



P-20W+ Data Governance
Tips from the States

Brief 4

Updated February 2017

best practices brief



Statewide
Longitudinal
Data Systems
Grant Program



SLDS Best Practices Brief

P-20W+ Data Governance

Brief 4

Updated February 2017

In general, data governance refers to the overall management of the availability, usability, integrity, quality, and security of data. A sound data governance program includes a governing body or council, a definition and allocation of authority, a defined set of procedures, and a plan to execute those procedures. This brief offers best practices for establishing and maintaining a P-20W+ data governance structure. It includes examples from states with experience setting up P-20W+ data governance.

Data governance is both an organizational process and a structure. It establishes responsibility for data, and it organizes program area and agency staff to collaboratively and continuously improve data quality and usability and to ensure data security and confidentiality through the systematic creation and enforcement of policies and procedures. Data governance also ensures efficient and effective collaboration, authority, and accountability by defining and assigning clear roles and responsibilities.

This document relates to P-20W+ or interagency data governance rather than K12 or intra-agency data governance. While there are many similarities in structure and process between interagency and intra-agency data governance, there are key differences. For example, among the various P-20W+ agencies, there may be varying security requirements, processes for data sharing, data use and reporting priorities, and timelines. There also may be a different, broader research agenda at the P-20W+ level, whereas intra-agency priorities tend to focus on the education programs overseen by that specific agency. (See figures 1 and 2 on pages 2 and 3, respectively, for depictions of single-agency versus P-20W+ data governance structures.)

When data governance is effectively established, the quality of data collected, reported, and used by state and local education agencies—as well as early childhood, postsecondary, and other agencies such as the Department of Labor or Department of Health—is enhanced; staff burden is reduced; and communication, collaboration, and relationships with the various agencies, information technology (IT) staff, and program areas are improved.

P-20W+ refers to data from prekindergarten (early childhood), K12, and postsecondary through post-graduate education, along with workforce and other outcomes data (e.g., public assistance and corrections data).

The specific agencies and other organizations that participate in a P-20W+ initiative vary from state to state.



This product of the Institute of Education Sciences (IES) SLDS Grant Program was developed with the help of knowledgeable staff from state education agencies. The views expressed do not necessarily represent those of the IES SLDS Grant Program. We thank the following people for their valuable contributions:

Melissa Beard
*Washington State Education Research
& Data Center*

Jan Kiehne
*Connecticut State Colleges
and Universities*

Tod Massa
*State Council of Higher Education
for Virginia*

Ron Parrish
Virginia Department of Education

Kirsten Rowe
*Virginia Department for Aging and
Rehabilitative Services*

Jeff Sellers
SLDS Program, State Support Team

For more information on the IES SLDS Grant Program, additional Best Practices Briefs, or for support with system development or use, please visit <http://nces.ed.gov/programs/SLDS>.

Initial Steps to Establishing P-20W+ Data Governance

1. P-20W+ partners identify education policies, priorities, challenges, and needs that span multiple agencies.
2. Partners develop the initial draft of the P-20W+ data governance policy, using the P-20W+ education policies as a foundation and the initial step for governance. Authority for P-20W+ data governance should be granted via executive order, state statute, or as part of memoranda of understanding (MOUs). Either a single MOU for all partners or a separate MOU for each participating agency can serve as the administrative vehicle for responding to P-20W+ education policy initiatives.
3. State leadership¹ reviews and approves the P-20W+ data governance policy.
4. The leadership disseminates the policy to staff and to executive and legislative leadership, including a reference to where it will be available electronically.
5. The leadership identifies P-20W+ data governance leads, establishing the membership of the P-20W+ Data Governance Committee.
6. The P-20W+ Data Governance Committee uses the policy as the foundation for its P-20W+ data governance manual², which details how the policy will be put into practice.

Figures 1 and 2 depict two data governance approaches: intra-agency (single agency) and interagency (multiple agencies). The intra-agency approach (figure 1) is an agencywide approach to managing information from collection through use. There should be distinct

roles for and relationships among program areas, IT, and leadership, as well as local education agency representation and an agencywide data governance coordinator. Information is owned at the program-area level, and each data element has one owner.

