

Easing the Path to Data Modernization Initiative Engagement: Approaches to Address Legal and Governance Challenges

This white paper summarizes key legal and governance challenges to collaborative and agile data exchange, a critical component of data modernization, and potential approaches to address these challenges.

Background

In the wake of recent public health events such as the opioid epidemic and the COVID-19 pandemic, there has been a collective call to action to modernize public health's data infrastructure. Specifically, cumbersome and slow data exchange and information system interoperability challenges prevent the flow of needed data to the right people at the right time. Clinical, academic, and public health partners need to be able to receive and share data so that our responses to public health threats are coordinated, targeted and timely. CDC's Data Modernization Initiative (DMI), launched in 2020, is a "multi-year, multi-billion-dollar effort to modernize data across the federal and state public health landscape."ⁱ DMI takes aim at these data exchange challenges and is working to re-imagine data infrastructure and build a shared data ecosystem that breaks down barriers to agile data exchange.

A key feature of CDC's DMI is the North Star Architecture (NSA), a data ecosystem intended to offer "flexible, interoperable, and secure digital tools" that CDC and state, tribal, local and territorial public health partners (STLTs) can use. These tools will improve the flow of data between and across jurisdictions, breaking down data siloes. Although the DMI and NSA hold tremendous promise for public health partners, STLTs face several legal and governance challenges. In the past, these challenges have created roadblocks to development of data strategy, coordinated efforts to improve interoperability and agile data sharing. These challenges may prevent STLTs from fully leveraging CDC DMI resources and are most relevant to **DMI 2023-2024 Public Health Goal #4: Ensure access, exchange, and use of interoperable data and solutions**".ⁱⁱ

Legal Challenges

Complex and ambiguous legal landscape. The public health data legal landscape includes a patchwork of federal and state laws and regulations that in some instances interact smoothly and at other times create conflicting and ambiguous rules of the road for data exchange. Federal laws and regulations such as the Health Insurance Portability and Accountability Act of 1996 (HIPAA) and subsequent revisionsⁱⁱⁱ, the Health Information Technology for Economic and Clinical Health Act of 2009 (HITECH)^{iv}, 42 CFR Part 2 which requires confidentiality of substance use disorder patient records^v, the Information Blocking Rule^{vi}, and the Common Rule^{vii} address a range of issues including health information privacy, confidentiality, adoption and meaningful use of technology, and ethical treatment of human research subjects. These laws overlay similar and sometimes dissimilar state laws, raising questions of which laws apply or how might disparate laws interact and affect data sharing and technology acquisition decisions. Tribal Nations have inherent sovereign authority to govern the collection, ownership, and disclosures of their own data.^{viii} Tribal data governance approaches are nuanced and evolving, requiring hopeful data partners to have a deep respect for and understanding of tribal rules, perspectives, and practices. State legal landscapes are nuanced as well, including a constellation of public health laws related to open records, confidentiality of certain records and information, data protection and consent requirements, associated regulations and administrative rules.

State public health laws that relate to health data protection vary greatly in scope, clarity, and alignment with DMI priorities and FAIR (findability, accessibility, interoperability, and reusability) principles. Some state laws are general while others prohibit disclosures of specific data elements, datasets, or information systems; for example, individually identifiable information related to sexually transmitted diseases, emergency medical services run data, cause of death from vital statistics death files, name of employer from hospital discharge data, and certain Prescription Drug Monitoring Program (PDMP) data. State data privacy laws are often outdated, silent on key disclosure questions, inconsistent with other applicable law or overly broad, creating both roadblocks and grey areas for data sharing. How states and their legal teams apply and operationalize state laws relating to data management also varies from state to state, making it difficult to develop streamlined and uniform approaches to data sharing.

Lack of clear standards for data de-identification. Data protection laws generally prohibit disclosures of individually identifiable data and then provide, albeit not in all cases, exceptions to the general nondisclosure rule. In the current legal landscape, data sharing can more easily be accomplished when data are deidentified. Federal and state law provide some guidance on how to properly deidentify data. Federal laws such as HIPAA^{ix} and Family Education Rights and Privacy Act^x (FERPA) include definitions and standards for deidentification for protected health information and educational information and records.

Occasionally, but not often, state laws provide a standard or preferred method of data deidentification. Public health practitioners often refer to generally accepted statistical methods for guidance on data deidentification. These methods, legal and statistical, vary in their effectiveness and applicability across data sets. Deidentification can be particularly tricky with integrated and detailed data sets with a high number of discrete data elements, such as hospital discharge data and PDMP data sets.

