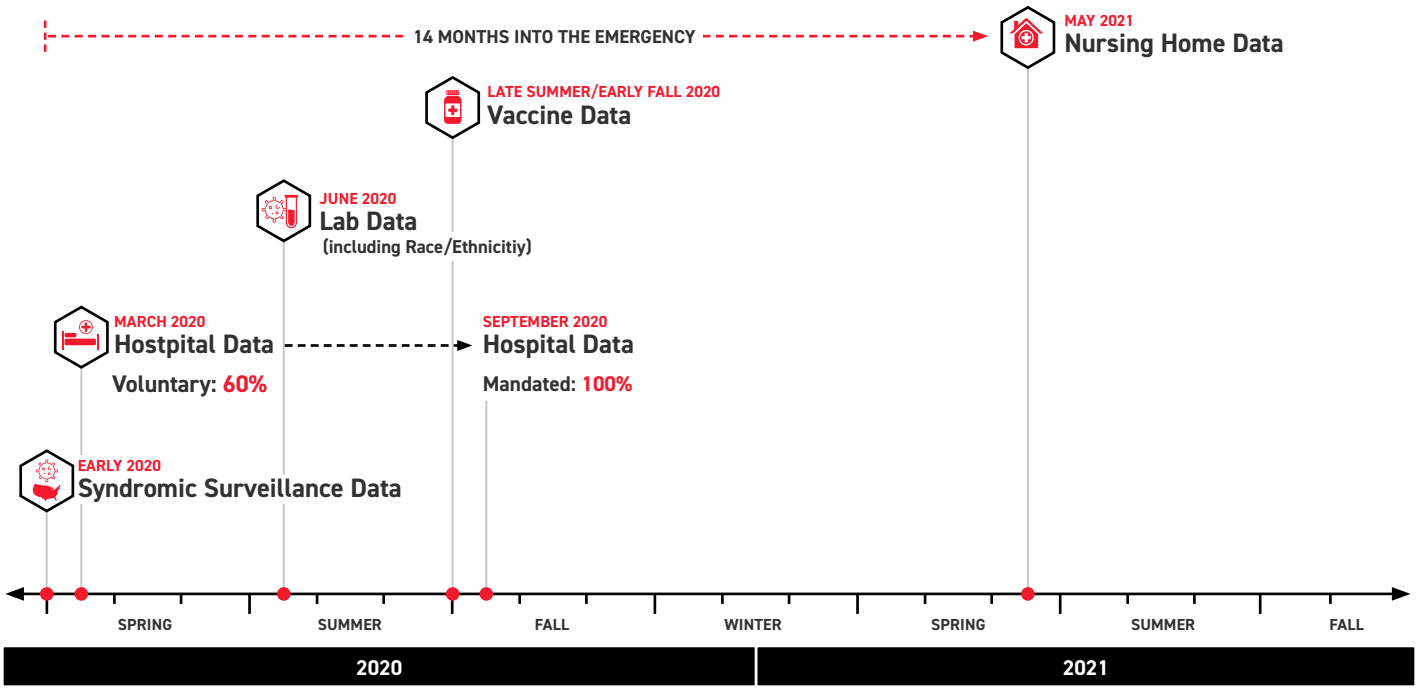
THE SOLUTION:
As a nation, we need to coordinate our data so that:

- **Local health departments** can target interventions where they're needed most
- **States** can react to developing or ongoing situations based on actionable and complete information
- **We have a clear, national picture** that transcends state borders to guide federal decision-making and improve our country's ability to act quickly

"We wouldn't attempt to predict severe weather like this, so why do we try to forecast the spread of deadly diseases this way?"

— FORMER CDC DIRECTORS TOM FRIEDEN, JULIE GERBERDING, JEFFREY KOPLAN, WILLIAM L. ROPER, AND DAVID SATCHER, THE HILL, MARCH 2022

WHEN THE COVID-19 PANDEMIC HIT, IT TOOK TOO LONG TO ACCESS IMPORTANT DATA FOR MAKING DECISIONS ABOUT THE NATION'S PUBLIC HEALTH.



Read the full stories in the 2022 DMI Snapshot

THERE IS NO TIME TO WASTE. WE NEED AGREEMENTS IN PLACE BEFORE EMERGENCIES STRIKE.

THE CHALLENGE:
When important information about disease outbreaks or other emergencies needs to be shared across jurisdictions or with CDC, agencies establish DUAs to ensure valuable disease information is shared securely. However, it can take 6 or more months to develop a new DUA or rework an existing one. Sometimes multiple DUAs may be needed to get data where it needs to go.

THE SOLUTION:
CDC's investment in modernizing data began prior to the COVID-19 pandemic, and we continue to improve national systems for communicating public health data. We're working to streamline the process for getting DUAs in place by:

- Establishing templates, so we're not starting from scratch when we need to exchange data quickly
- Ensuring the language in the agreements protects privacy
- Making sure we're gathering the minimal amount of data necessary for determining how to halt transmission and save lives.

However, it can still take multiple months to get agreements in place, and emerging threats move far more quickly across state borders. Without data authority, CDC will not be able to act rapidly and efficiently to enable states and jurisdictions to act to address these threats.

We need to remain proactive in finding ways to make sure data sharing can happen as efficiently and effectively as possible, so that state and local jurisdictions, federal agency partners, and CDC programs can better access and use the data they need, right when they need it most.

DMI
@ CDC

This report was developed by the Office of Public Health Data, Surveillance, and Technology, Centers for Disease Control and Prevention

CDC Data Modernization Leads
Jennifer Layden, MD, PhD
Acting Director, Office of Public Health Data, Surveillance, and Technology, Centers for Disease Control and Prevention

Suzi M. Connor
CDC Chief Information Officer
Director, Office of the Chief Information Officer

Special thanks to our partners across public health and the nation, who have driven the success of the Data Modernization Initiative from its inception, and who continue to transform the future. And to the many people at CDC whose dedication improves people's wellbeing every day.

Snapshot Co-Leads
Jessica Alexander
Dana Pitts

Project Team
Matthew Dempsey, Amy Ferzoco, Jennifer Jester, Allison MacNeil, K. Ceres Wright

Graphic Design and Digital Leads
Deirdre Launt
Joanne Stein

Design Team
Chris Herman
Meredith Newlove

Suggested Citation
Office of Public Health Data, Surveillance, and Technology, Centers for Disease Control and Prevention, *Data Modernization Initiative 2022 Snapshot: Transforming Data to Protect Health.* Atlanta, GA: Centers for Disease Control and Prevention: April 2023.

Read the full stories in the 2022 DMI Snapshot