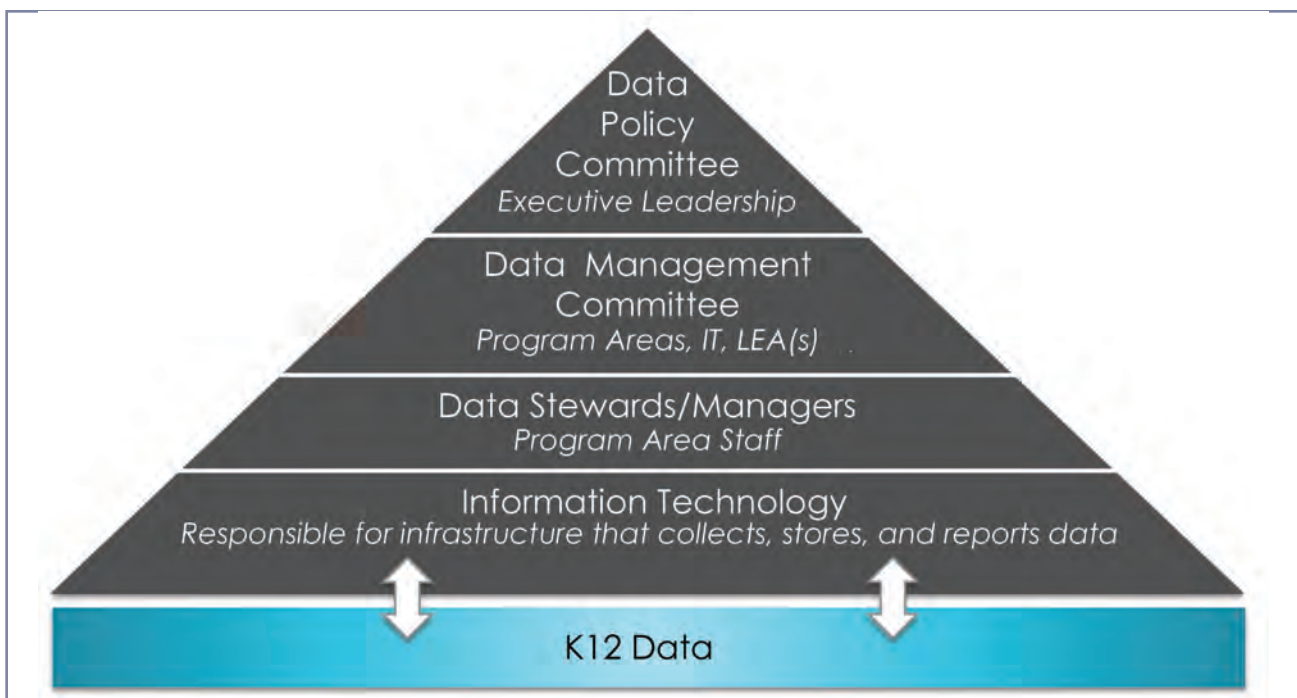


Figure 1. A traditional data governance structure for a single agency (a K12 agency is used as an example)

¹ State leadership includes leadership from those agencies participating in the P-20W+ system, along with leadership from the governor's office and legislature.

² The SLDS State Support Team has developed a P-20W+ Data Governance Manual Example and a P-20W+ Data Governance Manual Template that are available in the Public Domain Clearinghouse. See the Additional Resources section at the end of this brief.

The interagency approach (figure 2) offers a statewide perspective on data use to inform education policy and priorities. Roles for and relationships among state leadership and the leadership, program, and data management staff of participating agencies should be distinct and defined. In addition, state leadership and agencies should work together on data sources,

integration, reporting, and use. Information is owned at the agency or program-area level. A P-20W+ Data Governance Committee chairman serves as point of contact for, and facilitates communication between, the Executive Leadership and the Data Steward Workgroup. Both groups should include representation from all agencies participating in the P-20W+ SLDS.