The HIPAA Privacy Rule provides two de-identification standards for protected health information.^{xi} These standards, the Safe Harbor Method and Expert Determination Method, have been adopted as policy by some covered and non-covered jurisdictions. Each standard, even the seemingly straightforward Safe Harbor Method, requires a professional judgement call on the risk of re-identification. The need for awareness and compliance with legal and statistical standards may challenge the resources and expertise of STLs, particularly the smaller ones who benefit most from strong data dissemination capabilities.

High Burden Data Sharing Agreement Process. Data Sharing Agreements (DSAs) are formal contracts that describe what data are being shared, appropriate uses of data, and expectations for secure and compliant management of data. By ensuring there is a mutual agreement and understanding of data use, DSAs help to develop trust among parties. DSAs require front-end collaboration between parties and result in careful documentation of agreed upon terms.

The downside to data sharing agreements is that the road from collaboration to a final and approved agreement can be long, redundant, and cumbersome, slowing the flow of data to and between public health partners. Adoption of best practices for getting to yes on agreements, streamlined processes for development of data use agreements, and innovations in automated user authentication and access hold promise for reducing the time and administrative burdens required to implement DSAs.

Governance Challenges: People, Policies, Processes and Technology

Data governance is the “exercise of authority, control, and shared decision-making (planning, monitoring, and enforcement) over the management of data assets.”^{xii} CDC’s DMI and STLT engagement with DMI resources will require data governance at both levels to achieve effective shared governance as DMI partners. Data governance occurs within each jurisdiction, often informally, in siloes and with varying operating frameworks and processes. Despite how data governance operates within each STLT, common challenges persist across three key organizational domains: people and culture, policies and processes and technology.

People and Culture. Data governance is about people – the squad of people creating, managing, analyzing, integrating, and disseminating data to advance science and improve population health. The people and culture of an organization determine the extent to which data is seen as a valued asset and whether the organization has critical mass to create a common vision for data management. Common roadblocks in the people and culture domain include lack of leadership buy-in or bandwidth, reluctance to share data, ambiguity around specific data roles and responsibilities including perceptions about the meaning of “data ownership” vs. “data stewardship” and missed opportunities for collaboration across siloes. At the heart of many of these challenges is the need for trust between data exchange partners. A critical pre-requisite for STLT’s data modernization engagement and success is confidence that data governance is compliant, informed by local data stewards and subject matter experts, and upholds the mutual expectations of data exchange partners.

Organizational Policies and Processes. Modernization of data infrastructure, at the federal and local levels, will require policies and processes that authorize and support these aims. Policies and processes that lag science and the imminent need for accurate, real-time create roadblocks for agile and timely data exchange. Policy and process gaps, particularly for STLTs, may include lack of data governance strategy and operating framework, absence of comprehensive and transparent data use and sharing policy, missed opportunities to streamline and automate data lifecycle workflows, high burden and time consuming data sharing agreement processes, and lack of detailed and transparent documentation of data assets.

Technology. Many STLTs, through past and current funding vehicles, have acquired new technical tools or resources to support improved data management and data exchange. However, STLTs still face significant challenges including interoperability across surveillance information systems, ambiguous rules of the road for adoption and onboarding of new technology, workforce skills and bandwidth gaps, misalignment between business needs and IT policies, and uncertainties related to sustainability of dedicated funding to support public health data infrastructure. STLT ability to leverage shared technology is critical to advancing improvements in data exchange.

Relevant RTI Service Offerings

The sections above identify common challenges to agile and meaningful data exchange between public health partners. RTI is prepared to offer client-specific service data management solutions which leverage our experience working within state government, our understanding of cultural, operational and technical levers that promote agile data exchange, and our strong team of subject matter experts. These solutions align most closely with **DMI 2023-2024 Public Health Goal #4: Ensure access, exchange, and use of interoperable data and solutions.** As next steps, we suggest consideration of the following service offerings:

Data Landscape Assessment: Assessment of governance structures and data sharing policies and practices will provide a deeper understanding of governance and business challenges that impact STLT DMI engagement and participation and inform efforts to manage change and governance.

Data Asset Inventory: Assessment of public health information systems and data sets will support organization-level views of data assets, promote deeper understanding of interoperability challenges, inform efforts to close technology and interoperability gaps.

Data Sharing Tools: Facilitated co-development of administrative data sharing tools, e.g. data sharing and data access methods and policies, model data sharing agreements, and data strategy roadmaps, will create shared low-burden resources designed to improve data sharing and associated administrative processes.

