Public Health Data + IT Modernization at CDC

Daniel B. Jernigan, MD, MPH Acting Deputy Director for Public Health Science and surveillance (DDPHSS)

April 1, 2021



WHAT IS THE DATA MODERNIZATION INITIATIVE?

DMI

@ CDC

CDC is at the heart of a national effort to create modern, integrated, and real-time public health data and surveillance that can protect us from any health threat.

The Why: Making Modernization Meaningful

Data is Moving Slower than Disease



The nation's public health data systems are antiquated and in dire need of security upgrades - paper records, phone calls, spreadsheets and faxes requiring manual data entry are still are in widespread use and have significant consequences including **delayed detection** and response, lost time, missed opportunities and lost lives. "

Testimony of Janet Hamilton, Director of Science and Policy at CSTE, speaks at Public Witness Day, Labor, Health and Human Services, Education, and Related Agencies (116th Congress)

The What: A Strategy for Connected, Modern Public Health Data

THE NEW WORLD OF PUBLIC HEALTH DATA

TIMELY, GRANULAR, ACCURATE, ACCESSIBLE...



DMI IS A UNIFYING FOUNDATION FOR CHANGE

DMI IS BOTH RESOURCED AND COMPREHENSIVE, AND IT UNIFIES US IN WAYS NO OTHER STRATEGY HAS BEFORE.

PARTNER SUPPORT

 Unprecedented connection to public health and healthcare partners, state and local health departments, researchers, academics, innovators, and industry leaders

CONGRESSIONAL SUPPORT

• First-ever funding dedicated to modernization, accelerated by CARES

PUBLIC SUPPORT

• New threats change awareness and demands

CDCSUPPORT

Unified, whole-of-agency approaches

FUNDING SUPPORT FOR THE INITIATIVE

FY 2020 appropriation provided CDC with \$50 million to modernize its IT and data systems. FY 2021 appropriation provided CDC with \$50 million to continue data modernization activities. The Coronavirus Aid, Relief, and Economic Security (CARES) Act provided an additional \$500 million to CDC to advance surveillance goals for the nation.

BENEFITS TO PUBLIC HEALTH





EMPOWER SCIENTISTS Focus on Knowledge Discovery and Public Health



GET BETTER DATA Access to Complete, Accurate, and Up-to-Date Information



SAVE TIME Faster and More Streamlined IT Development and Implementation



ENCOURAGE INNOVATION Enhance Productivity and Creativity



PROMOTE COLLABORATION Ensure Alignment & Enhance Productivity



ENSURE SUSTAINABILITY Maximize Value and Flexibility

STRENGTHENING THE CORE OF PUBLIC HEALTH

Gives faster understanding of emerging health threats through electronic reporting of emergency department visits

SYNDROMIC

SURVEILLANCE

NOTIFIABLE

DISEASES

VITAL

RECORDS

ELECTRONIC CASE

REPORTING

ELECTRONIC

LABORATORY

REPORTING

Reduces burden on states for reporting notifiable diseases to CDC through modernized electronic messages

Offers earlier disease detection and intervention through automated reporting of certain diseases and conditions from electronic health records

Supports faster, more complete automated laboratory reporting of notifiable conditions to local and state health departments

Captures data from ~6 million births and deaths annually that can signal changes in trends, monitor urgent public health events, and provide faster notification of cause of death





Data to Partners

Technical and policy solutions for timely, complete, and accurate data from EHRs, labs, and other primary and new data sources to STLT partners and others in government, academics, and industry



Data to CDC and USG

Streamlined, coordinated, and interoperable public health reporting via API gateways supporting timely, complete, and accurate bi-directional data flows between STLT public health partners



Building a Public Health Workforce

Reskilling, upskilling, recruitment, and retention of a data science workforce with skills to design, implement, sustain, and innovate data modernization efforts



Ongoing Data Modernization and Innovation

Leverage state-of-the art analytics and data visualization capabilities to integrate data from new or non-traditional sources with minimal IT assistance to strengthen the detection, response, prevention, and forecasting of health threats

DMI PRIORITIES



The How:

Laying a Sustainable Foundation

DMI FOCUS AREAS







ENGAGEMENT



SUSTAINABLE IMPLEMENTATION

ELEMENTS OF SUSTAINABILITY

SUSTAINABLE



ADAPTABLE

Derived from https://nam.edu/procuring-interoperability-achieving-high-quality-connected-and-person-centered-care

THE PUBLIC HEALTH ECOSYSTEM

TOWARD COORDINATED AND SEAMLESS EXCHANGE OF DATA

Reduce Burden + Add Value

Enhance + Promote Interoperability

Put Data to Action



Roadmap

- Lays out a path from where we are now to where we need to be:
 - Activities
 - Outcomes for Short-Term, Intermediate, and Long-Term
- Presents a vision
- Guides resources
- Tracks progress



