National Network of Public Health Institutes

Data Governance for Children's Mental Health Surveillance: What is It and Why Does It Matter?

To support children's mental health, we first need a solid foundation of information based on data.

The power of data to serve this purpose is multifaceted. Data can be used to:

- Paint a picture – What is the state of the state now or at any other point in time?
- Look at trends to see what has happened over time – Is the change systemic or random? Are there outliers and, if so, where?
- Prepare for the future – Can we predict and prepare needed resources? Can we improve outcomes?

But to have faith in what the data say, there must be confidence that the data are valid (i.e., measure what they purport to measure) and reliable (i.e., the findings or information are repeatable), collected with fidelity and protected following protocol, and used appropriately. Data governance is the means to ensure confidence in the data and in the information that comes from analyzing current data. Currently, no systemic and comprehensive surveillance system on children's mental health exists, so work is needed to create data governance tools to move toward that goal.

Children's mental health (CMH) data could come from a variety of sources, including health, human services and education sectors. As a result of the diverse data sources and the variety of assessments and indicators across a child's life span, potential challenges and threats to data validity, reliability, and quality could emerge. Each state or local agency or service provider that collects, stores, and analyzes data should have its own data governance policies and procedures to oversee data standards, security, privacy, access and use. When different agencies, state and/or local, plan to share and combine data, it is essential that cross-sector or interagency governance and communication occurs too. The mission of interagency data governance is to ensure that the highest quality data are used and made available to key stakeholders through coordinated efforts across organizations for the purpose of providing critical information to policymakers, educators, state and local agencies, service providers and the general public.

This report is first in a series intended to help users of educational, health access, children & family, mental health, and health data understand how to communicate and share data collaboratively with the ultimate goal of coordinating children's mental health surveillance. The report describes developing data governance structures, activities, and data standards to engage the right people in the right ways at the right time to improve long-term outcomes in children's lives.

To be clear, data governance itself is not a product, deliverable or program that a state agency is mandated or asked to produce. Data governance provides the mechanism to oversee and implement in a coordinated way the data-related policies and practices that are used to manage, monitor or evaluate services or programs, such as home visiting services or special education services. If a state wants to create a CMH tool that uses data from multiple agencies or establish a new CMH data collection within the health department, the affected agency/agencies would either tap into an existing data governance program or establish a new one to create CMH-specific data-related policies and processes.

For example, the state of Washington established and funded the Education Research & Data Center (www.erdc.wa.gov) to compile data about students as they move through school to the workforce. The data are transformed into insights that inform policymaker, parent and educator decision-making. As a part of establishing the ERDC, the Center created a data governance program with representatives from key agencies and organizations to oversee the data-related policies and processes. Minnesota created the Early Childhood Longitudinal Data System (www.eclds.mn.gov) to gain insight into children's development and learning. The ECLDS uses data from the departments of human services and education, as well as the office of higher education, to generate useful reports and metrics. As with Washington, data governance committees were established to oversee data sharing and research requests.

What is Data Governance?

A strong data governance program is specifically designed to provide data oversight that ensures confidentiality, integrity, and availability of the data by reducing data security risks due to unauthorized access or misuse of the data. A strong data governance program also provides transparency into how the data are generated, managed, and consumed. Data governance helps ensures that data are reliable, valid, complete, timely, available to those with a legitimate need for and authority to access. Coordinated data governance also provides the opportunity to decrease data collection redundancies, standardize data-related processes and systems, and increase data system and resource efficiencies within and across agencies. Figure 1 displays 10 essential elements of data governance.