We believe these approaches, led by capable and nimble RTI teams, will strengthen public health partner capacity to navigate common governance challenges related to DMI engagement and implementation.

Selected RTI Capabilities

RTI is an independent nonprofit research institute dedicated to improving the human condition. Our partners rely on us to deliver high-quality data and associated systems managed through all project lifecycles. To achieve this, we apply modern technologies, best practices and innovative approaches to collect, manage, analyze, and disseminate data effectively within a secure governance framework.

We lead data transformation as part of a larger community working toward a modular national ecosystem that allows participants to connect, collaborate, share data, and create insights that inform decisions. Our expertise and capabilities in the data modernization space include data management, advanced analytics, infrastructure and security, software, systems and maintenance and operations.

Examples of our work in this space include the NIH NHLBI BioData Catalyst® (BDC)^{xiii,xiv} and the NIH Helping End Addiction Long-Term InitiativeSM (HEAL). RTI co-coordinates the BDC, a community-driven ecosystem, where researchers find, access, share, store, and compute on large-scale datasets. Secure, cloud-based workspaces, search, tools and workflows, applications, and innovative features address community needs, including exploratory data analysis, genomic and imaging tools, tools for reproducibility, and adoption of open APIs to enable data exchange. As the HEAL Data Stewardship Group, RTI provides solutions for managing and coordinating diverse data across the 1000+ studies in the NIH HEAL Initiative. We collaborated with HEAL to define the data strategy and governance plan, developed a sustainability plan, and provided materials to upskill HEAL contributors and support translational science by researchers, healthcare providers, community leaders, and policymakers^{xv}. RTI's approach to technical transformation can inform CDC's efforts to modernize public health infrastructure.

Disclaimer

The information contained in this paper is for background and general informational purposes only and does not, and is not intended to, constitute legal advice. Readers of this white paper should contact their attorney to obtain advice with respect to any legal matter.

-
- ⁱ Data Modernization Initiative (DMI) Basics. (2023, January 5). Retrieved from <https://www.cdc.gov/surveillance/data-modernization/basics/index.html>.
- ⁱⁱ Layden, Jennifer. (2023, February 27th). *Office of Public Health Data, Surveillance, and Technology (OPHDST) Strategic Priorities and Needs from Industry*. [Conference presentation]. CDC-ONC Industry Days 2023, Washington, DC, United States.
- ⁱⁱⁱ Health Insurance Portability and Accountability Act. Pub. L. No. 104-191, § 264, 110 Stat.1936.
- ^{iv} Title XIII of Division A and Title IV of Division B of the [American Recovery and Reinvestment Act of 2009 \(ARRA\)](#), Pub. L. No. 111-5, 123 Stat. 226 (Feb. 17, 2009) ([full-text](#)), *codified at* 42 U.S.C. §§300jj *et seq.*; §§17901 *et seq.*
- ^v 42 USC §290dd-2, 42 CFR Part 2 (2020)
- ^{vi} CFR 45 § 171.103 (2020)
- ^{vii} CFR 45 § 46 (2018)
- ^{viii} NCAI Policy Research Center (2018, October). The State of Tribal Data Capacity in Indian Country: Key Findings from the Survey of Tribal Data Practices. Washington, D.C.: National Congress of American Indians. Retrieved from https://www.ncai.org/policy-research-center/research-data/prc-publications/Tribal_Data_Capacity_Survey_FINAL_10_2018.pdf
- ^{ix} US Department of Health and Human Services. (2022, October 25). Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule. Retrieved from <https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html>.
- ^x US Department of Education. Protecting Student Privacy. Retrieved from <https://studentprivacy.ed.gov/node/548>.
- ^{xii} Earley S. (2017). *Dama-dmbok : data management body of knowledge* (Second edition first printing). Technics Publications.
- ^{xiii} National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health Human Services. The NHLBI BioData Catalyst. 2020. doi:10.5281/zenodo.3822858
- ^{xiv} Manning AK, Avillach P, Boyles RR. NHLBI biodata catalyst and the future of cloud computing. Genetic. 2021. Available: <https://pesquisa.bvsalud.org/global-literature-on-novel-coronavirus-2019-ncov/resource/pt/covidwho-1458029>
- ^{xv} Collins FS, Koroshetz WJ, Volkow ND. Helping to End Addiction Over the Long-term: The Research Plan for the NIH HEAL Initiative. JAMA. 2018;320: 129–130.