DATA INTEGRATION
TOOLKIT

CENTER *for the*
INTEGRATION *of* **IDEA DATA**

Table of Contents

CENTER *for the*
INTEGRATION *of* **IDEA DATA**

The Center for the Integration of IDEA Data (CIID) has a unique and specific mission: to provide technical assistance to state education agencies (SEAs) to increase their capacity to report high quality data required under the Individuals with Disabilities Education Act (IDEA) Part B Sections 616 and 618 by integrating IDEA data systems and processes with the Statewide Longitudinal Data System (SLDS).

In the IDEA data and SLDS technical assistance (TA) communities an array of resources and tools exist that SEAs and TA providers can use to support these efforts. The CIID team is working to develop resources and tools to support the integration of IDEA data. The CIID Data Integration Toolkit describes a series of steps, tasks and activities – a process – to organize these tools and SEA resources to achieve data integration goals.

How was this toolkit developed? Our CIID TA providers have years of experience as former SEA staff doing this exact work. Drawing on these experiences, and using best practices in the fields of data integration and project management, we have built the CIID Data Integration Toolkit. CIID TA providers will engage SEAs to support data integration efforts and will use the Toolkit to guide their work. While the CIID team is available to help, the Toolkit is also a resource for any SEA to use when integrating a data system or collection on their own.

By using the toolkit and integrating data, your agency can increase its capacity for reporting IDEA Part B Sections 616 and 618 data. Data integration will also result in higher quality IDEA data for better informed decision-making for improved educational outcomes for students with disabilities.

The CIID team will continue to update and refresh the Data Integration Toolkit based on interactions with, and feedback from, SEAs. If you have feedback on any aspect of the Toolkit, or want to request TA, please email us at CIIDTA@aemcorp.com.

Thank you!

Bill Huennekens
Director, Center for the Integration of IDEA Data

Anna Mark
Deputy Director, Center for the Integration of IDEA Data

CIID Data Integration Toolkit Overview

Purpose of the CIID Toolkit

The CIID Data Integration Toolkit is designed to support State Education Agency (SEA) efforts to integrate Individuals with Disabilities Education Act (IDEA) data with their statewide longitudinal data systems (SLDSs), and includes:

- A sample road map articulating typical steps and tasks required for integration
- A sequential process for integrating data (note: the order of completion may vary by SEA)
- Specific activities and resources to support integration including checklists, templates, sample process documents, and state examples

How to Use the Toolkit

This Toolkit outlines the major steps and supporting tasks required to complete any data integration process. Each task description details its purpose, suggested activities, timeline, lessons learned, and available resources. The intended users of the Toolkit are key staff in each SEA who will work collaboratively to coordinate efforts across program areas, and may include:

- Special Education Directors
- Part B Data Managers
- ED*Facts* Coordinators
- SLDS Staff
- Data Governance Coordinators
- IT Staff
- Data Stewards

SEA staff can use the Toolkit to plan and support integration efforts in their SEAs in collaboration with CIID TA providers or on their own in collaboration with their colleagues within the SEA. To access additional integration resources or receive technical assistance from CIID, contact CIIDTA@aemcorp.com or visit CIIDTA.org.

Toolkit Assumptions and Assertions

The toolkit and integration process make the following assumptions and assertions:

- State toolkit users should have expressed interest to integrate
- Although data integration can be complex and vary by state, it includes a set of standardized steps
- Integration can be a sequential process but can vary by state
- The toolkit is focused primarily on technical steps
- The toolkit is focused on state-level integration of data only (not district to state)

Web & Print Versions

(7) **Major steps to data integration**

(25) **Tasks in support of 7 steps**

... and a host of other resources including checklists, timelines, useful links, & more

Major Steps and Tasks

Steps	Tasks
Step 1 Define the goals for integration	**Task 1.1:** Define use case and goals for integration.
	Task 1.2: Present use cases to state education agency (SEA) leaders for support and approval.
	Task 1.3: Define the scope of the data integration project.
Step 2 Establish project plan and structures for data integration work	**Task 2.1:** Establish project team members and regular meeting schedule for project team.
	Task 2.2: Develop project plan and team responsibilities for data integration.
	Task 2.3: Determine and secure agency resources required.
	Task 2.4: Create resource repository to support integration efforts.
	Task 2.5: Create communication plan for external and internal stakeholders.
	Task 2.6: Determine how vendors and contracts will be included in all steps and tasks.
Step 3 Locate and organize all potential data elements and associated attributes for integration into a data repository	**Task 3.1:** Identify datasets or systems to be integrated.
	Task 3.2: Document, in the Common Education Data Standards (CEDS), all data systems and elements associated with the data integration effort.
Step 4 Complete master integrated dataset	**Task 4.1:** Determine redundant data elements.
	Task 4.2: Resolve element redundancies.
	Task 4.3: Compile master dataset based on data dictionary for each dataset/system to be integrated.
Step 5 Implement and perform extract, transform, and load (ETL) procedures	**Task 5.1:** Develop/update ETL target database schema.
	Task 5.2: Identify all program business requirements needed for ETL(s).
	Task 5.3: Create an ETL requirements document for the programmers.
	Task 5.4: Conduct pilot/test cases for ETL and validate data load with the program team.
	Task 5.5: Perform ETL.
Step 6 Conduct review and validation of data integration	**Task 6.1:** Document and perform data quality checks.
	Task 6.2: Secure final sign off on overall process.
	Task 6.3: Review integration processes and plans to sustain integration.
Step 7 Conduct post integration supports and activities	**Task 7.1:** Create and disseminate final calendar containing all processes and key dates for newly integrated data system.
	Task 7.2: Create documentation for users to understand how to access integrated data.
	Task 7.3: Create and disseminate report for stakeholders summarizing reduction of risk, realized efficiencies, lessons learned, and resources available.

CENTER *for the* **INTEGRATION** *of* **IDEA DATA**

Contact Us: By email at CIIDTA@aemcorp.com. Visit the CIID website for more information at CIIDTA.org.

DATA INTEGRATION
TOOLKIT
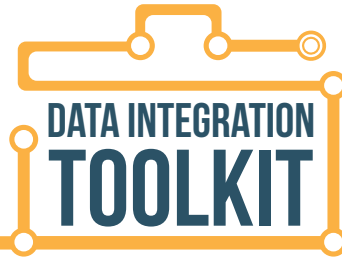
Step 1: Articulate Use Case(s) for Integration

During Step 1, SEA staff identify why they want to integrate their data. Using the "use case" process, staff clearly articulate the reasons and challenges of data integration. By completing Step 1, a shared understanding of the scope, outcomes, and support of data integration efforts are established.

**Task 1.1:** Identify use cases for integration and complete use case justification.

Task 1.2: Present use cases to state education agency (SEA) leaders for support and approval.

Task 1.3: Define the scope of the data integration project.

CII **CENTER** *for the* **INTEGRATION** *of* **IDEA DATA**

Contact Us: By email at CIIDTA@aemcorp.com. Visit the CIID website for more information at CIIDTA.org.

CENTER *for the*
INTEGRATION *of* **IDEA DATA**

**DATA INTEGRATION
TOOLKIT**
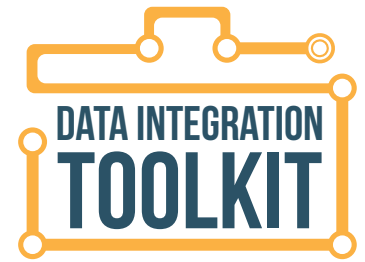
The CIID Data Integration Toolkit

STEP ① Articulate Use Case(s) for Integration

Task 1.1: Define use case and the goals for integration.

Overall Description of the Task

A use case describes what you want to accomplish through data integration. In this first task, you will document why you need data integration, the challenges you want to solve, and the individuals who will champion or challenge the data integration effort. Completion of this task will produce one or more use case documents that will help team members explain and build support for data integration (Task 1.2) and define the scope of the data integration project (Task 1.3) and understand and describe the purpose of integration and its expected impact on stakeholders.

Why Is This Task Important?

• The use case describes the goal of integration including the impact and benefits in general terms. It can be used to build support within your state education agency (SEA) for the integration effort.
• The use case creates a common understanding of the problem, current need, or goal that data integration will address.
• The process of describing the use case and the expected outcomes of integrated data will:
  » Improve support for the project by demonstrating the benefits of data integration to critical stakeholders within the SEA;
  » Establish the general scope for the project and inform the development of an integration plan and ongoing maintenance; and
  » Provide the project team with an initial list of stakeholders to support the project and help gather requirements, provide feedback, and test the final functionality.

Activities

1.1.1 Write a need statement and associated questions for each use case.

Begin developing a use case by describing a current need, challenge or set of critical questions that could be addressed if the data were integrated. Working with a core group of interested staff, justify why data integration is important for you and your colleagues' work. The need statement should be a short narrative that explains what data integration will accomplish. If you identify multiple needs, consider developing multiple use case documents to demonstrate the different benefits data integration could support. An example of developing a use case is provided below. A use case template is provided under Resources.

Example:

- Use Case: Improved understanding of the population of students served by various federal programs, especially those with special needs, by integrating federal program participation and special education program data with our SLDS.
- Need Statement: State policy makers are working collaboratively to provide technical assistance to school districts across federal programs including title programs and IDEA. State policy makers find it difficult to identify how many students are served by multiple programs because federal program participation data and special education data are separate from each other. To facilitate cross program technical assistance to school districts, we will integrate federal program participation data and special education student data with our SLDS.
- Stakeholders: Policymakers, IT staff, program office staff and other data users
- Impacts:
  » Policy makers will be able to determine which student populations are served by multiple federal programs including IDEA and whether students served by those programs have better outcomes, thereby informing technical assistance and support provided to school districts;
  » Implementation and maintenance of data integration may involve additional tasks for SEA IT staff (e.g., new manual, yearly process for ongoing support, increase of workload for one-time integration, or temporary increase to set up automated process);
  » Data users will require re-training to understand the new reporting process; and
  » Financial – The data integration project will require shifting of staff resources and additional costs for technology up front, with the potential for long-term savings based on ease of gathering data to inform decisions once data is integrated.