Figure 1. **10 Essential Elements of Data Governance**

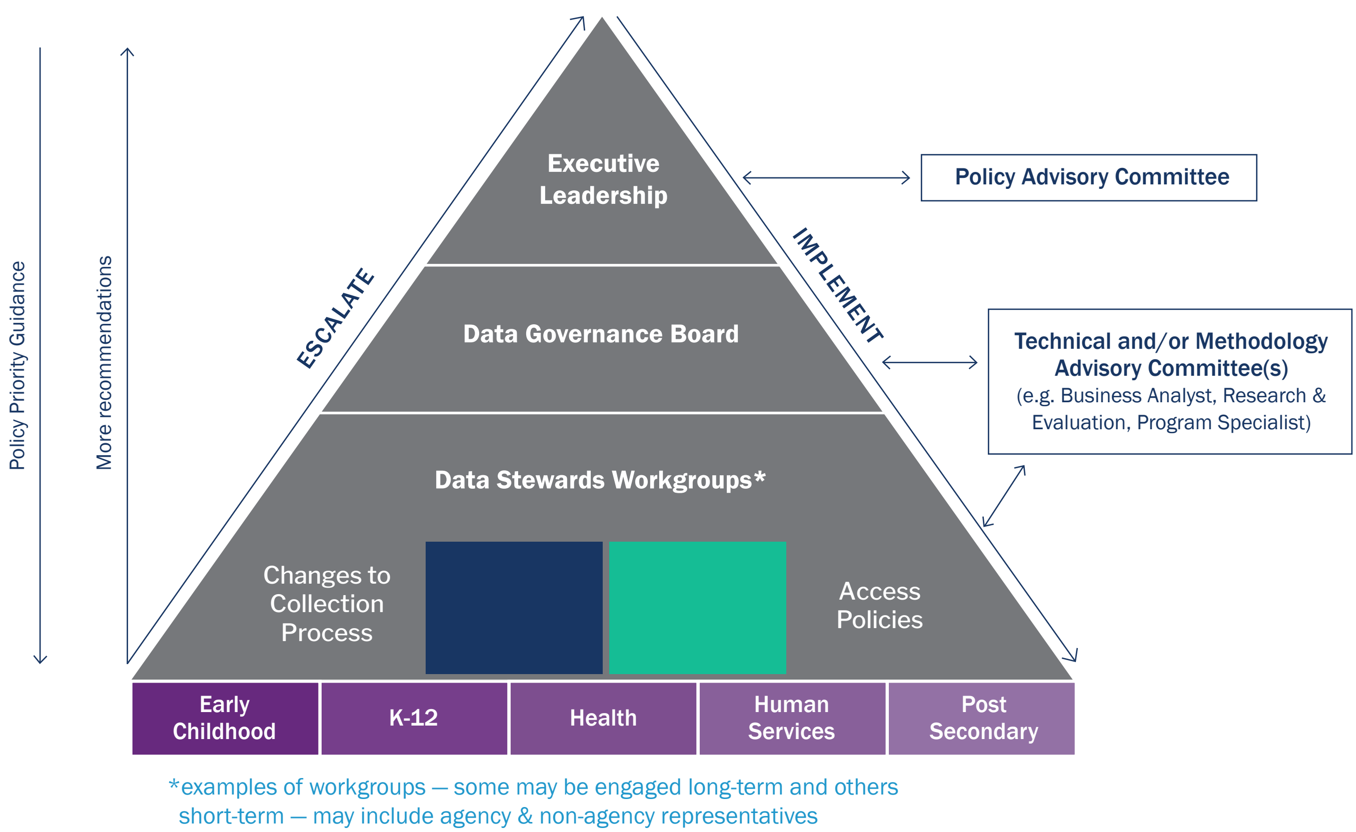
Processes	• Cross-functional decision-making hierarchy • Data policies aligned to organizational goals • Ongoing oversight, change management and assurance reviews
People	• Executive sponsorship - organizational commitment and participation • Data stewards - authority & responsibility to define the meaning, business rules and use • Field participation - engagement of district, school, and program staff • Data Governance Coordinator - oversee governance and metadata management, conduit between data stweards and IT
Data & Technology	• Standards - definition, names, code values, format • Collections - tools and procedures • Privacy, access, and use - policies and applications

Who Should be Engaged in Data Governance?

Successful data governance involves the vision, leadership and cooperation of people at all levels of implementation: leadership, project managers, program staff, research, Information Technology (IT) and subject matter experts (SMEs). SMEs can represent a variety of perspectives engaged in the data system, including children's mental health program staff who bring content knowledge, IT database administrators, and research analysts. SMEs could be engaged through specific workgroups or advisory committees to address topics such as, but not limited to, operational and technical issues, data quality standards, research priorities and processes, and security protocols. SMEs can also help ensure regulatory compliance of data access, use and reporting.

One approach to data governance programs is to create a set of committees responsible for varying levels of detail and authority, so that each committee only focuses on issues within their purview, as shown in Figure 2 below. This approach allows each committee to focus on their areas of expertise, e.g., high-level policy versus detailed-level implementation decision-making.

