SLDS Topical Webinar Summary

Introduction to ECIDS System Design – The Basics 101

System design represents a critical step when building an early childhood integrated data system (ECIDS). As a state begins to design and build its own system, it is important to consider the overall vision and goals of the system and what questions may arise from stakeholders. This webinar defined a four-level process involved in system design for an ECIDS. Utah, North Carolina, and Maine shared methods, best practices, and lessons learned using system design to inform their ECIDS.

System Design as It Relates to an ECIDS

An ECIDS is a data system that integrates early childhood education, health, and social service information from key participating state agencies. System design can take on many different meanings and definitions depending on how it is applied. For all ECIDS projects, system design is a component of a larger system lifecycle and can be a very complex IT process to translate a state's needs into a technical solution. This process often involves four stages: assessment, design, development, and implementation of the data system. This process is illustrated in figure 1.

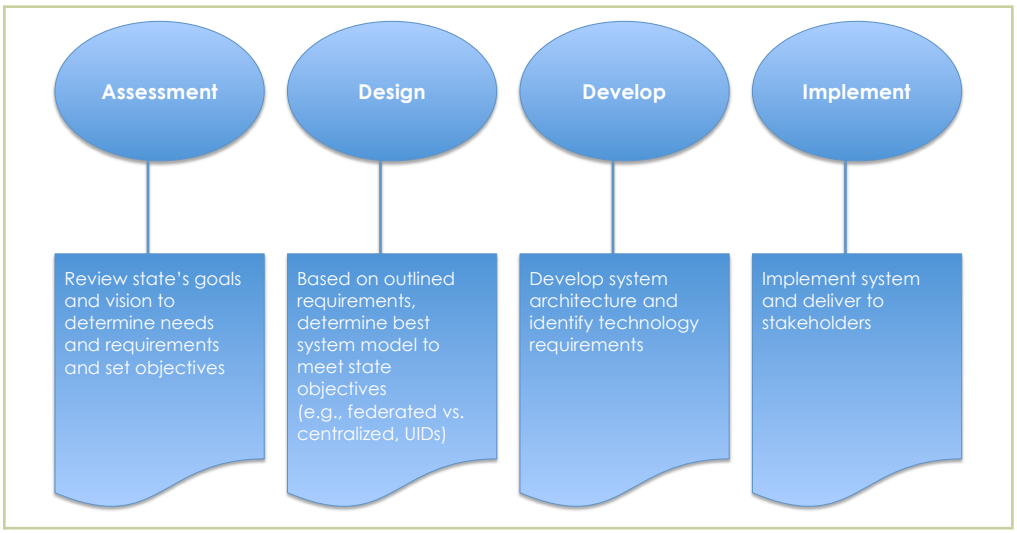

Figure 1. The Four-Stage Process of ECIDS System Design

1. Assessment

At the assessment stage, states should first review their mission and vision to ensure that the goals for the ECIDS will be met with the system design. From there, states should create a communications plan, establish business requirements, and then identify which requirements to prioritize. At this stage, states should also work on creating data sharing agreements among the agencies that will contribute data to the ECIDS. While this is considered one of the most challenging aspects of system design, it is a critical piece that will help various stakeholders understand what data are being shared and how the data can be used.

2. Design

Design is the second stage of system design, and it is the most important technical stage. During the design phase, states determine the type of model they would like to implement for the integrated data system: federated, centralized, or a hybrid approach. For many

This product of the Institute of Education Sciences (IES) was developed with the help of knowledgeable staff from state education agencies and partner organizations. The content of this brief was derived from a Statewide Longitudinal Data Systems (SLDS) Grant Program monthly topical webinar that took place on May 28, 2014. The views expressed do not necessarily represent those of the IES SLDS Grant Program. We thank the following people for their valuable contributions:

Webinar Presenters:
Stephen Clyde
Utah

Gary Alexander
North Carolina

Ted Fournier
Maine

Moderator:
Steve Duarte
SLDS Grant Program, State Support Team

Jeff Sellers
SLDS Grant Program, State Support Team

For more information on the IES SLDS Grant Program or for support with system development, please visit http://nces.ed.gov/programs/SLDS.