1.1.2 Identify the data sets that, if integrated, would solve the need identified in activity 1.1.1 and describe how the data will be integrated.

For each use case, describe the specific data sources or collections that, if integrated, could address the needs or questions identified in activity 1.1.1. A short paragraph should explain the data source(s) in relation to the need statement and how the data could be integrated. A clear statement of the data to be targeted for integration is essential for accurately establishing the scope of the data integration project (Task 1.3).

1.1.3 Define the stakeholders and the impact of data integration on each of them.

For each use case, compile a list of the stakeholders who will benefit from integration and how they will benefit. Also identify those who may be impacted unwillingly.

1.1.4 Identify how different stakeholder groups might use integrated data

For each use case, identify how different stakeholders will use integrated data. List the current and proposed tools, applications, or systems stakeholders will use to access the data once integrated. The narrative should include:

- The stakeholders who will access the data;

- The benefits and impacts of the integration for these stakeholders; and

- Examples, if relevant, of the data they need to access and the tool, application or systems they will use to access the data (e.g., existing application, new application, database access, Generate, etc.).

1.1.5 Identify champions and challengers to data integration for each individual use case. (This does not have to be included in the written use case.)

Identify individuals or groups who will champion the data integration use case (list them by name where possible). Champions are those who will support or advocate on behalf of the proposed data integration effort. These champions can help vet the data integration requirements documentation, present the case for data integration to decision makers, and keep the project moving forward.

Identify individuals or groups who may challenge the need for data integration, its relative level of importance, or not understand how they will benefit from the effort. Identifying challengers will help you:

- Understand who may be adversely affected by integration, and how the team can mitigate these impacts or the perception of them; and

- Prepare documentation and communication to address potential objections to integration and the security processes in place to ensure data privacy.

Champions and challengers can include, but are not limited to SEA or local education agency (LEA) staff members, SEA leaders, policy makers, or parent privacy groups and advocacy groups.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
1.1.1 Identify the need statement and associated questions for the use case.	Project team	2-hour meeting
1.1.2 Identify the data sources that would need to be integrated to solve statements from 1.1.1 and describe, if possible, how the data would be integrated.	Project team	2-hour meeting
1.1.3 Identify the stakeholders and the impact of data integration for each individual use case.	Project team	2-hour meeting
1.1.4 Identify how the stakeholders will access the integrated data for each individual use case.	Project team	2-hour meeting
1.1.5 Internally identify champions and challengers to data integration for each individual use case. (This does not have to be included in the written use case.)	Project team	2-hour meeting
1.1.6 Internally identify champions and challengers to data integration for each individual use case. (This does not have to be included in the written use case.)	Project team	1 day
1.1.7 Review completed use case and determine additional use cases.	Project team	1 day

What Are the Lessons Learned From Previous Efforts?

- Regularly revisit the use case to ensure the data integration process is achieving the stated goals. Revise or develop new use cases as needed.

- Identify a champion or owner of the use case to maintain the documentation and ensure the use case(s) are reviewed and modified regularly (at least annually and when there are staffing changes).
  - » Embed reviewing the use case into a scheduled routine (e.g., as part of data quality or data governance routines) to keep the purpose of the project at the forefront of conversations.
  - » Revisit the use case when leadership, policy, data collection, data systems, reporting requirements, or priorities change, which may impact the efforts or focus on data integration.

What Are the Lessons Learned From Previous Efforts?

- Consider more vocal and influential stakeholders when identifying the project champions. They may be likely to influence the decision-makers because of their organization, role, or level of influence.

- Do not overthink it. This process is intended to be extremely high level in scope. The reader of the use case should be able to understand the purpose and impact of data integration; however, avoid getting too specific as it may overwhelm the intended audience, making them think the effort is too complicated.

- Consider using an outside facilitator to support this task; it is important that as many perspectives are included in this task as possible. Using a facilitator allows more perspectives to be included.

- Ensure that you develop a use case for each unique situation. Consider the perspectives of other stakeholders, other data that may also be integrated, or which use case may be highest priority.

Resources

- CIID Use Case Template
- CEDS Connections

Contact Us: By email at CIIDTA@aemcorp.com.
Visit the CIID website for more information at CIIDTA.org.

CENTER *for the* **INTEGRATION** *of* **IDEA DATA**

DATA INTEGRATION
TOOLKIT

The CIID Data Integration Toolkit

STEP 1 Articulate Use Case(s) for Integration

Task 1.2: Present use case to stakeholders, champions, and state education agency (SEA) leaders for support.

Overall Description of the Task

This task provides stakeholders and champions with the opportunity to review use case and suggest modifications. It also enlists their assistance in presenting the use case(s) to SEA leadership. The activities in this task may be iterative and you may repeat the tasks before finalizing the use case.

Why Is This Task Important?

- Provides SEA and other leaders with an understanding of the goals and benefits of integrated data.
- Helps to engage stakeholders, champions, and other supporters by soliciting their input on the scope of integration and offering insight into its purpose and benefits.
- Engaged stakeholders and champions can smooth the transition from planning to implementation.

Activities

1.2.1 Present the draft use case to identified stakeholders.

Schedule one or more in-person meetings with stakeholders to present and explain the goals and expected benefits of data integration. If in-person meetings are not feasible, this can be accomplished less formally through an email solicitation and document review process. The result should be to develop a common understanding and garner support from stakeholders and champions for the proposed data integration.

Possible outcomes of this activity include:

- Stakeholders disapprove of the proposed data integration project and the project ends;
- Stakeholder involvement reveals new use cases through the discussion and are documented;
- Stakeholders agree with the use case with or without revision.

1.2.2 Adjust use case to reflect stakeholder input.

Revise use case based on feedback from the stakeholders and project champions. If appropriate, also draft new use cases. Iterate activities 1.2.1 and 1.2.2 until the stakeholders and champions agree on a final use case. Share revisions with stakeholders for approval (i.e., return to 1.2.1) before presenting to SEA leadership (1.2.3).

1.2.3 Present use case to SEA leadership for support.

Schedule time with SEA leadership to present and explain the use case, gain support and, if required, formal approval to proceed with data integration. Leadership support is often necessary to make staff resources available for the project. If possible, include champions and stakeholders in the meeting and have them explain their need for the data integration project. Their involvement is most effective when leadership views them as an important voice. Meetings with leadership should be done in-person if possible.

1.2.4 Adjust use case based on SEA leadership feedback.

Revise use case based on feedback received from SEA leadership and share revisions with stakeholders and champions for approval before sharing with leadership for approval.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
1.2.1 Present the draft use case to identified stakeholders.	Project lead, stakeholders, project champions	1-hour meeting
1.2.2 Adjust and finalize the use case.	Project team	2-hour meeting
1.2.3 If needed, present the use case to leadership for support.	Project team, leadership, stakeholders, project champions	1-hour meeting
1.2.4 If needed, adjust use case based on leadership feedback.	Project team	1-hour meeting

What Are the Sustainability Considerations?

• As the data integration project progresses, keep SEA leadership apprised of the progress to keep the project visible and keep them aware of the project's value and the need to support it.

• Look for opportunities to leverage the use case to demonstrate capabilities and value of the project, along with a tangible display of why this project is important and the importance of support from leadership.

• When leadership changes, present the use case to the new leaders to inform them of the purpose, answer questions, learn their priorities, and garner support.

What Are the Lessons Learned From Previous Efforts?

• Be prepared to respond to leadership questions and requests for information about potential challenges for data integration. Understanding potential challenges that can have cost and resource consequence is a priority for state or SEA leaders.

• When meeting with champions or SEA leaders, encourage their feedback and emphasize a willingness to revise. If leaders can mold the use case to mirror or support their priorities they will be more likely to support integration.

Contact Us: By email at CIIDTA@aemcorp.com.
Visit the CIID website for more information at CIIDTA.org.
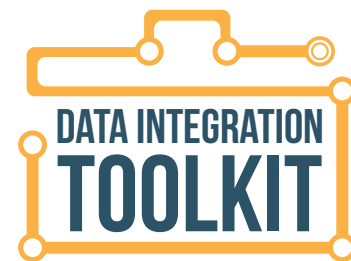
The CIID Data Integration Toolkit

STEP ① **Articulate Use Case(s) for Integration**

Task 1.3: Define the scope of the data integration project.

Overall Description of the Task

Based on the goals of integration and use case identified in Task 1.2, develop a scope document that defines the extent and limits of the project and guides the evaluation of the project.

Why Is This Task Important?

- A documented project scope defines the limits and extent of the project, and serves as a management tool to ensure the project goals are met.
- The project scope is the foundation for the project management plan, resource allocation, budgeting and prioritization of the work (covered in Step 2).
- A scope document supports common understanding of the project parameters, frames expectations and reduces scope creep.

Activities

1.3.1 Assess the current data integration environment.

Using the *CIID Part B Data Framework Self-Assessment* or similar tool, assess the current environment to support data integration including data governance policies and procedures, communication processes, the current alignment of data to be integrated, and staff resources available to support data integration.