Figure 2. **Interagency Data Governance Hierarchical Structure**



In this approach, each committee would engage a specific type of staff and execute particular responsibilities commensurate with their roles and responsibilities within their organizations as described below. The committees should include representatives from each participating organization. The graphic above displays a structure that spans early childhood through postsecondary education and includes health and human services. Many states have established a statewide longitudinal data system that engages preschool, K-12, postsecondary education and workforce agencies. Ideally, each organization participating in an interagency data governance program also has a similar internal governance structure that guides its own data system.

Robust data governance programs, especially interagency programs, require a position such as a data governance coordinator or a Program Management Office to provide coordination, documentation and communication services. Table 3 in the Appendix describes in more detail possible committee membership and responsibilities for an interagency data governance program.

- The **Executive Leadership** team, comprised of the senior executives from each partner organization, sets the overall mission and strategic goals and crafts policy for the data sharing and analysis program and for its governance. It also obtains needed funding and resources and maintains final authority and responsibility for all activities.
- The **Data Governance Board** is comprised of project/program managers (e.g., early intervention, special education services), research, information technology staff, and various subject matter experts (SMEs) from each partner organization and the data governance coordinator. Much of the design, planning and implementation of the program could be accomplished through topic-specific workgroups that rely heavily on input from partner SMEs and project managers, with input from advisory committees as needed.
    - The **Data Governance Board** reviews and approves the high-level task plan, processes and procedures produced by workgroups and/or the advisory committees as necessary to achieve the strategic goals outlined by the executive leadership team.
    - **Data Steward Workgroups and Advisory Committees** generally focus on specific technology, research or legal topics and are comprised of subject matter experts and the representatives from each partner organization who review and make recommendations about logistical issues and operating procedures that guide the implementation activities. External stakeholders are engaged in these groups.
    - The **Data Governance Coordinator** provides dedicated support for day-to-day operations, coordinates governance activities and provides support to the governance bodies.

Types of Data Governance Scope and Activities

The scope and goals of data governance activities differ for data management, project management and overall program coordination, though it encompasses all three, as outlined below:

> **Data Management** addresses issues such as data quality, data standards, common vocabulary, and data matching standards for cross-agency data alignment. It supports processes to more easily integrate, synchronize, and consolidate data across different programs and organizations.

> **Project Management** provides a framework for decision-making around specific projects within a larger program. Projects have specific start and end dates and are focused on established and agreed upon scope, outcomes, and deliverables that are to be completed on time and on budget and include activities such as conducting analyses and producing reports related to a specified policy question.

Program Coordination provides a structure and framework for goal setting, strategic planning, and decision-making for a program. The overarching governance plan identifies key roles and responsibilities for each organization and the people involved in the program. It identifies the key stakeholders involved in program management and the individuals authorized to approve program activities and priorities.

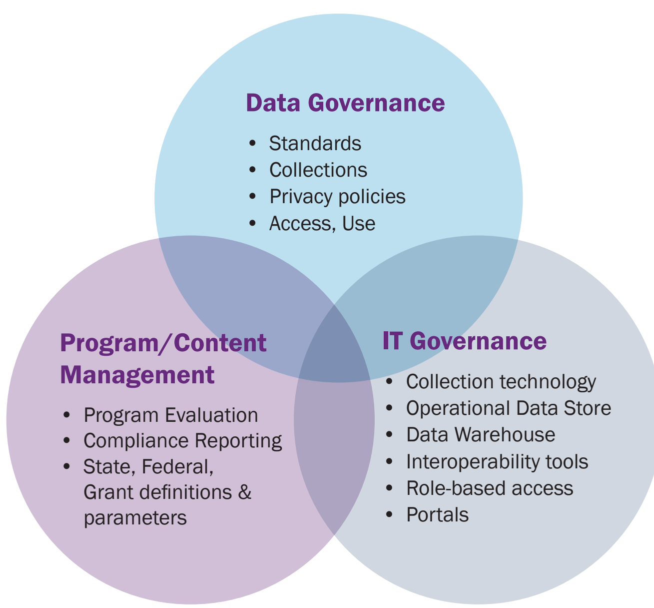
Data governance integrates a wide variety of activities across the various committees. In general, the program addresses, but is not limited to, standard operating procedures, process management, data-related business rules, data standards, documentation, communication, and data/research request review and approval processes.