ies NATIONAL CENTER FOR EDUCATION STATISTICS
Institute of Education Sciences

states, the choice of model depends on the culture of the agencies involved and the degree of comfort the state has with sharing data. All three models will require business rules to determine how the data are extracted and to what degree the data will be integrated. Depending on how the data will be matched and linked together, these business rules may cover processes for transforming data as they enter the ECIDS. While many data elements will be unique to their source systems, there may be data elements—such as those with similar names but different code sets or definitions—that will need to be discussed and reconciled through data governance.

During the design phase, states should also do the following:

- *Perform an inventory of current systems and data that are available.* As states are approached with policy questions, an inventory of current data will help them recognize if enough data are available to answer the questions or inform research projects. This inventory should also identify any restrictions on the use of specific data.
- *Heavily document every aspect from beginning to end of the process.* Documentation will continue to be an important tool as the system design is developed.
- *Use a reliable unique identifier (UID) and data matching process.* The use of a UID or other identifiable matching data is the "backbone" of a state's integrated data system. Without reliable linkages, the resulting analysis will be difficult to use.
- *Describe desired features of the system and create system diagrams.* Because an integrated data system is theoretical, states should create tangible items such as screenshots and handouts to help stakeholders understand the system.

Types of Data Models

Centralized
Once data integration occurs in a centralized system, the data linkages are loaded into a central database or warehouse so that they can be used for multiple purposes, such as reports and data extracts.

Federated
In a federated model, datasets are created from linked data but the linkages do not persist after each use. The data linkages are generated upon every data request. An advantage of using a federated data model is that states do not need a central data warehouse for the integrated data system.

Hybrid
A hybrid data model combines aspects of the federated and centralized data models. The identifiable data that a state uses to generate linkages among records is maintained, but the records themselves are not linked.

3. Development
At the development stage, states should adopt a development lifecycle process. Project management remains a crucial step at the development stage and will help to ensure that states are bringing in the data components that will eventually be linked in the ECIDS. During this stage, states will also begin engaging with vendors, issuing requests for proposals, and managing the work of vendors once they are brought on board with the project. "It's a bit more involved beyond just coding," said Steve Duarte of the Statewide Longitudinal Data Systems (SLDS) Grant Program State Support Team. "It's a lifecycle event."

4. Implementation
In the final stage of implementation, states should be ready to deliver the ECIDS to stakeholders and end users. At this level, the ECIDS should be stable and secure and all measures need to have been taken to ensure confidentiality and compliance with federal, state, and local regulations.

Utah

Defining Goals and Overall Vision
Utah works with the state Department of Health and Utah State University in building the state's ECIDS. "Longitudinal data systems are the 'Batmobile' and can get complicated," said Stephen Clyde of Utah. "But by focusing on basic ideas you can manage the complexity and not let it get out of hand."

Utah's team of stakeholders met early on to discuss the overall vision and goals for the system. Through those meetings, stakeholders identified how the state would like to use the ECIDS: to create pre-defined and quality reports, data analytics, and ad hoc reports. These reports would be used to answer the state's policy questions and to create a research platform. The data would be pulled from a number of data sources, including early childhood programs such as Help Me Grow, Head Start, and Early Intervention, as well as state registries for birth certificates, immunizations, etc. Another major goal of Utah's ECIDS is to feed the early childhood data into the state's existing longitudinal data system, the Utah Data Alliance (UDA).

Importance of Data Inventory
At the beginning phase of its integrated data system, Utah performed a data inventory to see what data were available. By looking through the available data from the early childhood programs and registries, the state was able to align those pieces of information to see where there was duplicative and complementary data. From this initial inventory, Utah uncovered a number of alignment issues with the data, including subtle differences in data granularity, semantic differences, and timing issues.