1.3.2 Define the goals and exclusions of the data integration project.

Based on the use case from Task 1.2, identify the data integration project needs. Review the use case to identify what is excluded or outside of the bounds of the project. Exclusions may refer to data or reporting needs that will not be addressed by the integration project. Develop a "parking lot" of issues that are out of scope.

1.3.3 Identify assumptions that serve as the basis for project planning.

Describe the underlying assumptions that will be incorporated into the project planning process. If any of these assumptions turn out not to be true, it could adversely affect the project timelines or deliverables.

Include assumptions related to available resources, technologies, general timelines or pre-requisite work that must be completed prior to the start of the data integration project, such as:

- The amount of staff FTE available for the project;
- Specific technologies such as software and hardware tools, including the version;
- Time allotted for leadership decision making or sign off;
- Date upon which project work may begin; or
- Additional technology and resources needed by a specified date.

1.3.4 Identify major timing- and calendar-related implications for the data integration project.

Consider the timing and calendar issues related to the integration project. At a minimum consider the following integration questions:

- When will the data be available for integration?
- Will the timing and availability of the data affect current uses or processes, IT staff resources, or the proposed use of the data?
- What certification or validation process are in place and how do they impact timing of data integration?
- How frequently will the data be integrated (annually, quarterly, more frequently)?

1.3.5 Identify potential risks that may affect the project definition.

Identify known risks that could impact the project implementation. These may include potential risks that could impact funding, resources, or available technologies. Possible risks may include:

- Dependencies on specific technologies or outside resources;
- Software or hardware upgrades that are scheduled to occur during the project;
- Decision making processes that impact or delay the project schedule;
- Staff or resource vacancies that must be filled prior to the effort;
- Changes to data collected or collection requirements;
- Competing agency priorities that could affect available resources;
- Tight timelines required to meet desired project outcomes by a specific date; or
- Potential choke points related to staff expertise or technology constraints that could have an adverse effect on the project schedule.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
1.3.1. Assess the current data integration environment.	Project team, internal stakeholders	1-hour meeting
1.3.2. Define the goals and exclusions of the project.	Project team	30-minute meeting
1.3.3 Identify assumptions that serve as the basis for project planning.	Project team	30-minute meeting
1.3.4 Identify major timing- and calendar-related implications for the data integration project.	Project team	30-minute meeting
1.3.5. Identify potential risks that may impact the project definition.	Project team	30-minute meeting

What Are the Sustainability Considerations?

- Document decisions made relative to the scope and the reasoning leading to the decision so that if the composition of the project team changes, the same issues will not need to be revisited.

- Use a "parking lot" list to capture ideas for future enhancements or additions to the data integration project for consideration in a future phase.

What Are the Lessons Learned From Previous Efforts?

- If the scope document is not accurate or adhered to, resources may not be sufficient to complete the project or work performed may not meet the stated goals of the project.

- If underlying assumptions are no longer accurate, review the scope document to determine how that may affect the desired outcomes or milestones.

- Outline a process for addressing changes in scope to ensure all changes are intentional and agreed upon.

Resources

- CIID Part B Data System Framework Self-Assessment
- Project management plan template

DATA INTEGRATION
TOOLKIT

Step 2: Establish Project Plan and Structures for Data Integration Work

During Step 2, SEA staff develop a clear project plan to support data integration efforts. The project plan will include a project schedule, communication plan, and project team members, roles, and responsibilities. Completing Step 2 identifies the activities, timelines, and resources required for the data integration efforts.

Task 2.1: Establish project team and regular meeting schedule.
Task 2.2: Develop project plan and assign team responsibilities for data integration.
Task 2.3: Determine and secure required resources.
Task 2.4: Create resource repository to support integration efforts.
Task 2.5: Create communication plan for external and internal stakeholders.
Task 2.6: Determine how vendors and contracts will be included in all steps and tasks.

CIID
CENTER *for the*
INTEGRATION *of* **IDEA DATA**

Contact Us: By email at CIIDTA@aemcorp.com. Visit the CIID website for more information at CIIDTA.org.

The CIID Data Integration Toolkit

STEP ② Establish Project Plan and Structures for Data Integration Work

Task 2.1: Establish project team and regular meeting schedule.

Overall Description of the Task

The state education agency (SEAs) will identify a project manager with the authority to make decisions for the project. The project manager will then identify necessary expertise or skills required for successful implementation. The project team will be responsible for managing the SEA's data integration efforts and will meet regularly to discuss project progress, identify and resolve issues, and adjust the project plan and activities as necessary.

Why Is This Task Important?

- Establishing a project team to manage the implementation helps to ensure data integration is implemented on time and with quality.
- Establishing regular meetings provides the project team with scheduled opportunities to discuss project progress or delays, communicate information from stakeholders, and decide what actions are required to modify or improve the project plan.

Activities

2.1.1 Identify the project manager.

Identify who (one individual) has the authority to make project decisions and lead the project team. This individual will be a senior leader on the project or will report to a senior leader. The project manager (also known as the project lead) must understand the SEA's data integration plan, current data systems, and data governance processes. In addition, the project manager must also have experience in managing data coordination projects and a good working and trusting relationship with SEA leadership and staff.

2.1.2 Identify remaining project team members.

Data integration projects often require representation from a broad range of experts. Identify the most important areas of expertise and invite staff with this expertise to serve as part of the project team. Establishing the project team may require coordinating with managers in different parts of the agency to identify individuals with the necessary backgrounds or expertise. Limiting the number of team members to no more than 10 will improve personal accountability.

Areas of expertise for data integration may include:

- Statewide longitudinal data system (SLDS) master student data;
- Individuals with Disabilities Education Act (IDEA) datasets;
- Common Education Data Standards (CEDS);
- ED*Facts*; and
- IDEA program management.

Additional areas of expertise include:

- Vendor systems;
- Annual Performance Reports (APRs);
- Data privacy and security;
- Legal counsel;
- Database administration;
- Requirements analysis;
- Content area experts;
- Information technology (IT) team or information officer;
- Local education agency (LEA) or charter school expertise.

2.1.3 Review project scope and use case for integration.

Convene a project team meeting to review the project scope and the use case created in Step 1 to ensure team understanding of the purpose of the integration project. The purpose of this meeting is to create a common understanding of the integration goals of the project.

2.1.4 Schedule regularly occurring project team meetings.

Determine how often the project team meetings will occur (e.g., weekly, bi-weekly, monthly, quarterly) and establish a standard day, time, and location for project team meetings. These project team meetings allow team members to stay up-to-date with integration activities and make project decisions. Establish a regular routine and expectation for all project team members to attend and contribute.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
2.1.1 Identify the project director.	Initial project team members	30-minute meeting
2.1.2 Identify remaining project team members.	Initial project team members	30-minute meeting
2.1.3 Review the project scope and use case for integration.	Initial project team members	3-hour meeting
2.1.4 Schedule regularly occurring project team meetings.	Project director and project team	1-hour meeting

What Are the Sustainability Considerations?

- Consider what skillsets and roles may be needed for ongoing support and maintenance efforts for data integration.

- Refer to the use case(s) for talking points about the purpose and benefits of the data integration project that can be shared with new staff and team members.
- Develop documentation about the general expectation of project team members, including information to help new team members better assimilate into the project team.
- As you are initiating the project think about overall agency mission and how to fit the project into those structures to maintain project visibility.

What Are the Lessons Learned From Previous Efforts?

- Make sure all team members understand the purpose of the project. Though many of them may have helped create or finalize the data integration use case (Step 1), implement activity 2.1.3 to ensure that all staff understand the project's intent.
- Include team members from different departments or offices in the SEA, while preventing the total number of team members from becoming cumbersome. Do not be afraid to invite additional SEA staff to participate in team meetings when needed; however, keep the official project team small.
- Be aware of political relationships when establishing your project team and include representation from program offices across the SEA supporting these integration efforts.
- Gather information about the staff schedules and dates when staff may not work to ensure staff are available for regular team meetings.
- Establish meeting norms of behavior so team members are encouraged to collaborate and feel at ease sharing their thoughts.
- To maximize the efficiency of full project team meetings, convene subgroups of the team to accomplish specific objectives, such as agenda planning and document review, that doesn't require the whole group.

Resources

- Traveling Through Time: The Forum Guide to Longitudinal Data Systems, Book 2: Planning and Developing an LDS
- Seven Norms of Collaborative Work, Garmston & Wellman

Contact Us: By email at CIIDTA@aemcorp.com.
Visit the CIID website for more information at CIIDTA.org.

CENTER *for the*
INTEGRATION *of* **IDEA DATA**

DATA INTEGRATION TOOLKIT
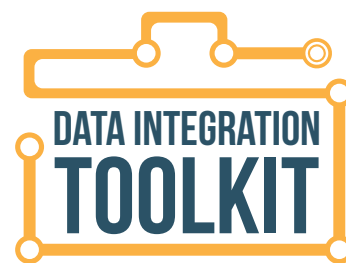
The CIID Data Integration Toolkit

STEP ② Establish Project Plan and Structures for Data Integration Work

Task 2.2: Develop project plan and assign team responsibilities for data integration.

Overall Description of the Task

The project lead will identify the project management routines, major activities of the data integration effort, as well as specific project team responsibilities to achieve the goal of integration identified in Step 1.

Why Is This Task Important?

- Establishing a clear project plan to support the data integration effort will help complete the effort in an efficient and timely manner.
- Establishing project team responsibilities allows project members to prepare for their roles and clarify what is expected of them before the project begins.
- Assigning individuals responsibility for specific activities based on their knowledge and skills ensures accountability for each task.

Activities

2.2.1 Develop a project management plan.

Using the CIID Data Integration Toolkit and any other applicable resources, identify the major steps or phases of the data integration project and their sequence. This list, typically based on the Toolkit steps and tasks, will serve as the basis for the project plan.