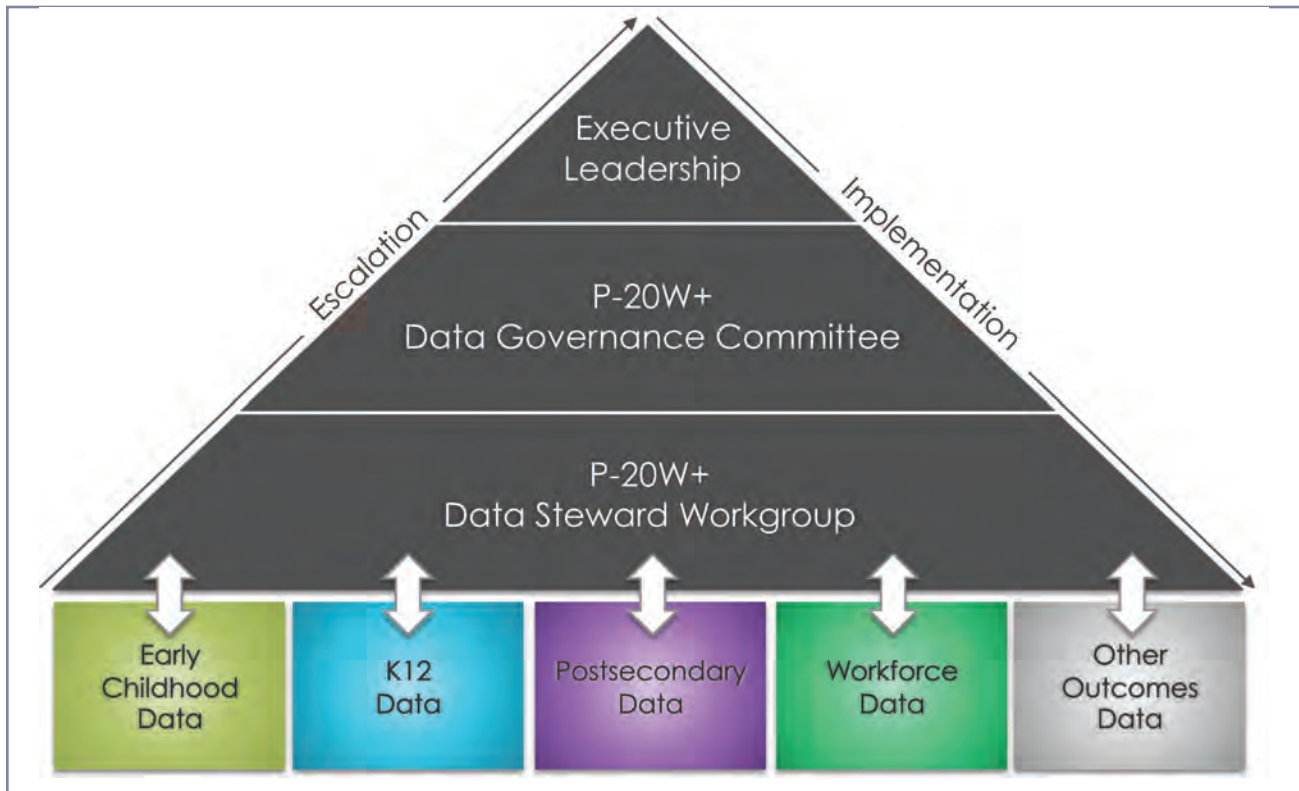


Figure 2. Data governance across P-20W+ partner agencies

Do:

According to states, the following practices are effective ways to establish P-20W+ data governance.

Engage state leadership.

Fully engage state leadership (Executive Leadership levels of figures 1 and 2) early in the development of P-20W+ data governance. Without this engagement, both short-term effectiveness and long-term sustainability of the P-20W+ system may be in jeopardy. Support from Executive Leadership helps foster the initial establishment of P-20W+ data governance, provides authority for cross-agency

decisions, and defines the commitment of personnel and other resources so that the ongoing needs of P-20W+ data governance can be met.

Pay attention to legislative activity in your state.

Many states have legislation impacting how data can be collected, where they can be stored, and who has access to them. During the planning phase, research all possible legislation that could impact your work. If state laws limit data sharing, there may be an opportunity to lobby to change the laws—particularly if the state

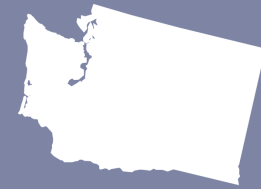
already has legislative and executive support for P-20W+ work. Additionally, demonstrating how the P-20W+ system can support legislative interests can help secure legislative backing and build momentum for a state's data collection efforts.