Data Sharing Snapshots
After reviewing its goals, the state began looking at the design activities for its integrated data system. For its

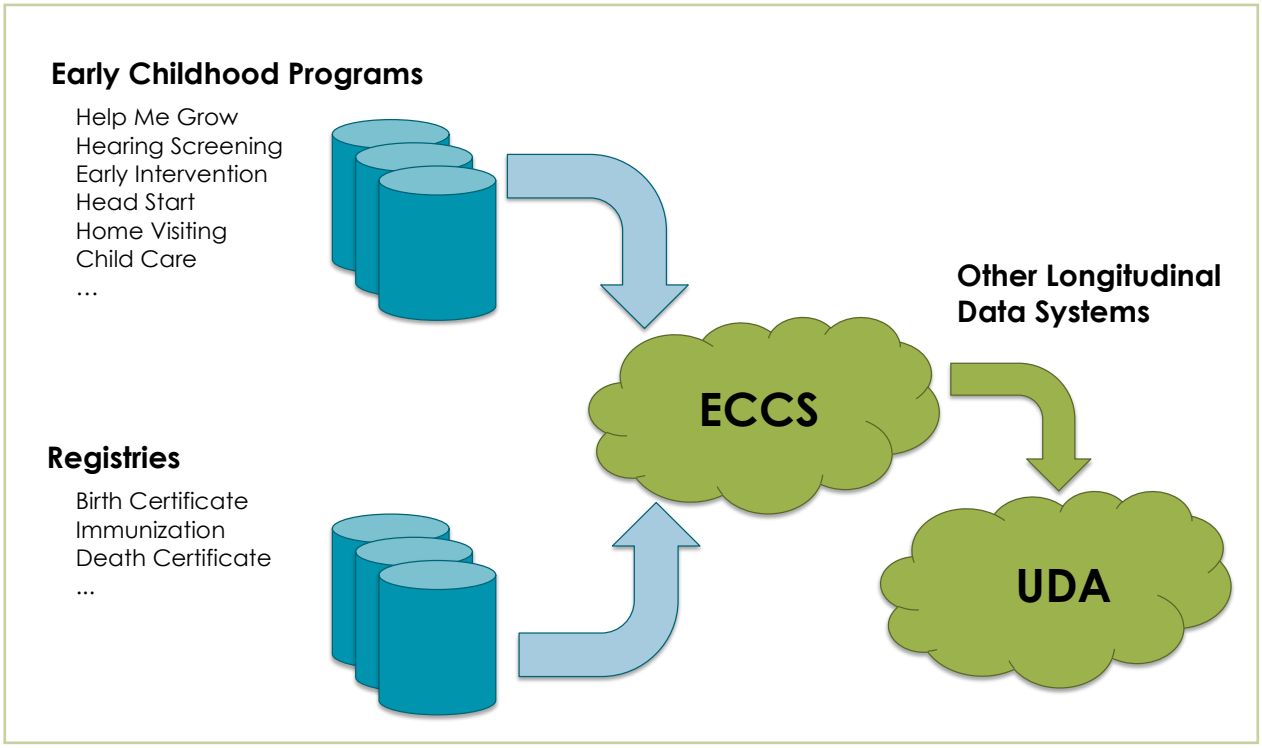

Figure 2. Overall Vision for Utah's ECIDS

ECIDS, Utah uses different levels of data at different levels of granularity to create a "person snapshot." For example, a person snapshot for student "John" from May 1, 2014, might include data from the immunization registry and his Head Start program. Those data will sit on John's timeline and can be used to answer questions about John at a given time or over a range of time. "This ability to work with data in time was central to the vision [for Utah]," said Clyde.

Importance of Data Sharing Agreements and Documentation
Utah used data sharing agreements in the assessment phase of its integrated data system to define responsibilities among the system's stakeholders. The data sharing agreements also helped to establish the type of data that would be extracted and translated from the system, and they spoke to policy enforcement and data analytics.

"The biggest thing that helped us with our data sharing agreements was documentation, which help[ed] us with mapping and sorting out semantics and granularity issues," said Clyde.