For each step, define its length, requirements to complete the step (what must be prepared or completed prior), and outcome. Identify start and end dates for each step keeping in mind the calendar information gathered in Task 1.3.4. In addition, identify a project team member responsible for each major step. This team member will serve as the individual responsible for developing a plan for the step, managing its implementation, and reporting on its progress to the project team.

Finalize a single document containing the project plan, timeline and team member assignments.

2.2.2 Assign responsibility for routine project management activities.

Identify the routine project management activities and assign specific team members responsibility for these activities. These include:
- Developing agendas and facilitating scheduled project team meetings;
- Documenting project team meetings and distributing meeting notes to team members;
- Serving as a point of contact with vendors, stakeholders or contract staff;
- Collecting regular status updates from the project team;
- Developing and maintaining a system for tracking integration component progress;
- Analyzing requests and issues identified by project members;
- Maintaining and updating the project plan; and
- Elevating relevant issues and decisions to state education agency (SEA) leadership.

2.2.3 Review and update the project plan and supporting documents.

The project team will provide feedback on the plan's feasibility, timing and appropriate assignments. Specifically, the project team will review and approve individual team member assignments for project plan components and ongoing project management responsibilities. The project plan should be reviewed and updated at regular intervals based on the progress to date.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
2.2.1 Develop a project management plan.	Project Lead	2-hour meeting
2.2.2 Assign responsibility for routine project management activities.	Project Lead	30-minute meeting
2.2.3 Review the project plan and supporting documents to ensure team understanding and support.	Project lead and select project team members	30-minute recurring meeting

What Are the Sustainability Considerations?

- Documentation is an integral part of a successful project. While creating the project plan, include tasks for documenting processes and activities.
- Identify what tasks, such as data dictionary maintenance and ETL development, are required for ongoing support and maintenance of the integrated data system.
- Develop a plan for how the ongoing support and maintenance tasks will be staffed post-integration.
- Consider how roles currently provided by vendor or technical assistance center staff will be supported in the future.
- Consider what steps need to be taken when roles and responsibilities change (e.g. how does the new person need to access documents, what needs to be shared with that person, what access needs to be provided for them).

What Are the Lessons Learned From Previous Efforts?

- Be specific when documenting the project plan. This plan will change many times throughout the project, however the more specificity that can be provided now, the better the team will understand and implement it.
- Use the CIID Data Integration Toolkit as a guide for major project activities. Consult the Toolkit to determine the potential expertise required to support the step.

- Be thoughtful when assigning responsibilities to team members. Recognize the political sensitivities as well as potentially overlapping responsibilities among team members.

What Are the Lessons Learned From Previous Efforts?

- Be specific when documenting the project plan. This plan will change many times throughout the project and should be updated regularly as needed. Including more specificity in the initial plan will give team members a better understanding of the project tasks and effort that will be required.

- Be thoughtful when assigning responsibilities to team members. Recognize the political sensitivities as well as potentially overlapping responsibilities among team members.

- Identify a single point of contact when working with technical assistance centers or vendors.

- Look for logical chunks of work to organize the project plan. For example, it may be helpful to organize the project plan based on data sets, so that all the tasks related to that data set are in one place, knowing it is possible that different activities may be occurring on two different data sets at the same time.

- Be realistic when setting timelines; this is not about identifying the shortest path, rather it is a tool to help track progress. Consider building in buffers. Missing timelines because it was drafted too aggressively can be demoralizing for the team. Conversely, getting work done ahead of time can help create momentum and excitement.

- After each project team meeting send summary notes with next steps, assignments, and timelines to participants and project team members or sponsors as appropriate.

Resources

- Project Management Institute
- Project management plan template

Contact Us: By email at CIIDTA@aemcorp.com.
Visit the CIID website for more information at CIIDTA.org.

DATA INTEGRATION TOOLKIT
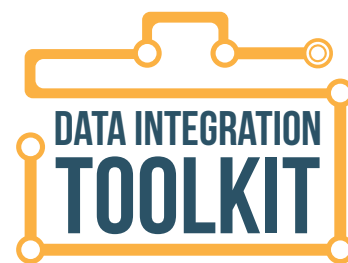
The CIID Data Integration Toolkit

STEP **2** **Establish Project Plan and Structures for Data Integration Work**

Task 2.3: Determine and secure agency resources required.

Overall Description of the Task

The project manager and senior leadership will establish a budget for the project, including estimated costs and secured funding source(s) in order to achieve project implementation.

Why Is This Task Important?

The budget for data integration will provide a structure for project planning and for managing personnel and resources.

Activities

2.3.1 Create a line item budget spreadsheet detailing the costs of the data integration project.

In consultation with fiscal staff, review the list of project activities and the personnel identified in Tasks 2.1 and 2.2 to get a realistic picture of the budgetary needs of the integration project. List each component of the project plan and its anticipated costs in a spreadsheet. Include costs required for staff time, travel, equipment, supplies, contractual support, and other costs. Document important fiscal deadlines and processes (e.g., budget request timelines, procurement procedures) to ensure continuity of funds. Designate which budget items are already funded (state education agency (SEA) staff currently with time available to help) and costs that will continue after the integration effort is complete.

2.3.2 Identify potential sources of funding or support.

Support can be provided in a number of different forms, staff time, information technology (IT), or financial resources. At a project team meeting, identify potential funding or resources to support data integration efforts. Establish a clear list of potential supports, documenting the estimated type and amount of their contribution. This allows the project team to understand the remaining resources and funding required. Support for the project can come from diverse sources, including:
- SEAs;
- State discretionary budget;
- Local funds;
- Federal funds;

- Private/philanthropic contributions; or
- Higher education source.

2.3.3 Develop integration resource memo and communicate with potential funding sources.

Develop a one- to two-page memo to distribute to potential sources of funding or support. The memo should describe why the integration effort needs to occur, including the use case established, how it will affect the SEA and its work, as well as the resources or support required. Once reviewed and approved by the project team, use this memo along with individual outreach to gain support for the project.

2.3.4 Review progress and adjust as needed.

Update team members on the progress of acquiring resources. Determine the remaining funds still required and adjust fundraising strategy or communication materials as needed. Consider cost savings and efficiencies such as automation and increased staff capacity, as the project progresses. Determine if the project budget, plan, or scope needs to be adjusted depending on the amount of funds or resources secured or required system changes.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
2.3.1 Create a line item budget spreadsheet detailing the costs of the project plan.	Project lead	1-hour meeting
2.3.2 Identify potential sources of funding or support.	Project lead	Over the course of 1 week
2.3.3 Develop integration resource memo and communicate with potential funding sources.	Project lead	Over the course of 2 weeks
2.3.4 Review progress and adjust as needed.	Project lead	2-hour meeting

## What Are the Sustainability Considerations?

- Identify funding and personnel needs for ongoing support and maintenance of the project after it has been implemented, including resources required for future expansions.
- Schedule a time with SEA leadership to ensure continued support beyond the scope of the initial project.
- Include any realized efficiency or cost savings when communicating with leadership. Consider creating a "turnover plan" if it is known that key positions change after a certain time frame (e.g., the superintendent may change every two years along with other key leadership).

What Are the Lessons Learned From Previous Efforts?

- Do not expect to get the budget exactly right; it will require estimation. The project team will not understand the full scope of activities until after this step. Establish a good understanding of the project plan requirements to ensure that all funding estimates are aligned with realistic cost estimates.
- Do not be afraid of selling the benefits of data integration. Gaining financial or other support for any project can be very challenging and at times feel unnatural. Do not let these feelings limit the team from reaching out to potential supporters who also would benefit from this work.
- Think creatively about how you approach potential supporters and the types of resources they may be able to provide. This project will require much more than just financial support. Consider the potential supporter and his/her expertise and resources—could the supporter provide a percentage of staff time, host the test environment, or help disseminate communication materials?

Resources

- CIID - Benefits of an Integrated Data System
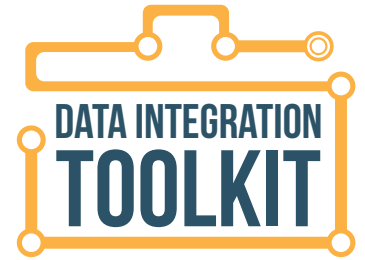
CENTER *for the* **INTEGRATION** *of* **IDEA DATA**

DATA INTEGRATION TOOLKIT

The CIID Data Integration Toolkit

STEP **2** **Establish Project Plan and Structures for Data Integration Work**

Task 2.4: Create a resource repository to support integration efforts.

Overall Description of the Task

Create a centralized electronic system for the purpose of updating project members on project issues; storing relevant documents for all project members to access, addressing project members' questions in an easily accessible location and format, and performing other functions that serve to inform all project members and centralize communication.

Why Is This Task Important?

- Creating a central resource repository maximizes efficiency by eliminating individual communications that would otherwise be needed to relay updates, important documents, and relevant information.
- A central resource repository ensures that all project members stay informed of important project updates, documents, and information.
- A central resource repository gives all project members access to common documents and resources and allows project members to track and archive different versions of the same document efficiently.

Activities

2.4.1 Establish the purpose of the central repository.

Define the project management needs and capabilities to be supported by the repository system. Project needs include:
- Uploading and disseminating documents;
- Maintaining version control;
- Acting as a discussion forum for project members' questions; and
- Tracking project plan status, risks, project decisions, etc.

2.4.2 Select a central repository tool or system.

- Identify a tool or system that meets the needs identified in activity 2.4.1. Consider solutions that may already be available and familiar to the project team members. Also consider the tool's usability, flexibility, and support of system users. The length and breadth of this selection process may vary depending on whether the SEA already has management/central repository software or not.