If a lack of state legislation affecting P-20W+ data collection, storage, and access is hindering the effort,

develop relationships through appropriate channels to begin the process of crafting legislation that will support the project. Otherwise, use the lack of legislation as an opportunity to develop the system while involving policymakers from the beginning so that there is buy-in for the project and it can be sustainable in the future.

Washington's Approach to P-20W+ Data Governance

The Washington State Legislature created the Education Research & Data Center (ERDC) in 2007 to help the state answer questions about education outcomes that no single agency at the time could answer. ERDC manages a centralized data warehouse containing data from state early learning, K12 and postsecondary education, and workforce agencies that are used frequently for reports and research. Datasets from additional sources, such as corrections and justice programs, can be linked to education data for specific reports or analyses but are not regularly collected in the warehouse.



Although ERDC's role in linking and managing data from multiple state agencies was established in law, the center soon recognized that the success of its P-20W+ work depended on building strong relationships and trust with its data contributors. Because ERDC is housed in the Office of Financial Management, which also oversees the state's budget, some agencies feared their data would affect the funding they receive. Some also worried about the time and work involved in contributing data to ERDC.

ERDC's data governance efforts began before it decided on the technical architecture for its data warehouse. Leaders spent eight months meeting with data contributing agencies at the start of the project. They often met multiple times with stakeholders who were most concerned about data sharing, discussing with them the valuable information they would gain by sharing their data.

Clarifying and documenting ERDC's role and data governance processes also helped address concerns. Memoranda of understanding among all the state agencies outline the expectations and responsibilities of each partner, including ERDC. They also cement ERDC's focus on questions affecting multiple P-20W+ sectors rather than single sectors and agencies.

Three groups contribute to data governance in Washington:

- The **Research Coordination Committee** is composed of research and policy representatives from agencies that contribute and use ERDC data. The group meets quarterly to revisit data sharing policies, discuss upcoming work, and hear presentations from data users about recent projects.
- The **Data Custodian Group** includes IT personnel from data contributors concerned with how data are transported to and stored at ERDC.
- The **Data Stewards Group** consists of data experts from each contributing agency who are familiar with what their data mean and how they can be used.

ERDC meets with data custodians and data stewards at contributing agencies as needed to discuss data sharing procedures and to make adjustments for any changes to the datasets they submit. To minimize burden on its data contributors, ERDC built its data warehouse to accept datasets from source systems in a variety of formats rather than asking contributors to conform to a single standard.

Although ERDC maintains the linked data and has the authority to release data on behalf of individual agencies, its data governance culture recognizes each agency as the expert on its own data. Questions about specific data are directed to the contributing agency rather than to ERDC. ERDC also works with researchers and state agencies to provide data that are not regularly collected in the statewide warehouse.

Convene all relevant parties. Educate each other to develop data knowledge and buy-in.

Once the Executive Leadership has defined the vision and direction for the P-20W+ system, data governance is needed to effectively implement this vision. Bring all involved stakeholders to the table and discuss why data governance is important and how it will benefit each of the participants in the P-20W+ partnership, as well as the state as a whole. In addition, discuss and document what data are already being collected—and by whom—within the state. The group can then begin to visualize which data could be shared, as well as what benefit sharing data can have for each of the relevant partners. Discussions are also needed to determine

how these data can address and support the vision defined by Executive Leadership. Such conversations can create a sense of ownership and common interest, and they can help alleviate any concerns participants have about misuse or loss of authority over their data. Establishing trust and common interest is vital to data governance efforts.

Possible partners in this discussion include key leadership at the executive level, director-level representatives from each agency, agency program area data steward representatives, and IT and data analysis staff from participating agencies. The key is to establish a committee structure that meets the needs of the state and is responsive to the culture. Generally there are three types of committees: (1) Executive

Connecticut's Approach to P-20W+ Data Governance

The five agencies involved in Connecticut's P20 WIN SLDS collaborate voluntarily to share data for cross-agency projects, reports, and research. When approaching data governance, the project leaders knew they needed to formalize the collaboration they started during the state's 2009 SLDS grant. They wanted a data governance approach that engaged the right people from each organization, gave them equal representation, and facilitated decisionmaking without becoming a burden on participants.