CDC Roadmap of Activities + Expected Outcomes for DMI

ACTIVITIES	SHORT-TERM OUTCOMES	INTERMEDIATE OUTCOMES	LONG-TERM OUTCOMES
If we (CDC and partners) do this	then we expect these changes to occur	which will lead to	our ultimate goals.
COORDINATE PEOPLE AND SYSTEMS			
Create interoperable systems: federal, state, local, and healthcare Coordinate investments, decisions, and policies across CDC and with partners Make data sharing easier through common policies, practices, and standards Advance academic and private partnerships	Increased collaboration, communica- tion, and messaging among CDC and partners Reduced data collection and reporting burden at state, tribal, local, and territorial levels Improved data sharing and interoper- ability through common standards like <i>HL7</i> ® <i>FHIR</i> ® Increased capacity to quickly analyze, interpret, and act on data	Effective coordination on complex health and emergency response challenges Timely and complete data reporting to CDC Efficient, secure data access and exchange between systems across the country A more comprehensive picture to improve decision-making and protect health for all	CDC can rapidly identify and effectively mitigate emerging threats Trusted data promotes evidence-based behaviors, interventions, and solutions to protect health
ACCELERATE DATA FOR ACTION			
Identify data for priority public health needs Upgrade and modernize IT infrastructure Strengthen the data science workforce Adopt open standards and tools while protecting data security Translate data into evidence-based recommendations	Increased electronic reporting and specific enhancements to flagship CDC surveillance systems Stronger workforce in data science, analytics, modeling, and informatics Targeted real-time communication of data and results	 Real-time, linked systems that recognize threats early to inform timely response A highly skilled workforce that applies state-of-the-art data skills and tools High-quality information and guidance to protect people's health 	Every American has equal opportunity to attain the highest level of health possible
SUPPORT STRATEGIC INNOVATION			All people have the right information at the right time to make decisions
Seek partner-driven data solutions Develop next-generation tools (e.g., modeling, visualization, predictive analysis, machine learning) Strengthen predictive analytics	Integration and use of data from new or non-traditional sources Improved pathways to explore, develop, and deploy next-generation technologies	Open-source, enterprise-level technologies and coordinated systems New approaches to address present and future threats	Our country is better prepared for, and protected

https://www.cdc.gov/surveillance /images/factsheets/318212-A_DMI_LogicModel_July23b330 0x2523.png

THEMATIC AREAS FOR DMI CARES ACT FUNDING



THEME 1

Data Sharing Across the Public Health Ecosystem

Getting data to all levels of public health with upgrades to core surveillance systems.



THEME 2

CDC Systems and Service Enhancements for Ongoing Data Modernization

Laying the foundation for cloud, machine learning, artificial intelligence, and advanced analytics at CDC.



THEME 3

New Standards and Approaches for Public Health Reporting Adopting new standards and approaches that will give public health agencies access to richer data.

THEME 1 SHARING DATA ACROSS THE PUBLIC HEALTH ECOSYSTEM



KEY OBJECTIVES

- Develop and implement cloud-based approaches for automating data collection and supporting multi-directional data flows among STLT partners and CDC and USG
- Reduce burden for data providers and public health agencies
- Ensure systems and services are scalable, interoperable, and adaptable to meet evolving needs

KEY PROJECTS + PROGRESS

- Modernization of critical core surveillance systems
- Moving to "program-agnostic" data sources from EHRs, Emergency Departments, and laboratories
- Rapid extension of data lakes and services that support public health reporting and immunization data

THEME 1 SELECT ACCOMPLISHMENTS

SHARING DATA ACROSS THE PUBLIC HEALTH ECOSYSTEM

- Gene Sequencing
 - Enabled over 800,000 unique SARS-CoV-2 genomes to be analyzed since August 2020 through CDC's expanded sequencing efforts
 - Led FDA to change treatment recommendations
 - Expanding platform for multiple respiratory illness surveillance

• Vaccine Response

- Over 140 million COVID-19 vaccination administration reports received, processed, and made available through CDC's Immunization Data Lake
- Average of 3.1M records per day

Lab Reporting

 Increased capacity for states to report electronic lab data: From thousands per week for all conditions to millions per week just for COVID-19

recommendationExpanding place

- Electronic Case Reports
 - Over 7100 healthcare facilities in 47 states can send COVID-19 eCR
 - Working with EHR vendors to incorporate automated case reporting and public health needs into their products

- Environmental Health Monitoring
 - 2.7M interactive maps, tables, and charts covering 525 environmental health measures are now available for sharing and embedding