2.4.3 Provide login information to project members, along with some training and/or guidelines on central repository use.

Give each project member access to the central repository via an individualized login. Provide each project member with guidance on how to use the repository, whether through a live or electronic training, reading materials, or other formats that give users the opportunity for hands-on learning.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
2.4.1 Identify what project needs a central repository must meet.	Project lead, project team	30-minute meeting
2.4.2 Identify or create a tool or system to use as a central repository.	Project lead, tech team	30-minute meeting
2.4.3 Provide login information to project members, along with training and/or guidelines on central repository use.	Project team support	1 day

What Are the Sustainability Considerations?

- Consider the process for granting new project team members access to the resource repository.
- Periodically review who has access to the resource repository to ensure all current project members have access and that access is removed for former project members who no longer need access.
- Ensure that all project documents that may help retain institutional knowledge about the project goals, decisions, or processes are posted in the resource repository.

What Are the Lessons Learned From Previous Efforts?

- Be aware that your needs may change throughout the project and that the functionality of the central repository should be flexible enough to allow for this potential for change.
- Do not reinvent the wheel: use familiar systems where possible.
- Depending on the scope, budget, and needs of the project, the repository could be a simpler solution using a shared location on a network or cloud that is accessible by project members and a naming convention for version control.
- Be mindful of the organization of the repository such that documents are easy to find. This is especially helpful when new members are added to the project team.
- Encourage team members to send links to documents in the repository rather than attachments via email. This will encourage regular use of the repository.
- Choose a repository that all team members have easy access to.
- Create agreed upon document naming and document management conventions.

CENTER *for the* **INTEGRATION** *of* **IDEA DATA**

DATA INTEGRATION
TOOLKIT
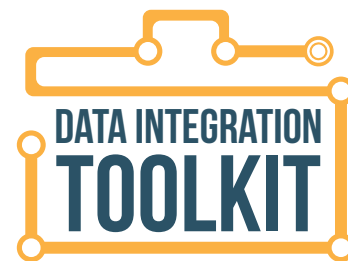
The CIID Data Integration Toolkit

STEP (2) Establish Project Plan and Structures for Data Integration Work

Task 2.5: Create a communication plan for external and internal stakeholders.

Overall Description of the Task

Develop a communication plan that informs and engages stakeholders throughout the integration effort as well as establishes decision making processes for issues requiring state education agency (SEA) leadership review.

Why Is This Task Important?

- This plan outlines the process and responsibilities you can use to ensure timely communication with internal and external stakeholders to inform them about project status, convey information about problems and other issues, enable them to raise issues and brainstorm solutions, and establish roles and responsibilities that prevent duplication of efforts.
- Some issues cannot be addressed or resolved without input from external governance or policy bodies. An "escalation process" enables the project team to forward matters that are not within its immediate control.
- A reporting process will ensure that internal stakeholders are aware of the timeline, create an accountability structure for all involved parties, and allow for mid-project corrections.

Activities

2.5.1 Identify the internal and external stakeholders who should receive communication about the integration effort.

Develop a list of stakeholders to engage. Include all organizations, users, and individuals affected by integration. When identifying stakeholder groups, consider how integration will affect them, what information they need to know (the purpose of the communication), and what type of feedback from them is necessary for the success of the project. Where possible, group stakeholders based on the type of information they should receive. For example, internal stakeholders like program teams require a different level of detail from external groups like local educational agencies (LEAs) and advocacy groups.

For each group of stakeholders, establish the mode for communicating, the appropriate level of detail to include, and how external stakeholders will provide feedback.

2.5.2 Establish a process to escalate issues to governance or policy bodies.

Identify an ongoing process for determining what topics or issues the project team can address on its own and which issues must be escalated to a governance or policy body with the authority and ability to address them. Establish a point of contact at each entity and a method for tracking and resolving escalated issues.

2.5.3 Identify specific communication needs within the project plan.

Review each step of the project plan, identifying specific instances where the project team should engage stakeholders for feedback or provide project updates. Specify what feedback is required and which stakeholders should be engaged.

2.5.4 Establish regular reporting processes for the project team.

Project team members should submit reports to project management at regularly scheduled intervals, using a standard format for consistency. Define the different reports project team members must submit, including the content, level of detail and audience. The reports could include:

- A bi-weekly status report or status update meeting, with notes, for tracking tasks to inform the project team of their status;

- Monthly reports on the progress of specific activities, including upcoming tasks or decisions to be made, intended for internal or specific external stakeholders;

- Quarterly reports that describe the integration progress overall, including a focus on tasks or issues that are not time sensitive or that are applicable to external audiences.

2.5.5 Define the responsibilities for managing and implementing project communication.

Identify the team member responsible for tracking the communication plan, identifying instances for engagement or reporting, and maintaining the contact information for internal and external stakeholders. Clearly document the responsibilities for communicating within project plan steps as well as regular reports from internal and external stakeholders clearly.

2.5.6 Distribute the communication plan.

Distribute the communication plan to project team for feedback and approval to ensure a common understanding of the communication plan, including project reporting and issue escalation. Periodically review, discuss, and revise the plan to ensure that it reflects changes to circumstances and new information acquired.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
2.5.1 Identify the internal and external stakeholders who should receive communication about the integration effort.	Project lead and select project team members	1-hour meeting
2.5.2 Establish a process to escalate issues to governance or policy bodies.	Project lead and select project team members	30-minute meeting
2.5.3 Identify specific communication needs within the project plan.	Project team members identified in the project plan	30-minute meeting
2.5.4 Establish the regular reporting processes for the project team.	Project lead and select project team members	30-minute meeting
2.5.5 Define the responsibilities for managing and implementing project communication.	Project lead and select project team members	30-minute meeting
2.5.6 Distribute the communication plan.	Project lead and select project team members	1-hour meeting

What Are the Sustainability Considerations?

- Periodically review the communication plan to ensure that internal and external stakeholders are being kept up to date.
- When developing the plan, consider the ongoing communication needs beyond the scope of the initial project, which may include leaving open the possibility of expansion of the project, additional funding, or resource requirements.
- Consider how the data integration project ties into larger agency initiatives or goals and how the ongoing support and maintenance of the implemented project to support those initiatives or goals.

What Are the Lessons Learned From Previous Efforts?

- If required by policy, consult with the SEA public information officer or communication manager.
- Differentiate your communication. Be thoughtful about which stakeholders should be engaged and to what degree. Some groups, including the project team, may require a high level of detail and consultation, where others will only need to know when the effort has been completed.
- Consider the mode and frequency of communication; some topics are more easily addressed through direct conversation, while email will suffice for other issues. Moreover, some people are easier to reach via email, while others prefer telephone contact.
- Communication methods should encourage two-way communication between internal project team members and stakeholders.
- Consider the political sensitivities within and outside of the SEA when determining the communication plan. Be careful to include all stakeholders and consider their needs and expectations when developing the plan.
- When developing a communication plan, consider where to include technical assistance centers that provide support to the state in project updates. Even if the centers are not directly involved with the data integration project, there may be opportunities for collaboration or their work may be affected by the outcomes of the project.
- As part of the communication plan, include a plan for regular updates to senior leaders (SEA and technical assistance providers).

Contact Us: By email at CIIDTA@aemcorp.com.
Visit the CIID website for more information at CIIDTA.org.

CENTER *for the* **INTEGRATION** *of* **IDEA DATA**

DATA INTEGRATION TOOLKIT

The CIID Data Integration Toolkit

STEP 2 Establish Project Plan and Structures for Data Integration Work

Task 2.6: Determine how vendors and contracts will be included in all steps and tasks.

Overall Description of the Task

The state education agency (SEA) will identify and review outstanding contracts with vendors and develop an understanding of the timing and important terms of these contracts in order to properly plan the integration of these systems.

Why Is This Task Important?

- Knowing how vendors and contracts support current data systems, their status, and dependencies between vendor systems will reveal valuable information that could influence an SEA's integration plan.
- Understanding vendor's expertise and role within current data systems allows the project team to identify where to involve them in the project plan, if required.

Activities

2.6.1 Document all relevant vendor contracts related to the integration scope.

Compile relevant documentation relating to all contracts that support data systems or collections identified in the project plan. Information for each contract should include vendor, purpose, timeline and expiration, services provided to identified data systems, services provided beyond integration scope, and dependencies on other systems or processes. It is also helpful to identify any political relationships associated with the contract that could affect vendor relations, communications, or timelines. Create a comprehensive document that identifies significant dates for all vendor contracts and a description of what happens during each important date to inform the project plan.

2.6.2 Identify impact of integration on all contracts and determine vendor engagement.

At a project team meeting, review the documentation of each contract and determine the impact the integration effort might have. Based on the impact of the contract, determine the need and process within the project plan for engaging the vendor. This engagement will range from simple communication of project progress, to data extraction, to redevelopment or modification of existing systems to accommodate newly integrated data.

2.6.3 Review plan for vendor engagement and a master timeline of significant dates.

During a project team meeting, review the planned engagement and anticipated impact on all vendors. Determine if vendor services and expertise are being properly leveraged to support integration efforts. Once approved, identify specific team members to serve as the primary contact for each vendor during the project. Consider the ongoing support and maintenance needs after project implementation and whether a long-term engagement is required.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
2.6.1 Document all relevant vendor contracts related to the integration scope.	Project lead, project team	30-minute meeting
2.6.2 Identify impact of integration on all contracts and determine vendor engagement	Project team	30-minute meeting
2.6.3 Review plan for vendor engagement and a master timeline of significant dates.	Project team	30-minute meeting

## What Are the Sustainability Considerations?

- Identify state staff who can partner with vendors and contractors to learn their processes to support the integration work after the contractors are gone.

