Types of Data Systems

Tate Gould, Amy Nicholas, Tony Ruggiero, William Blandford, Sara Thayer, & Bruce Bull

Planning for a new data system? Preparing to link data across agencies? Part C and 619 staff are major stakeholders in data system improvement efforts. Being an informed stakeholder requires being familiar with basic types of data systems. This brief describes and compares four common types of data systems.

What Are the Various Types of Data Systems?

There are several types of data systems that support data from Part C/619 programs. Although the system types have similarities, each has its own unique characteristics and purposes. The attributes that make one type of data system a particularly good fit for one data-related need or function can be less desirable for another need or function. In fact, a Part C/619 program may have several different types of data systems that meet different needs. This brief describes the characteristics of four common types of data systems: **static/reporting, transactional, federated,** and **centralized**. It is important to note that a data system may have attributes of more than one type.

Static/Reporting Data System

A **static/reporting** data system is a one-way system in which the end user is able to see data but is unable to modify those data. Some static/reporting systems do, however, allow users with permissions to run queries for analysis purposes. The data elements in this type of system are *static*, or fixed and unchanging. Data are commonly loaded in periodic batches (e.g., annually). These systems are well designed for sharing data with the public. Static/reporting systems usually have minimal maintenance needs. An example of a static/reporting data system is one that supplies the data to operate a public State Performance Plan (SPP) reporting website, such as Maryland's online IDEA Report Card (Figure 1). The full website is available at http://mdideareport.org/Indicators.aspx?IndicatorType=1.

Figure 1. Maryland's Online IDEA Report Card

Source: Maryland State Department of Education.

Transactional Data System

A **transactional** data system records and tracks data transactions. Much like a typical banking account tracks financial transactions, transactional data systems store events or "transactions," such as a child enrolling in a program. Each transaction has related data points to record relevant information about the event (e.g., date of enrollment, name and agency number where child enrolled). The use of a transactional data system allows for the back-and-forth exchange of data between users, such as local Part C service providers and their program administrators. A transactional data system records the iterative history of *dynamic*, or changing, data. The detailed data collection and storing that occur with each transaction allow for robust near real-time reporting and data use. An example of a transactional

data system is a comprehensive child-level online Individualized Family Service Plan (IFSP) that records data from referral to exit, such as data associated with screening, eligibility, transition, and the like.

Federated Data System

A **federated** data system is a collection of cooperating data systems that are autonomous and often different in design. Such systems are most common where multiple programs/agencies join data to form a single dataset, such as a statewide longitudinal data system (SLDS). In a federated system, individual programs/agencies maintain control over their own data but agree to share some or all of those data with other participating agencies upon request. This information sharing is accomplished through "interoperability." Interoperability means a set of rules, definitions, and transport processes that enable different software systems to share information and work together so agencies with different software systems can share information regardless of their operating system or platform. Email is an example. A Gmail user can email a friend who uses a Yahoo account, and the email, attachments, time stamps, and all other contents are easily transferred.

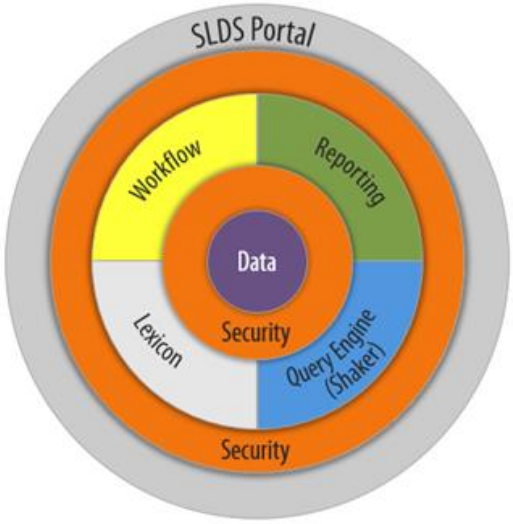
When conducting queries in a federated system, users submit requests via a shared intermediary interface that then searches the independent data systems for the requested information. Federated data systems have limitations, causing many agencies/states to move away from using them. For example, a dataset in a multi-agency federated data system can have different values for the same data element, such as date of birth, and the burden is on the end user to resolve the discrepancy.