To facilitate the data sharing agreements, Utah created a documentation database system to help with mapping and granularity issues.

Utah also noted that the actual implementation of the ECIDS was underestimated. Due to the complex work involved in setting up and testing the data system, the project team realized that documenting the system's implementation might be helpful in the future.

North Carolina

North Carolina's overall vision for its ECIDS is to build an application to collect data from a variety of sources that will be used to answer critical questions and help to make informed decisions about its early childhood programs and services. To do this, North Carolina is using a federated data model to create a web portal application for initiating and receiving ECIDS data requests. The portal will also maintain a data dictionary of all the data available and the source of those data.

'One Stop Shop' for Early Childhood Data
The web portal will be a "one stop shop" for requestors to view a pre-structured report or to select data they would like to see from specific data sources. The portal gathers and delivers data to requestors within a workflow-based process. The state will also continue to use a UID solution from eScholar to enable uniform data collection across all data stores.

North Carolina recommends using a methodology to manage an ECIDS project. The state currently uses Agile methodology software, which is an iterative application software that allows for more collaboration between the business and technical teams. This allows the state to have more frequent and iterative application releases, and to be more responsive to new requirements.

Importance of Data Agreements and Proofs of Concept
North Carolina suggests securing "top-down" support from business leaders who will prioritize their participation and commitment to the project. Each representative should be dedicated to the project; otherwise, continuity issues can arise.