- Develop a plan for transfer of knowledge from contractors or vendors to state staff, including documentation and training needs.

- Ensure staff are aware of state procurement procedures, who handles procurement, and necessary deadlines to secure resources for procurement staff. Include this documentation with the project work.

What Are the Lessons Learned From Previous Efforts?

- Use vendors as a resource. While a vendor's contract may have a well-defined scope, that does not mean their staff or services could not be beneficial to the integration project.

- Do not consider each contract in isolation. It is important to be aware of each individual vendor contract, while also being aware of how each vendor contract might impact other contracts.

- Consider the impact on current contracts but remain focused on the goals of the integration effort. The integration project will have a long-term impact on your SEA's data, do not let the priorities of current contracts or vendors dilute that impact.

- Knowing the appropriate channels for communicating with vendors and protocols can save time and frustration when interaction is necessary.

- Establish a single point of contact to interface with each vendor or contractor responsible for communication with the project team.

DATA INTEGRATION
TOOLKIT
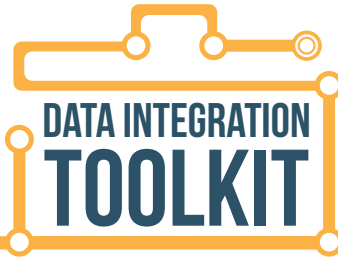
Step 3: Locate and Organize All Potential Data Elements and Associated Attributes for Integration into a Data Repository

During Step 3, the state education agency (SEA) staff is more involved with the technical aspects of the integration process by identifying the data systems and elements to be integrated and document the metadata associated with them. By completing Step 3, the project team documents data systems and elements to be integrated, allowing them to identify and resolve duplicate elements and/or discrepancies in Step 4.

Task 3.1: Identify datasets or systems to be integrated.

Task 3.2: Document, in the Common Education Data Standards (CEDS), all data systems and elements associated with the data integration effort.

CENTER *for the* **INTEGRATION** *of* **IDEA DATA**

Contact Us: By email at CIIDTA@aemcorp.com. Visit the CIID website for more information at CIIDTA.org.

CENTER *for the* **INTEGRATION** *of* **IDEA DATA**

DATA INTEGRATION
TOOLKIT

The CIID Data Integration Toolkit

STEP ③ Locate and Organize All Potential Data Elements and Associated Attributes for Integration Into a Data Repository

Task 3.1: Identify datasets or systems to be integrated.

Overall Description of the Task

In this high-level organization task, the project team assembles and reviews information on the data systems to be integrated. Note, beyond the identification of the elements, no work with data elements is required for this task.

Why Is This Task Important?

- Identifying the datasets or systems to be integrated sets the parameters for the project, loosely defined in Step 1.

- This task is necessary before identifying the elements that will be documented in Task 3.2.

Activities

3.1.1 Based on the scope identified in the use cases defined in Step 1, identify each data system and set that data integration will influence.

The project team will develop a list of data systems that will be affected by integration. The team will consider all systems that manage, collect, or report data that fall under the scope established in Step 1. For example, for datasets associated with school year, the team should identify which school years integrating the data will affect. Similarly, if datasets from prior years will be included, the team should determine which datasets will be back populated with integrated data.

3.1.2 Identify staff to lead documentation of information associated with each data system/set.

The project team will identify staff, likely the data steward, to document each data system/set. These staff members will compile the list of data elements contained within the data system/set. In Task 3.2, these staff members will document the data system/set.

3.1.3 Discuss and document reporting and process considerations of data integration.

The project team will determine the data steward for each data system/set. These data stewards will review each data system/set. Then, they should document each of the following items to inform future integration decisions:

- Federal and/or state requirements on any data integration;
- Timing of data collection;
- Data processing and reporting;
- Collection content;
- Collection processes;
- System where the data lives.

3.1.4 Review and update the scope of the project plan established in Task 2.2.

The project team, including data stewards, will review the scope and project plan established in Task 2.2, identifying changes and decisions made during activities 3.1.1–3.1.3. Changes to the project scope, timeline, or the resources required may be necessary based on the number of data systems and sets identified in activity 3.1.1 or additional considerations discussed during activity 3.1.3. These changes will be made to the original documents and redistributed to the project team for review and approval.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
3.1.1 Based on the use cases defined in Step 1, identify each data system and set that data integration will influence.	Project team and additional data stewards as necessary	1-3 1-hour meetings
3.1.2 Identify staff to lead documentation of information associated with each data system/set.	Project team and data stewards	1-hour meeting
3.1.3 Discuss and document reporting and process considerations of data integration.	Project team and data stewards	1-hour meeting
3.1.4 Review and update the scope project plan established in Task 2.2.	Project team and data stewards	1-hour meeting

What Are the Lessons Learned From Previous Efforts?

- Element-level decisions (e.g., regarding duplication, discrepancies) do not occur here, and no data element-level work is required in this task or in this step.
- Confirm the scope of data systems and dataset infrastructure that data integration may affect.
- Improve understanding of the integration effort. Affected data stewards, data governance structures, and administrators will become increasingly knowledgeable about upcoming data integration work and its implications.
- The assigned data steward should advocate for the data system to be integrated and bring any potential data integration issues to the team as soon as possible for consideration and resolution.
- Creating a data flow diagram to illustrate where data is collected, stored and moved for common uses can be helpful in describing the data environment for integration.

Resources

- Traveling through Time: The Forum Guide to Longitudinal Data Systems. Book Two of Four: Planning and Developing an LDS
- Data Flow Ideagraph tool

How Does This Connect to CEDS?

- 3.1.3: Review CEDS to see if any data dictionaries already exist as Align maps. If the data dictionaries do exist already, some of the discussions on data integration may be able to fast forward as you can run the maps side by side using the CEDS Data Dictionary Only report, to determine which elements exist in more than one system so you can see how they can integrate.
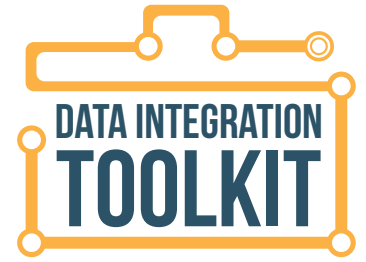
The CIID Data Integration Toolkit

STEP ③ Locate and Organize All Potential Data Elements and Associated Attributes for Integration Into a Data Repository

Task 3.2: Document, in the Common Education Data Standards (CEDS), all data systems and elements associated with the data integration effort.

Overall Description of the Task

The project team will organize all data system information into a document that includes detailed information on each element using the CEDS Align tool. This list will support efficient decision making about elements when data systems are integrated. Upon completion of this task, the project team will have created and organized a list of all elements with all element attributes to be used during element analysis in Step 4. No decisions about data elements are required for this task. Document any concerns or thoughts that surface for review in the next step (Activity 3.2.1).

Why Is This Task Important?

- Clearly identified data elements from all data systems targeted for integration will allow the project team to fully understand nuances of data elements in Step 4.

- This completed task allows the project team to identify elements with differences and redundancies so they can be addressed in Step 4.

- Organizing additional metadata about use of data, policies, procedures, schedules, etc. allows the project team to identify areas of data system support that they may need to address in Steps 6 and 7.

Activities

3.2.1 Document the relevant information about each data system/set and metadata in the CEDS Align tool.

Data stewards will compile the following metadata or information for each data system/set and its associated elements:

- Data policies (including federal and state requirements) for the data housed within the data system, including reporting date or period, will be added to the Comments section within an Align map in CEDS.

- Metadata information including:
  - » element name;
  - » definition;
  - » system/set name;
  - » database name;
  - » database table name;
  - » field type;
  - » character length;
  - » permitted values;
  - » validation logic (store in the comments section of the Align tool); and
  - » any other metadata information that is currently documented, and
- Data system manual and/or those procedures or policies associated with collection, cleaning, analysis, and reporting of data.

The project team, with data stewards and data governance committee, should agree to the common description categories prior to documenting.

3.2.2 Verify the list of data elements and descriptors.

The project team, including data stewards, will review the document to ensure all elements needed to address the scope as defined in Step 1 are included. The project team will also verify that the metadata and information are included and accurate for the system/set and elements.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
3.2.1 Document the relevant information about each data system/ set and metadata in the CEDS Align tool.	Data stewards	1 meeting per system
3.2.2 Verify the list of data elements and descriptors.	Project team, data stewards, and/or data governance committee	1-2 meetings

What Are the Sustainability Considerations?

Establish procedures and assign responsibility to someone, e.g., data stewards or those responsible for the data, to maintain CEDS documentation in order to keep it current. As federal reporting requirements change, new data elements will be added or removed, and keeping this information up to date helps to ensure the data remains clean and accurate.

Review and update your data dictionary on a regular basis, at least annually, reflecting any federal or state data element or definition changes. Each year ED*Facts* publishes the release notes specifying changes in ED*Facts* reporting requirements. These changes need to be reflected in the local data dictionaries and definitions.

What Are the Lessons Learned From Previous Efforts?

Ideally, a state should do the CEDS mapping for all data sets at the beginning of the project.

*   Compile metadata and information. Document any thoughts or discussions that occur, including where duplication may exist, and table for discussion during Step 4."

*   This task reveals what procedures and policies, if any, the project team may need to address to support data integration.