Which Departments are Involved in Data Governance?

Data governance should not be considered as solely a function under Information Technology. In fact, data governance is distinguished from Information Technology (IT) governance and from program/content management, although it should be guided and informed by all three perspectives (see Figure 3). As described above a comprehensive data governance program will include representatives from program areas, policy, research, and IT. These groups will work together to best determine what data to collect, how, when and to manage the storage, privacy and access processes. In general, data governance addresses data-related policies and procedures, while IT governance addresses decisions about the technology infrastructure, architecture, hardware and software that best meet the agency's or program's needs.

IT governance principles serves a resource for collecting, managing, protecting and sharing data that is required through state and federal law, policy requirements or for research and evaluation needs, but it is not responsible for deciding what data to collect and when. Program staff (e.g., early intervention, special education) are responsible for making sure that they have the data that is mandated or needed to manage, monitor and evaluate programs, but they are not typically SMEs in the state-of-the-art technology solutions. The data governance program can bring together representatives from each group, along with research and evaluation, to determine the best enterprise-wide solutions for the data system, with the goal of minimizing data redundancy and maximizing useful and actionable information.

Figure 3. **Relationship between Data Governance, IT Governance and Program Management**



What are the Data Sources?

Data sharing across state agencies is complicated by the fact that each state agency is guided by different federal agencies and laws, as well as by sector-specific state laws. What state agencies collect about clients, or students in the case of education, how the data are defined and then aggregated for reporting differ even when collecting similar information for the same children, often because of the nuances in federal or state legislation. Federal agencies have been trying to support and coordinate interagency data sharing efforts over the last 10 years, and they have provided financial and program incentives to states to create state-level interagency data sharing, but legislatively mandated collections take time to change.

As stated previously, there ought to be a rationale for collecting data within agencies and sharing data across agencies with a specific intended use. As state departments of education have built student-level data systems over the past 15 years, many states have put the onus on the state education agency to ensure that they do not collect any data elements that have not been mandated in state or federal law. School districts, however, typically collect more information than is shared with the state agency. For example, school districts maintain transportation, health, library, food services, and athletic data, among others, that is not shared with the state. While each state has developed its own data system, data collection process, and data documentation, the documentation processes and data standards are difficult to find or understand in some states. Federal law allows states to establish their own definition and calculation of common performance indicators, such as graduation and dropout rates. In fact, states set their own graduation requirements.

By the same token, health and human services agencies and programs also vary within and across states in terms of what they collect, how and when. The data may come directly from service providers or from state and federally sponsored programs, and local programs likely have much more data on individuals than the state agencies.

Data Standards and Mapping

When sharing data across state agencies, good documentation about each data collection and the data standards (e.g., the data dictionary that includes data element definitions, code sets, level of aggregation, etc) can help to make sure that each variable is matched, aggregated and used properly. Different agencies may have identical outcome indicators given their distinct and separate purposes. However, if analysts need to match records for individual children across programs and agencies in order to look at long-term outcomes, they will have to match on person-specific data elements to ensure proper linkage across datasets. If the state does not have a unique person identifier across agencies, then the matching will likely be done by each individual's first name, middle name or initial, last name, date of birth, gender and race/ethnicity. The table below demonstrates how disparate the race/ethnicity codes can be across early childhood and education programs and highlights the need for analysts to plan for the time and resources necessary for data cleansing and element matching before conducting analyses.

Table 1. **Data Standards in a New England State for Race/Ethnicity Across Collections in Human Services and Education**

Bright Futures Information System		Children's Integrated Services		Agency of Education	
Data Element	Code Set	Name	Type/Code Set	Field Name	Code Set
Individual Ethnicity	Hispanic	**Ethnicity**	0-missing; 1=American Indian or Alaskan Native; 2=Asian; 3=Black or African American; 4=Hispanic or Latino; 5-White (not Hispanic); 6-Native Hawaiian or Other Pacific Islander; 7=2 or more races	**Ethnicity**	1/2; 1=HISPANIC/ LATINO; 2=NOT HISP/LATINO
Individual Race	American Indian or Alaskan Native, Asian, Black, Pacific Hawaiian, White			Race - White	
				Race - Black or African American	
				Race - American Indian or Alaskan Native	
				Race - Asian	
				Race - Native Hawaiian or other Pacific Islander	