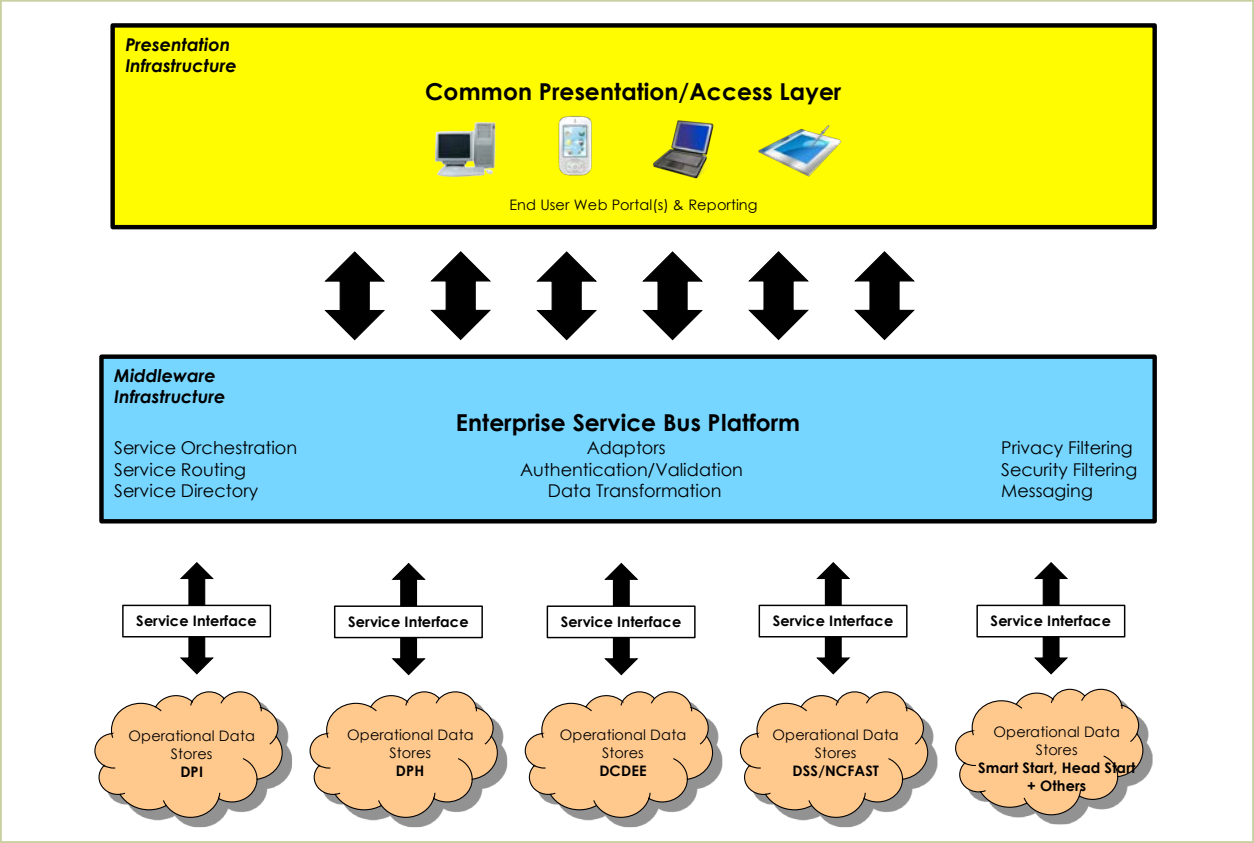

Figure 3. Overview of North Carolina's ECIDS

The state also notes that developing data sharing agreements among the stakeholders is one of the most challenging aspects of creating an ECIDS and can take a long time to complete. However, it is a critical step for the foundation of the ECIDS.

North Carolina suggests creating a proof of concept or an early prototype for key features and functions of the application. A proof of concept helps stakeholders to visualize the key features and functions of the application and can be used to estimate operating and supporting costs.

Maine

Maine's ECIDS uses a federated model that leverages the state's current SLDS. The SLDS uses data from the state's K12 student information system, which uses a common State Student ID (SSID). A federated system was selected to adhere to the state's cultural norms and to allow agencies to retain ownership of their data.

Maine is able to create data matches with the state's Department of Health and Human Services (DHHS) by linking SSIDs with DHHS IDs in various agency data systems through a Cognos/Oracle pilot system.

In addition to the K12 SSIDs, Maine's SLDS has access to the unique identifiers assigned to all children in 4-year-old programs in public schools as well as direct linkages to Maine Educare for all children assigned state K12 SSIDs. The state also has a project to match or assign all K12 SSIDs to all children enrolled in Child Development System programs.

Data Sharing Umbrella Agreement

An interagency group formed by the Maine Department of Education and DHHS has developed a cross-agency governance structure that also includes representation from the state's Department of Labor and the University of Maine System (UMS). This governance body developed a data sharing umbrella memorandum of understanding to

- define the scope and parameters of data sharing;
- establish UMS as a trusted third-party broker to handle identifier matching and facilitate data sharing and research across agency data systems; and
- serve as a master agreement for adding specific data sharing and research requests.

Maine's biggest challenge was establishing and managing defensible legal linkages for secure data access across state and federal agency data systems that were subject to different legal mandates, such as the Family Educational Rights and Privacy Act (FERPA) and the Health Insurance Portability and Accountability Act (HIPAA). Maine also notes that addressing stakeholder concerns surrounding the cross-agency data sharing is critical, especially with regard to early childhood confidentiality and privacy.