*   When populating the template, include information such as the source collection and data uses, where it may help with Steps 4 and 5. For example, staff assignments may be collected through a personnel collection and used by special education, for Common Core of Data reporting through EDFacts, and by teacher licensing and educator effectiveness. Knowing the collection and users of the data may help identify who needs to be consulted when reviewing for duplicate elements in Step 4.

*   State teams may be interested in focusing on only specific or prioritized groups of elements in order to speed up the implementation process; however, data integration encompasses all data elements so it is important to complete the alignment for all systems and data.

*   The CEDS tutorials are helpful for new users of the Align tool.

Resources

*   Hawaii Department of Education – Data Governance
*   University of Nevada Las Vegas – Data Governance Council Website
*   The Data Governance Institute
*   The Common Education Data Standards
*   CEDS Mapping Toolkit
*   CEDS Align Upload Template
*   CEDS Map Assignment Tracking & Validation Worksheet

How Does This Connect to CEDS?

*   3.2.2. If an SEA has uploaded and aligned data dictionaries in CEDS for the targeted source systems, it may run an Align Map-Based Report and sort on the CEDS data element name to look for commonalities.

*   The CEDS Align tool is used to document the metadata in this task.

Contact Us: By email at CIIDTA@aemcorp.com.
Visit the CIID website for more information at CIIDTA.org.

CENTER *for the* **INTEGRATION** *of* **IDEA DATA**   35

DATA INTEGRATION
TOOLKIT
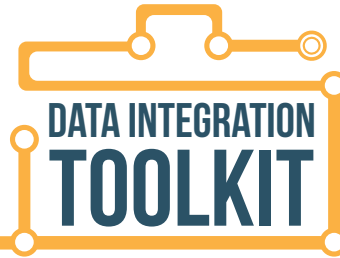
Step 4: Complete Master Integrated Dataset

During Step 4, the project team documents what the elements look like in the future integrated system by creating the master integrated dataset. Within this step, staff compare each element from the separate systems documented in Step 3 to determine the master integrated data set that will be the outcome of the future integrated system. By completing Step 4, the project team develops a master integrated dataset, including a list of redundant elements or changes that need to be made to current elements.
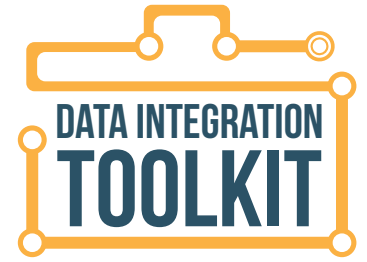

Task 4.1: Determine redundant data elements.

Task 4.2: Resolve element redundancies.

Task 4.3: Compile master dataset based on data dictionary for each dataset/system to be integrated.

CID
CENTER *for the*
INTEGRATION *of* **IDEA DATA**

Contact Us: By email at CIIDTA@aemcorp.com. Visit the CIID website for more information at CIIDTA.org.

CENTER *for the*
INTEGRATION *of* **IDEA DATA**

DATA INTEGRATION
TOOLKIT

The CIID Data Integration Toolkit

STEP ④ Complete Master Integrated Dataset

Task 4.1: Determine redundant data elements.

Overall Description of the Task

Identify potentially redundant data elements in the datasets/systems being integrated.

Why Is This Task Important?

- Carefully comparing and documenting redundant or similar element characteristics allows the project to determine which elements to eliminate, combine, or alter.

- Redundant data elements will cause confusion during data migration or extract, transform, and load (ETL) processes.

- Data queries and reports based on redundant data elements may not yield the expected results and may be very difficult or impossible to interpret correctly and accurately.

Activities

4.1.1 Generate a list of similar or potentially duplicative data elements by reviewing the list of all data elements created in Step 3.

The project team, with direction from the project lead and support from data stewards and other staff, reviews the list of elements to be integrated (from Step 3) and produces a list of data elements to review for potential redundancies. The project team reviews the name and definition of each element, placing each in a group of similar or potentially redundant elements. For example, data elements labeled "MI" or "Middle Initial" and "Student Middle Name" may be redundant. These elements should be listed in a single document where the team will make additional review comments. This review will help document the differences and similarities between the grouped elements.

4.1.2 Compare definitions and intent of grouped elements to identify and document differences in potentially duplicative elements.

Read through each definition in the groups of similar elements, highlighting the differences between/among definitions in the element list. For those with wording differences, note if they convey a similar intent with minor wording differences or differences in intent. Does one element cover a broader spectrum? Flag those elements that are worded identically. Ask the following questions:

- Are the elements for the same point in time?

- Do the elements have the same definition?

- Do the elements have the same intent, e.g., the difference between how race codes are captured at the point of the question vs. how it is reported federally?

4.1.3 Compare and document differences in code sets or permitted values, data type, and length.

For each group of similar elements, compare the code sets or list of permitted values, considering the questions listed below. In addition, review the data type and format for each data system and supported elements to be integrated. Note any potential issues that may arise between systems from the differences in data type or format, including the ability for them to be converted to a common format. Document any differences or duplication within the element list. Ask the following questions:

- Are they exactly the same?

- Do they have the same intent or options but use slightly different codes?

- Does one code set include a broader range of options? Is one more granular?

- When comparing the code sets, can you map individual values to those for another data element?

4.1.4 Review noted differences and potential duplications.

Review the grouped data element notes made during activities 4.1.2–4.1.3 with the project team to ensure consistency and verify the conclusions made.

4.1.5 Review and update the project plan and scope established in Task 2.2 and revised in Task 3.1.

The project team, including data stewards, should review the project plan and scope established in Task 2.2 and revised in Task 3.1, identifying changes needed to plan resulting from decisions made during activities 4.1.1–4.1.4. These changes will be made to the most up-to-date documents and redistributed to the project team for review and approval.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
4.1.1 Generate list of similar or potentially duplicative data elements by reviewing the list of all data elements created in Step 3.	Project lead, project team, data stewards	2 1-hour meetings*
4.1.2 Compare definitions of grouped elements to identify and document differences in potentially duplicative elements.	Project lead, data stewards	2 1-hour meetings*
4.1.3 Compare and document differences in code sets or permitted values, data type, and length.	Project lead, data stewards	2 1-hour meetings*
4.1.4 Review noted differences and potential duplications.	Project team, data stewards, additional staff with knowledge of collections	2 1-hour meetings
4.1.5 Review and update the project plan and scope established in Task 2.2 and revised in Task 3.1.	Project lead, data stewards	1-hour meeting

*Note: Systems with a large number of elements may require additional time

## What Are the Lessons Learned From Previous Efforts?

- Include subject matter experts or data stewards in this step because they may have knowledge that is not fully captured in documentation. They also understand the nuances of language differences, data quality issues, and state and federal reporting requirements.

- Use existing data governance organizations or practices to achieve the activities in this task.

- Do not leave any potential differences out during the comparison activities. Note the difference even if it is small or debatable. The project team can confirm if it should be noted in activity 4.1.4. Failing to note any differences could cause issues in future steps.

- Identify similarities as well as differences. This step helps identify what elements are appropriate to combine due to duplication as well as determine which elements need to remain separate due to the specificity of the element.

- When redundancies are identified, check if the elements in question are populated by the same source system. This can help you determine which system to use as the authoritative source.

- If integration is done in a series of phases, ensure you are revisiting redundancies across all phases of the integration process.

- If the team is doing this iteratively, or by process, consider beginning with more core/commonly used elements.

- By using Align "Map Based Reports" you can address Task 4.1.1 by generating a list of elements across maps to compare.

- Address Task 4.1.2 by comparing Align maps and elements using "Map Based Reports."

- Align data to CEDS to begin the process for working on Task 4.1.3, then use the "Map Based Reports" to compare and document any differences.

- When working on Task 4.1.4, export Align "Map Based Reports" to Excel then make notations on the documentation that already exists.

How Does This Connect to CEDS?

- See the CIID Brief The Use of Common Education Data Standards to Support Data Integration on how to use the CEDS tools to compare and evaluate elements in data systems.

Contact Us: By email at CIIDTA@aemcorp.com.
Visit the CIID website for more information at CIIDTA.org.

CENTER *for the* **INTEGRATION** *of* **IDEA DATA**

DATA INTEGRATION TOOLKIT
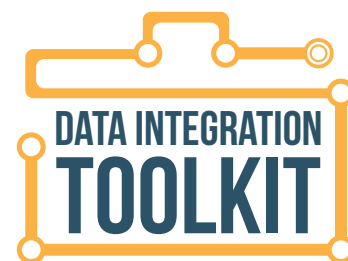
The CIID Data Integration Toolkit

STEP ④ Complete Master Integrated Dataset

Task 4.2: Resolve element redundancies.

Overall Description of the Task

Discuss and resolve redundant data elements for compilation into the master dataset for integration.

Why Is This Task Important?

- If redundant or very similar data elements are not resolved or carefully and thoroughly defined, documented, and renamed (if needed), then data elements may be unintentionally misused or misinterpreted.

- Redundant data elements will cause confusion during data migration or extract, transform, and load (ETL) processes.

- Data queries and reports based on redundant data elements may not yield the expected results and may be very difficult or impossible to interpret correctly and accurately.

Activities

4.2.1 Convene a meeting(s) with data stewards to review determinations made in Task 4.1.

Convene required data stewards to review the list of grouped data elements and the determinations made in 4.1. Additional team members should be included when data stewards require additional support. Consider holding multiple rounds or separate meetings to review the data in groups. During this meeting, review the determinations made in 4.1 to confirm the element groups and identified redundancies and identify any additional redundancies or similarities not yet documented. Also note any adverse impacts on data use or collection if redundant elements were to be integrated.