Data Element Definitions

When using data from multiple sources, care should be taken to understand what each data element represents. It is not safe to assume that data elements with the same name are measuring the same concept. For example, when combining datasets across K-12 and postsecondary sectors, a key concern in both the K-12 and postsecondary environments is student retention, so one could assume that a dataset from each would include one or more data elements about retention status. However, the definition of retention in K-12 usually means that a student is repeating a grade level in a subsequent year and has a negative connotation, while retention in the postsecondary arena indicates that a student has remained enrolled at an institution from one year to the next and has a positive connotation. It would be problematic to take the information in a 'Retention' data element and compare the data elements across sectors as though the meanings and code sets are equivalent.

For CMH data, indicators across programs might refer to behavioral problems. In K-12 datasets, these are often coded as instances of 'discipline' problems (e.g., fighting, self-harm); however, 'discipline' in postsecondary institutions refers to a student's area of study.

Children's Mental Health Assessment

There are a variety of ways to assess and document mental health indicators, such as questionnaires completed by parents, teachers or the children themselves, as well as academic indicators, such as attendance, behavior and participation in special education services. Many assessments are completed by private providers and are never entered into a health or human services data system. Some school districts administer short assessments for all students (e.g., Behavior Assessment System for Children©), but that data are not shared with the state education agency. If assessment data are shared with the state, they may be only a scale score or aggregate score, not at the item level for each child. All of this is to say that states may have some of indicators about children's mental health in state agencies, but they are likely to be limited in scope and quantity.

As indicated earlier, if CMH data are included in a state agency's data system, it is critical to document the data standards, particularly the data element name, definition, code set and format, which will provide valuable guidance to how to effectively use the data from various assessments and data collections. For example, common children's mental health questionnaires assess the existence and/or degree of anxiety, depression, attention problems, aggressive behavior, or social problems among other mental health indicators. A review of common children's mental health assessment tools demonstrates how diverse the data standards are across instruments assessing similar issues. For example, Table 2 shows the disparate types of coding used in children's mental health tools. If the data for each of the items in these instruments are translated to numeric coding, the data from different sources may look alike but have vastly disparate meanings. For example, items measuring anxiety on one assessment may be aggregated to represent a scale score of 16 (to represent 16 yeses out of 17 questions), while another anxiety measurement may be an index representing 16 (out of 48 yeses). Should the two scale scores of 16 be analyzed as though they have the same value?

Local service providers and mental health professionals have a plethora of assessment tools to choose from when assessing children's mental health issues, unless a state agency mandates a specific instrument. Program specialists within the state agency will likely determine the best children's mental health instrument(s) to include as part of its data collection system or to collect it via a statewide survey, but keeping IT and data governance program up-to-date with those decisions to ensure appropriate data standards documentation and use of the data in analyses.

Table 2. **Disparate Data Standards across Measures of Children's Mental Health**

Response Type	Common Code Set Options			
Yes/No	Yes/No	Y/N	1/0	
True/False	True/False	T/F	1/0	
3-point Likert	Never, Sometimes, Often	N/S/O	0/1/2	1/2/3
5-point Likert	Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree	SA/A/N/D/SD	5/4/3/2/1	2/1/0/-1/-2