The P20 WIN data governance model is composed of three groups:

- The **Executive Board** includes one leadership-level member from each of the five partner agencies. The board advocates for P20 WIN initiatives, sets the strategic vision for the project, approves new agencies joining the effort and related agreements, and addresses issues that affect all partner agencies. Members meet on an as-needed basis.
- The **Governing Board** is responsible for developing data governance policies to be approved by the Executive Board, ensuring that data sharing and use guidelines are followed, and considering research requests. It also considers and proposes processes for cross-state data exchanges such as WRIS2, and it revises data sharing agreements as needed. The Governing Board consists of one representative per agency—usually a senior policy or research official—and meets biweekly.
- The **Data Steward Committee** includes two members from each partner agency—one chief or senior-level IT representative and one key data user. The group meets monthly and is responsible for ensuring proper data management and compliance with agreements across agencies. It also works on aligning each agency's data definitions to the Common Education Data Standards to aid data sharing.

The original partner agencies involved in Connecticut's P-20W+ data work—the State Department of Education, Connecticut State Colleges and Universities, and the state Department of Labor—developed the data governance structure and drafted a data governance policy with guidance from the SLDS State Support Team. When the University of Connecticut and the Connecticut Conference of Independent Colleges joined P20 WIN, the governance groups approved the new partnerships and updated memoranda of agreement to reflect the goals and legal framework governing data sharing for each agency.

No staff members involved in data governance are devoted exclusively to that effort. They depend on their agencies prioritizing and making time for data governance activities, and their capacity to manage some tasks, like research requests, may be limited. Even so, Connecticut's data governance model has proved to be a strong one. It allows agency officials to hold important discussions and modify agreements to address changes to the collaboration, such as adding new partners and establishing cost-sharing and budgeting guidelines.

Leadership, (2) P-20W+ Data Governance, and (3) P-20W+ Data Stewards (see figure 2). The chair of the data governance committee should act as a “data coordinator,” facilitating the work of the various committees and ensuring that all work is documented and shared with the entire governance group.

A word of caution: Bringing too many players to the table may hinder the group’s efforts. Strive to include only those individuals whose input will bring value to the discussion.

Create clear, distinct roles for and relationships among program areas, IT, and leadership.

Each partner should retain ownership of its own data. Each partner’s role can be clarified in an MOU and administered or managed by the P-20W+ data governance committee. MOUs should align with the statewide vision for P-20W+ policy, which can be defined by legislation or executive order.

Create a data governance policy.

Without a data governance policy, participating agencies and program areas are more likely to question the state’s commitment to the changes that are required to implement a P-20W+ data governance process. The policy provides strategic direction by creating a framework for decisionmaking about and accountability for how data will be managed and used. In addition, the data governance policy

- assigns stewardship responsibilities for P-20W+ agencies’ data;
- establishes overarching standards for the management of the integrated data; and
- empowers the P-20W+ Data Governance Committee to establish more detailed interagency standards and processes for (1) the communication about and (2) the management and use of integrated information to ensure quality and protection of the data.

Consider the state’s existing capacity.

Assessing capacity may involve determining the role or function of state agencies in the process of planning and implementing data governance as well as current state needs for successful implementation. What data will be involved—all data from the source systems, or only the data that are integrated to address P-20W+

priorities? Where will the data be stored? Is existing infrastructure able to accommodate the additional data? How will the data be secured? While it is the responsibility of the Executive Leadership to secure the required resources, it may be necessary to perform a gap analysis to determine what additional resources are needed and to create a strategy for securing those resources. Planning can only take data governance so far; work with IT and other relevant parties to determine limitations and opportunities related to budget, capacity, security, and sustainability.

Convene a data governance committee.

A data governance committee may be mandated by legislation or it may be a more informal group of policy, data, and IT experts. The committee’s role is to disseminate the agreed-upon data governance policy to state and agency leadership and policymakers, and to respond to the leadership’s vision and direction for policy. Data management efforts are sustainable only if all participating agencies are held responsible for adhering to the data governance policy, if authority for making interagency decisions is defined, and if all partners adhere to the leadership’s overall vision.