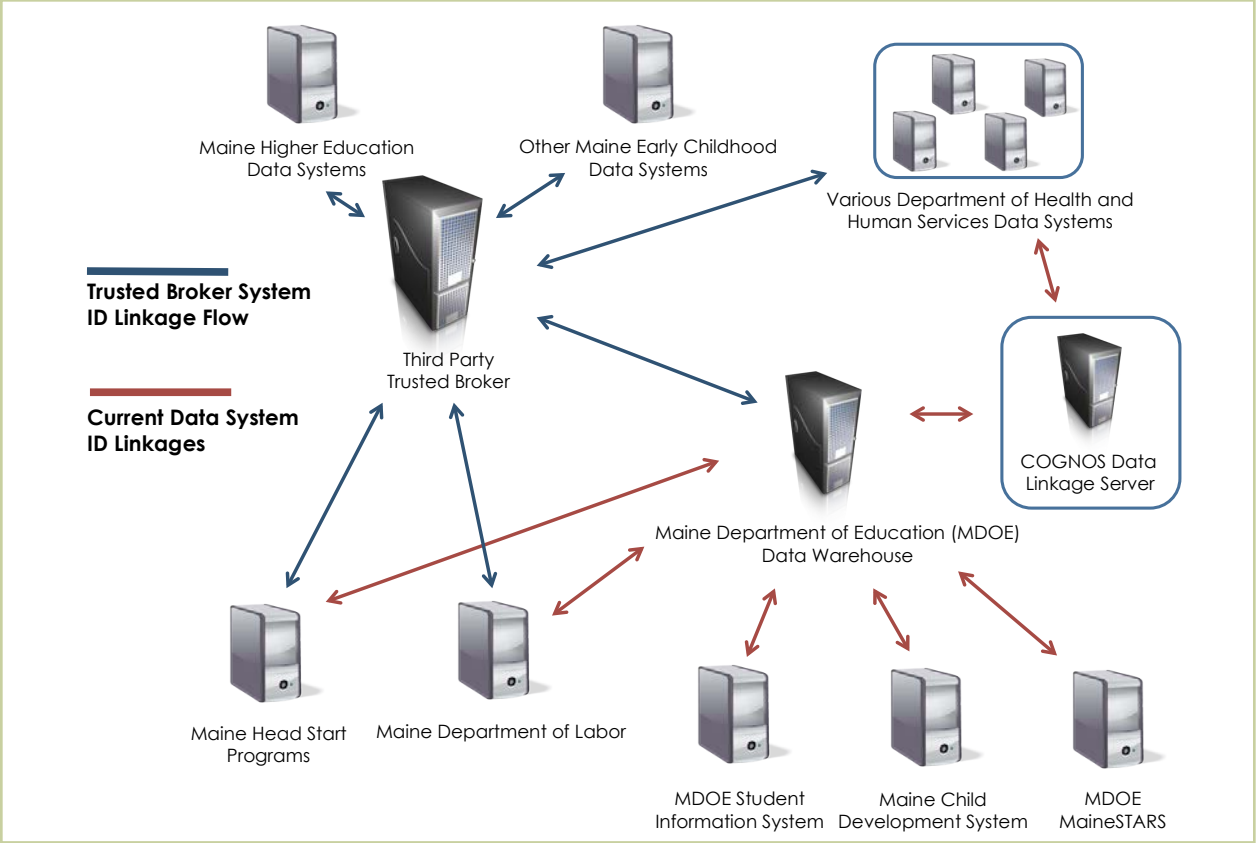

Figure 4. Overview of Maine's ECIDS

The ECIDS projects in Utah, North Carolina, and Maine have been designed to reflect the unique data sharing objectives, processes, and requirements in each state. Project leaders in all three states began the system design process by collaborating with data-contributing agencies and system stakeholders to assess their needs and establish a common vision and goals for the ECIDS. As the states moved through the design, development, and implementation stages of their ECIDS projects, they used the goals and expectations established with partners and stakeholders to determine the system design that would best meet those needs. Careful planning and conscious consideration of ECIDS design provide the foundation for systems that can grow and adapt to new early childhood priorities in the future.

Additional Resources

Early Childhood Data Collaborative
http://www.ecedata.org/

SLDS Early Childhood Integrated Data System Toolkit: System Design
https://slds.grads360.org/#program/ecids-toolkit:-system-design

SLDS Issue Brief: Early Childhood Data Governance in Action! An Introduction
http://nces.ed.gov/programs/slds/pdf/EC_DataGovernance.pdf

SLDS Issue Brief: Early Childhood Data Governance in Action! Initial Steps to Establish Data Governance
http://nces.ed.gov/programs/slds/pdf/EC_DataGovernance_Initial.pdf

SLDS Issue Brief: What is an Early Child Integrated Data System?
http://nces.ed.gov/programs/slds/pdf/WhatisanECIDS.pdf