Unprecedented upscaling



THEME 1 SELECT ACCOMPLISHMENTS

SHARING DATA ACROSS THE PUBLIC HEALTH ECOSYSTEM



- 49 states and 70% of all US EDs
- 75% of data received within 24hrs of visit
- 6 million messages/day all diseases
- Completeness of 93% for race

Pan-Respiratory Surveillance Initiative



- Scalable outbreak response to be prepared for next pandemic or big outbreak – flu, measles, Legionnaires' Disease
- Use of program-agnostic data sources for respiratory disease tracking
- Increased public visibility and access to respiratory disease data

- Faster Vital Stats
 - Reduced lag time on provisional death data (from months to weeks)



- Modernizing vital stats across the US
 - 67% of deaths reported now electronically within 10 days of death
- Faster overdose death data
- Expanded race and ethnicity data to inform COVID-19 response



THEME 2 CDC SYSTEMS + SERVICE ENHANCE-MENTS FOR ONGOING DATA MODERNIZATION



KEY OBJECTIVES

- Develop cloud-based, enterprise-wide infrastructure and services that redefine how CDC systems are designed, developed, interconnected, and maintained
- Enable data linking, sharing, analysis, and visualization across the agency
- Include new and non-traditional data sources to complement existing systems

KEY PROJECTS + PROGRESS

- Implementation of enterprise cloud services at CDC and modernization of legacy datasets to make data sharing faster and easier
- New cloud-based solutions for cataloging CDC's datasets, analysis of COVID-19 vaccine data, flu vaccination tracking, and API management
- Three new programs to train, hire, and grow CDC's state-of-the-art data science workforce

THEME 2 SELECT ACCOMPLISHMENTS

CDC SYSTEMS + SERVICE ENHANCEMENTS FOR ONGOING DATA MODERNIZATION



Greater Data Access

- Expanded open data with 164 new datasets added accessed by 6.1M users in data.cdc.gov in 2020 – a 32% increase from 2019
- Contributed 37 new COVID data sets into HHS Protect
 with over 1K users across Federal/State



Improved Visualization

- Over 243 unique data products developed
- Deployed a combined cloud-based dashboard showing influenza vaccination coverage for 2016-2021
- Deployed public dashboard showing confirmed COVID-19 cases for nursing homes by week, by state,• and by FEMA/HHS regions

Response Solutions

- Increased usable data to 95% for international passenger arrival contact tracing and decreased time to transmit to STLTs from days to overnight
- CDC's COVID-19 Symptom Tracker Healthbot has provided smart services to 43.1M people

Better Connectivity

- Built a cloud-based solution to ingest, analyze, and visualize COVID-19 vaccine ordering, delivery, and administrations supporting 100,000 providers
 Established 51 nodes of High Performance Clusters
- Hosting USDS Prime Data Hub and Simple Report

Enterprise Services and Data Management

- Deployed first phase of an enterprise, cloud-based data catalog and governance solution where CDC can publish details and define rights to shareable datasets
- Streamlined identity proofing and access management at CDC

Workforce

- CDC's Data Academy delivered more than 1000 hours of free training since its August 2020 launch
- CDC's Data Science Upskilling program began first cohort with 79 unique learners on 18 teams





THEME 3 NEW STANDARDS + APPROACHES FOR PUBLIC HEALTH REPORTING

KEY OBJECTIVES

- Develop and adopt new standards and approaches, including:
 - Real-world testing of new standards FHIR and standards for US Core Data Elements for Interoperability
 - Standardized APIs and electronic health record data elements
 - Use of hypothesis-driven discovery for fast-tracked implementation

KEY PROJECTS + PROGRESS

- Design & real-world testing of new FHIR-based approaches to Interoperability, including educating CDC and STLT partners about FHIR and coming changes for electronic health record (EHR) capabilities related to the 21st Century Cures Act
- Alignment and advancement of existing
 - **standards,** including ensuring public health needs were prioritized in the US Core Data Elements for Interoperability
- Assessment of policy and ethical considerations

Pandemic-Ready Interoperability Modernization Effort (PRIME): A Collaboration Between CDC and USDS





U.S. DIGITAL SERVICE

Mission Statement of PRIME





To get better, faster, complete and accurate data to state and local public health departments so that they can take appropriate timely action.

Overview of PRIME







SEEING DMI THROUGH THE DATA LIFECYCLE





Harmonizing efforts across the public health ecosystem will help us better collect, connect, track, and predict, allowing us to make more informed decisions and take more targeted action.

UNMET NEEDS AND PRIORITIES



Enhance STLT partners' capabilities to conduct effective surveillance and response Enhance CDC's capabilities to conduct effective surveillance and response, including use of innovative nontraditional sources of data and improved forecasting of public health threats



Foster an environment of continuous learning and ongoing innovation

Questions