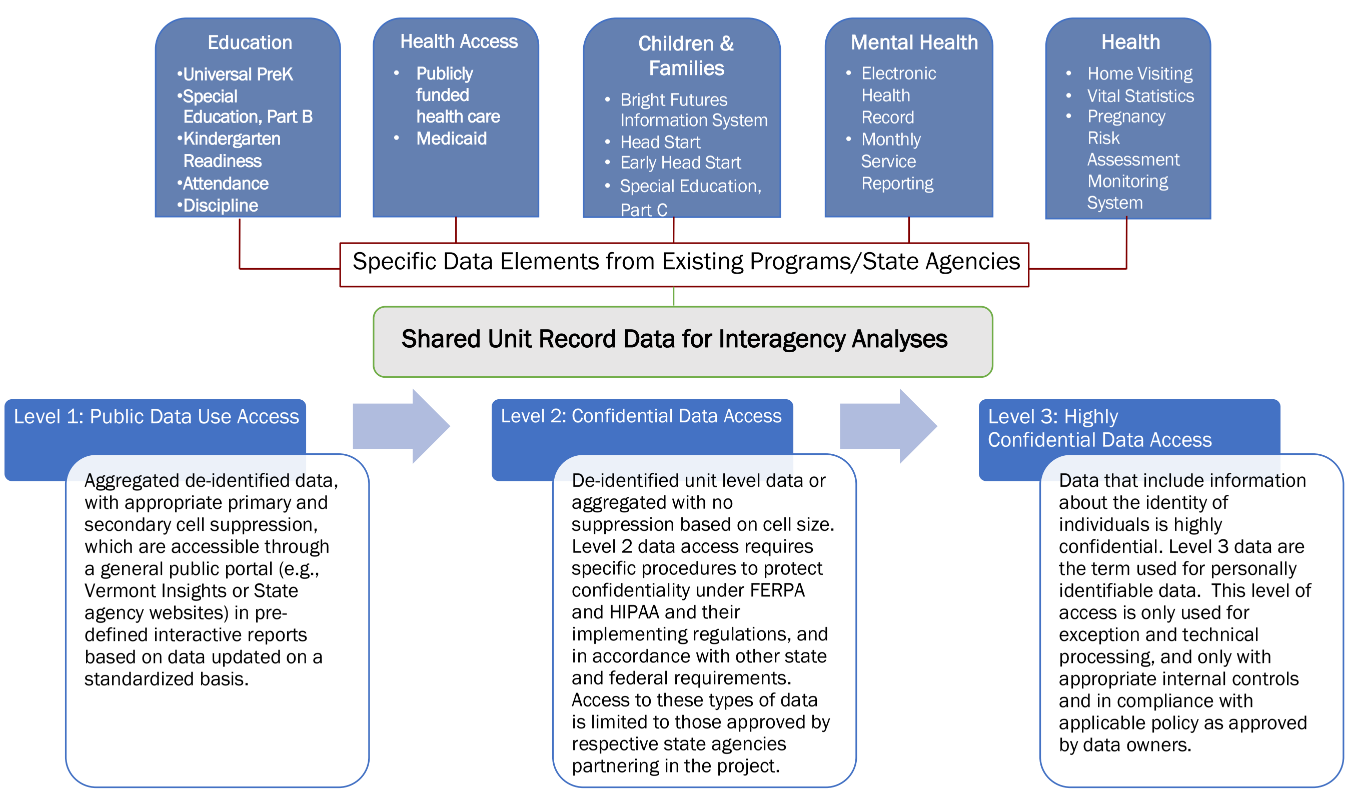
Multiagency Data Coordination and Privacy Protection

Numerous states, with funding from either the U.S. Department of Education Statewide Longitudinal Data System (SLDS) grant program and/or the Race to the Top (RTT) Early Learning Challenge (ELC) grant program, have built interagency data systems, either integrated (i.e., a copy of key data elements from multiple agencies are linked and stored in a separate data warehouse) or federated (i.e., copies of data elements from multiple agencies are only linked for each specific project and not stored separately) that facilitate the sharing of data to better to study the influences of early childhood and educational experiences on long-term education and workforce outcomes. Generally, states that establish these interagency data systems also establish privacy protocols or tiered access levels to ensure researchers and other stakeholders do not get access to data that reveals personal information.

Protecting individual's privacy is of paramount importance. It is especially important to protect the privacy and confidentiality of children's mental health and education data (e.g., participation in special education services, grades, and test scores) in order to protect children from abuses and biases by schools, institutions, government agencies and the public. Many families fear the information included in children's records could adversely impact opportunities in later years if the data became known. The Family Educational Rights and Privacy Act (FERPA, 20 U.S.C §1232g; 34 CFR Part 99) and the Health Insurance Portability and Accountability Act (HIPAA, Pub.L. 104-191. Stat. 1936) strive to ensure that individual's data remains private and confidential, while still allowing the use of data in research.