An example of a multi-agency federated data system is Virginia's SLDS (Figure 2). The state uses a federated data model and technical architecture composed of a web-based user interface, guided queries, a multilevel security structure, and rich business intelligence capabilities. For more information about Virginia's SLDS, visit http://vlds.virginia.gov/.

Figure 2. Virginia's SLDS

Partners:

- **Department of Education (DOE)**
- **State Council of Higher Education in Virginia (SCHEV)**
- **Virginia Employment Commission (VEC)**
- **Virginia Information Technologies Agency (VITA)**
- **Virginia Polytechnic Institute (VT)**
- **Center for Innovative Technology**
- **Department of Social Services**



Source: Virginia Department of Education.

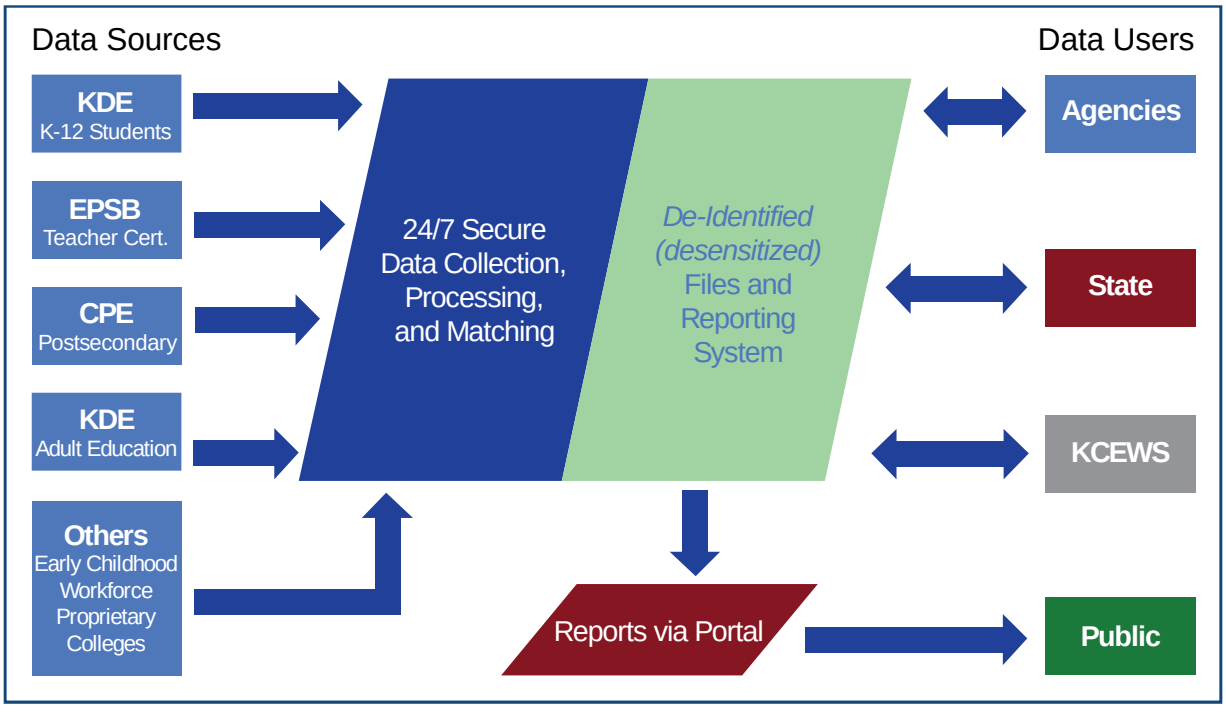
Centralized Data System

A **centralized** data system is located and maintained in one central location. Processing of data is done centrally on a single hosted server. Centralized data systems can be used by single agencies or for multi-agency initiatives. In a multi-agency centralized data system, source programs/agencies copy their data to a single, centrally-located repository where the data are organized, integrated, and stored using common data definitions and standards. Therefore, interoperability is less of an issue because information sharing is managed by the centralized system rather than individual participating systems. A centralized data system is advantageous for performing longitudinal studies, since the data are static and these systems are designed to store a large quantity of data.