Have regular data reviews with pertinent parties.

Ideally, data should be collected and audited within the source systems at the participating agencies. However, there is value in convening stakeholders to review the comprehensiveness of the integration into P-20W+; identify any gaps or holes in the data as they are being integrated from multiple sources; and discuss appropriate use, as similar data element names may have varied meanings depending on the source. Again, each participating agency should have a sense of ownership over its own data and be held accountable for adhering to the data governance policy that is implemented.

Additionally, individuals or groups who are interested in presenting and using data are important stakeholders. They can help inform decisions about which data need to be integrated and how to address the Executive Leadership’s vision. These stakeholders play a pivotal role in sustaining the P-20W+ SLDS’s position as a resource for informing state-level policy. They can use the SLDS data to help define models to inform policy development decisions, report on policy and program effectiveness, and determine whether policy changes are needed.

Consider data governance a policy-led, rather than an IT-led, initiative.

Although IT is crucial for success, agency or program personnel and managers are experts on the data—not IT. Designing a system that can follow a student's progress longitudinally requires much thought and

sensitivity across various agencies and sectors, and IT plays an important supporting role in making this happen. While IT has the technical expertise, program personnel may be more sensitive to or aware of issues directly affecting the data.

Virginia's Approach to P-20W+ Data Governance

The federated system model and data governance structure of the Virginia Longitudinal Data System (VLDS) evolved within the context of strict state privacy laws and a culture of state agencies wanting to retain control over their data.

Aided by a 2009 American Reinvestment and Recovery Act (ARRA) SLDS grant, Virginia has expanded the VLDS partnership from an initial collaboration between the Virginia Department of Education and the State Council of Higher Education for Virginia (SCHEV) that began in 2008. The system's partners now also include the Workforce Development Services Division of the Virginia Community College System, Department for Aging and Rehabilitative Services, Department of Social Services, Department of Health Professions, and the Virginia Employment Commission.



A statewide committee composed of representatives from every data sharing organization oversees data governance as well as general system management. The members, who meet monthly, represent different levels of leadership at their respective organizations. Four subcommittees—Technical, Communications, Legislative, and Research—allow representatives to examine topics in greater detail and offer guidance to the full group.

Each organization has one vote in committee decisions, which are made by consensus around an informal standard of "can we live with it?" Data requests for research or other purposes are channeled through a centralized VLDS web portal to be evaluated and approved by the agencies whose data are involved rather than by the full group.

Although VLDS partner agencies have different levels of capacity and cannot always participate to the same degree in governance activities, the state's data governance structure has successfully supported cross-agency collaboration by allowing members to get to know one another and each agency's interests and priorities. As the number of agencies involved in VLDS has increased over the years, the data governance structure has remained largely unchanged. Changes to the VLDS bylaws codified in the Book of Data Governance—which is maintained by the governance committee—have become less frequent as the partnership has expanded. Leaders believe the governance committee could accommodate several additional partners in the future without significant restructuring.

Prospective data sharing partners are invited to attend committee meetings and discuss their priorities with VLDS staff in advance of formally joining the system. To join, new agencies must be approved by the governance committee, sign onto the Book of Data Governance and a memorandum of understanding, and pay the cost of technical enhancements needed to enable data sharing—currently \$20,000.

Additional Resources

Connecticut P20 WIN

<http://www.ct.edu/initiatives/p20win>

P-20W+ Best Practices: SLDS Issue Brief

<https://slds.grads360.org/#communities/pdc/documents/5231>

P-20W+ Data Governance Manual Example

<https://slds.grads360.org/#communities/pdc/documents/3081>

P-20W+ Data Governance Manual Template

<https://slds.grads360.org/#communities/pdc/documents/3080>

Sustaining Core Processes for Data Governance: SLDS Webinar

<https://slds.grads360.org/#communities/pdc/documents/12690>

Traveling through Time: The Forum Guide to Longitudinal Data Systems, Book III: Effectively Managing LDS Data

<http://nces.ed.gov/pubs2011/2011805.pdf>

Virginia Longitudinal Data System

<https://vlds.virginia.gov/>

Washington State Education Research & Data Center

<http://www.erd.c.wa.gov/>