The New England state used grant funds to establish the federated Prenatal-Grade 12 Data Governance Program to facilitate data management and sharing between the education and human services (which includes the department of health) agencies. The state also established a set of operating principles for the governance program, including specific guidelines about data privacy and access. Figure 4 shows a sample of the types of data sources that are available for interagency studies and describes the three levels of access that guide the degree of granularity and personally-identifiable data included in a given research dataset.

Figure 4. **Federated Data System Data Sources and Privacy Protections**



Summary

This report is the first in a series of publications focusing on data governance around children's mental health data, with the goal of improving data systems which understand and track how children are growing and developing. Data governance processes provide the coordination and oversight within and across agencies necessary to ensure valid, reliable and high-quality data are available for public health, human services and education research, program evaluation and policymaker decision-making. As such, governance activities include data management processes, data standards and definitions, and program coordination that oversees interagency data sharing and analyses. Effective data governance includes engaging the right people in the right ways (i.e., commensurate with their skill and authority levels) at the right time to facilitate the collection to serve children's needs where they live, learn, and play.

National Network of Public Health Institutes

Funding for this publication has been provided to the National Network of Public Health Institutes (NNPHI) through a Cooperative Agreement with the Centers for Disease Control and Prevention (CDC-6-NU38OT000203-05). NNPHI collaborated with DataSmith Solutions, LLC, and the CDC's National Center for Birth Defects and Developmental Disabilities on this project. Contents are solely the responsibility of the author and do not necessarily represent the official views of the CDC, the US Department of Health & Human Services, and NNPHI.

Appendix

Table 3. **Possible Interagency State Governance Committee Membership and Responsibilities**

Response Type	Common Code Set Options	Types of Responsibilities
Executive Leadership	• Agency Secretaries or Commissioners • Deputy Secretaries or Commissioners • Agency Chief Information Officer	• Set overall mission and strategic goals • Secure funding, resources, and cooperation to support the data governance effort • Approve/edit/deny data governance recommendations or solicit more information • Update Governor, Legislature and/or public • Provide direction to data governance board
Data Governance Board (DGB)	• Data Governance Coordinator From each agency: • Research Director and/or Analyst • Business Architect and/or Information Enterprise Architect • Chief Data Officer • Program Director or Manager	• Implement policies of the agency leadership • Manage the scope and activities of the program • Develop and implement processes and procedures • Review possible projects and solicit input from Data Stewards Workgroups and/or Advisory Committees • Accept/edit/deny recommendations from Workgroups and/or Advisory Committees or solicit more information • Communicate with internal and external stakeholders. • Identify data stewards to participate on workgroups • Submit proposed plans of action, procedures and processes to leadership • Oversee scope of work of workgroups to implement approved changes
Data Stewards Workgroups	• Agency and non-agency subject matter experts • Program or division representatives who manage specific agency data collections, analyses or IT processes • Representative from local or regional agencies, school districts and/or non-profits may be engaged for their subject matter expertise	• Serve on short- or long-term workgroups as needed. Review project parameters and identify key data elements to share • Work with DGB and/or Advisory Committees to evaluate issue, possible solutions, cost/resource effectiveness, recommendations and timeline • Advise the DGB of recommended solutions • Resolve technical issues • Review and inform data use and access policies • Develop and recommend privacy and security policies and procedures • Implement approved policies and plans

Response Type	Common Code Set Options	Types of Responsibilities
Advisory Committees†	• State Advisory Council • Policy advisory committee comprised of executive leaders in local or regional organizations (e.g., school superintendents, non-profit service providers) • Technical or methodological advisory committee from local or regional organizations (e.g., school or district IT staff, local or regional service provider staff responsibility for data collection or management)	• Represent state entities, local providers, regional councils, non-profits and philanthropy • Liaison with DGB and agency leadership • Provide review, feedback and insight about legislative, programmatic or data-related issues • Support public information efforts of data governance program • May provide perspective on prioritization of reporting and research activities
Data Governance Coordinator	• Designated representative, usually from a participating agency	• Provide functional and organizational infrastructure support • Make decisions as necessary to fulfill the program's mission • Serve as liaison to executive leadership • Coordinate program communications

*Membership is usually drawn from participating agencies or organization, although external stakeholders can be engaged via advisory councils and workgroups.

†Advisory Council members or similar representatives from other organizations may be asked to participate in work groups to address specific topics, but it is also helpful to have a standing advisory committee to use to vet issues and possible solutions