An example of a *single agency* centralized data system, that is also transactional, is a state Department of Education-supported online Individualized Education Program (IEP) system. All local districts can access the system for data on their students, transfer students between districts, and perform transactional data collection functions associated with each student. The system is hosted by the Department on a cloud server though rather than each local district hosting on individual servers.

An example of a *multi-agency* centralized data system is Kentucky's P-20 Data Collaborative. Contributing agencies (including early childhood) provide data to the P-20 system on a regular schedule (Figure 3). Those data enter the system through a staging environment where they are validated, standardized, and matched. The data are stripped of all identifiable information, such as name, date of birth, social security number, and the agency or institution identification. Unique identifiers are assigned to protect individual personal and institutional information. Once the data are stripped of identification, they are placed in a de-identified data warehouse where analyses, reports, requests, and evaluations are created. Contributing agencies' staff have access to the merged, de-identified data through a web-based reporting tool. For more information about Kentucky's P-20 Data Collaborative, visit http://kcews.ky.gov/.

Figure 3. Kentucky P-20 Data Warehouse Architecture



Source: Kentucky Center for Education and Workforce Statistics.

Comparing Data Systems At-a-Glance

Figure 4 provides a comparison of capabilities across the four types of data systems described in this brief. Keep in mind that a data system can have attributes of more than one type of system.

Figure 4. Comparison of Data System Capabilities

	Data can be exchanged between multiple data systems in near realtime	Allows end users to enter dynamic (changing) data	Programs/agencies have ownership of their data	State and local program/agency staff have access to data for analysis
Static/reporting	No	No	Yes	Yes*
Transactional	Yes	Yes	Yes	Yes
Federated	Yes	No	Yes	Yes*
Centralized (multi-agency)	No	No	Yes	Yes*

*Possible but typically not local staff.

For more information about federated and centralized models, see the Statewide Longitudinal Data System Grant Program's brief *Centralized vs. Federated: State Approaches to P-20W Data Systems* (October 2012), available at http://nces.ed.gov/programs/slds/pdf/federated_centralized_print.pdf.

Selecting the Right Data System for Your Part C/619 Program

When designing a new data system or data system enhancement, Part C and 619 staff need to consider political, financial, administrative, and other factors. It is important to start any exploration of a new data system or enhancement by identifying your program's needs for data collection, reporting, sharing, and use. If your state is considering developing a multi-agency data system, Part C and 619 should be actively involved in the decision-making. Good data system planning with program staff involvement will result in a system that performs the functions staff need and provides them with access to all data collected.

Suggested Citation

Gould, T., Nicholas, A., Ruggiero, T., Blandford, W., Thayer, S., & Bull, B. (2015). *Types of data systems*. Menlo Park, CA: SRI International.

About Us

The contents of this brief were developed under a grant from the U.S. Department of Education, #H373Z120002. However, those contents do not necessarily represent the policy of the U.S. Department of Education, and you should not assume endorsement by the Federal Government. Project Officers, Meredith Miceli and Richelle Davis.

The DaSy Center is a national technical assistance center funded by the U.S. Department of Education, Office of Special Education Programs. The DaSy Center works with states to support IDEA early intervention and early childhood special education state programs in the development or enhancement of coordinated early childhood longitudinal data systems.

All graphics depicted in this brief were used with permission by the applicable state education agencies.

To learn more about the DaSy Center, go to http://www.dasycenter.org/ or call 650-859-3881.