During this review, consider the significance of the identified differences and their impact. Determine if the differences can be accommodated by combining lists of permitted values. Identify what information might be lost if a potentially redundant data element was merged with a similar element. Consider if a merged data element would provide the necessary level of granularity for all program areas that use it. Note any data quality issues that may arise as well as any benefits to using one source element over the other.

4.2.2 Develop and review recommendations for redundant data elements that may require changes.

Based on the information review in 4.2.1, develop a list of data elements considered redundant that require changes in the newly integrated system. The recommendations will specifically list which data elements will be kept, which may be merged or modified, and which may be dropped based on the review by data stewards in 4.2.1. If two elements of the same name are determined to be different enough to keep both, provide a recommendation for name changes to indicate the differences between elements.

These recommendations should be reviewed and approved by the project team, as well as additional leaders within the state, as required by data governance policies.

4.2.3 Compile and confirm the master dataset for integration.

Compile the finalized master list of data elements, including the elements from Task 4.2.2, with all other data elements from the datasets/systems to be integrated into a master dataset for integration. This list will be used extensively throughout the remainder of the data integration process. Ensure it contains all elements to be integrated and describes them in their final form. This includes elements described in 4.2.2 that required modification. The intent of this list is to document the characteristics of the final integrated dataset for project team members.

Project team members and data stewards must review this list to confirm that all necessary data elements are included in the master dataset. Additional review or approval of this final compiled list by data governance bodies may be required.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
4.2.1 Convene a meeting(s) with data stewards to review determinations made in Task 4.1.	Project team, data stewards	2 1-hour meetings
4.2.2 Develop and review recommendations for redundant data elements that may require changes.	Project lead, project team, data stewards, other state leaders	3 1-hour meetings
4.2.3 Compile and confirm the master dataset for integration.	Project lead, project team, data stewards, other state leaders	2 1-hour meetings

What Are the Lessons Learned From Previous Efforts?

- Data stewards alone may not know everything needed to resolve redundancies among data elements. Include additional personnel as needed to supplement their expertise. Those who collect the data may not be the same as those who use it, so include multiple perspectives to anticipate and avoid any unintended consequences.

- Include program and technical staff in the tasks and decision-making processes listed above.

- Meet with data stewards and technical staff individually first to obtain initial feedback, then convene them as a group to finalize their recommendations. It is considered a best practice to perform a review with at least two additional people to ensure the quality of the work so that one person is not reviewing in isolation, and held responsible for all tasks listed above.

- Don't be discouraged if you encounter resistance to change. Be prepared to provide multiple opportunities for discussion and feedback to refine the recommended changes.
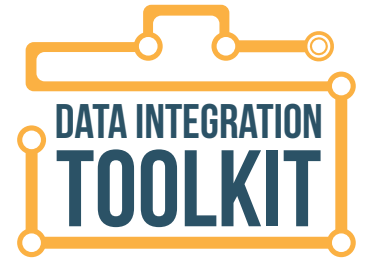
- The written recommendation prepared in Tasks 4.2.2 and 4.2.3 will be included as part of the project documentation and will affect the remainder of the integration process. Spend the time required to ensure it is clear and accurate. It will serve as a helpful reference in the future for how decisions were made and the intended outcome of the integration effort.

How Does This Connect to CEDS?

- For Task 4.2.2, consider using CEDS element names, definitions, or option sets when making modifications or changes.

- Create a new CEDS Align map for the final set of data elements and metadata determined in Task 4.2.3.

- CEDS provides the ability to view redundant elements side by side making it easier to determine discrepancies during Task 4.2.1. You can also export the data to Excel to manipulate and comment on it.

- For Task 4.2.3, use CEDS Align to create a new map of the final set of data elements and metadata. This can easily be done by exporting the existing Align maps and uploading into a single Align map using the Align tool upload wizard.

Contact Us: By email at CIIDTA@aemcorp.com.
Visit the CIID website for more information at CIIDTA.org.

42   CENTER *for the* INTEGRATION *of* IDEA DATA

CENTER *for the*
INTEGRATION *of* **IDEA DATA**

DATA INTEGRATION
TOOLKIT

The CIID Data Integration Toolkit

STEP **4** **Complete Master Integrated Dataset**

Task 4.3: Compile master dataset based on data dictionary for each dataset/system to be integrated.

Overall Description of the Task

Reconcile the data dictionary for each dataset/system to reflect the final decisions and updates made in Task 4.2 and use the revised data dictionaries to finalize the master dataset and the data dictionary for integration.

Why Is This Task Important?

- The data dictionaries for each of the datasets/systems being integrated will be used as the basis for creating the new data dictionary for the integrated dataset.

- The programmer or developer will use the updated data dictionaries and the finalized master dataset to ensure that all data elements needed from each of the source datasets/systems are correctly included in the integrated database and meet the needs and requirements of all program offices involved in the data integration process.

Activities

4.3.1 Reconcile the data dictionary for each dataset/system contributing to the data integration.

Ensure that all of the updated data elements and any other decisions made during the work completed in Task 4.2 are noted in the data dictionaries for each of the datasets/systems being integrated.

4.3.2 Finalize the master dataset and the data dictionary for the newly integrated dataset/system.

Present the master dataset for integration and updated data dictionaries to the project lead and agency leaders as determined by data governance policies. Obtain approval and sign off from all parties.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
4.3.1 Reconcile the data dictionary for each dataset/system contributing to the data integration.	Project team	3 1-hourmeetings, depending on the number of data dictionaries being updated
4.3.2 Finalize the master dataset and the data dictionary for the newly integrated dataset/system.	Project lead, department/ agency leadership	2 1-hour meetings

What Are the Sustainability Considerations?

- Identify all staff roles necessary for future maintenance of the dataset and ensure they are trained on the use of CEDS to maintain the work your team completes during the data integration process.

- Monitor changes to data definitions to identify new data elements to integrate.

## What Are the Lessons Learned From Previous Efforts?

- Document changes made. Including comment fields in the data dictionary and/or color coding changes made during the process of resolving redundancies in Task 4.2 allows for the communication of additional information to help data managers/stewards and subject matter experts and developers clearly understand any changes made to data elements and their relationship to the overall project.

- If data dictionaries for the legacy datasets/systems will continue to be used (even after the legacy dataset/ system is retired) in tandem with the new integrated dataset/system and its data dictionary, then ongoing maintenance processes will need to include the work to update those separate data dictionaries. For example, separate program areas such as special education (IDEA) and Title I/Accountability may decide to maintain separate data dictionaries, and those dictionaries will need to be included in the maintenance process for the integrated dataset/system.

- CEDS Align can be used to create a data dictionary when one does not exist.

How Does This Connect to CEDS?

- To address Task 4.3.1, you can export the existing Align maps and upload any changes to the consolidated Align map.

- Finalize the master data set using CEDS Align. Task 4.3.2 is a good time to share the map with your state team.

Contact Us: By email at CIIDTA@aemcorp.com.
Visit the CIID website for more information at CIIDTA.org.

DATA INTEGRATION TOOLKIT
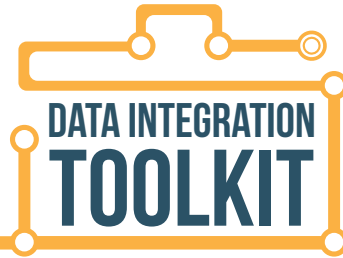
Step 5: Implement and Perform Extract, Transform, and Load (ETL) Procedures

During Step 5, the project team integrates the actual data identified in previous steps by pulling the data from the original systems and moving it over to the newly integrated system. By completing Step 5, the project team establishes a fully integrated data system, ready to be checked for quality and accuracy within Step 6.

Task 5.1:　Develop/update ETL target database schema.

Task 5.2:　Identify all program business requirements needed for ETL(s).
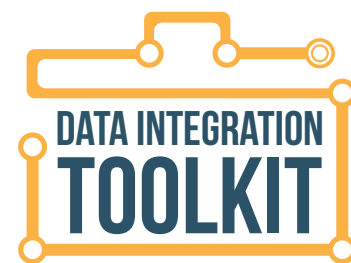
Task 5.3:　Create an ETL requirements document for the programmers.

Task 5.4:　Conduct pilot/test cases for ETL and validate data load with the program team.

Task 5.5:　Perform ETL.

CENTER *for the* **INTEGRATION** *of* **IDEA DATA**

CENTER *for the*
INTEGRATION *of* **IDEA DATA**

DATA INTEGRATION
TOOLKIT

The CIID Data Integration Toolkit

STEP ⑤ Implement and Perform Extract, Transform, and Load (ETL) Procedures

Task 5.1: Develop/update ETL target database schema.

Overall Description of the Task

Determine the necessary target database schema changes required to perform the ETL procedures. This could be the complete creation of a new target database or the updating of an existing target database.

Why Is This Task Important?

This task ensures that all elements have a defined location in the target database during the ETL process.

Activities

5.1.1 Verify destination location for each target element.

Using the list of elements defined in Step 4, verify the location in the target database where the ETL process will load the data. This task ensures that all mapped elements have a location to go to in the target database during the ETL process. Without this established schema, the technical team would not know where the elements identified in Step 3 would connect to the new target database.

5.1.2 Review impact on existing applications.

If the data integration plan involves the modification of an existing data store rather than the creation of a brand new location, changes to the target database may impact existing users and applications of the target database. Create a list of applications, reports, and/or processes impacted by the modification of the target database and determine required enhancements to ensure a seamless transition for users.

5.1.3 Update/create the target database schema documentation.

Update or create any target database documentation and models any time changes are identified in the previous activities (5.1.1 and 5.1.2). This may include:

• Data Models;

• Database definition documents; and

• Database creation scripts.

Apply the identified updates to the ETL target database.

What Is the Timeline for This Task?

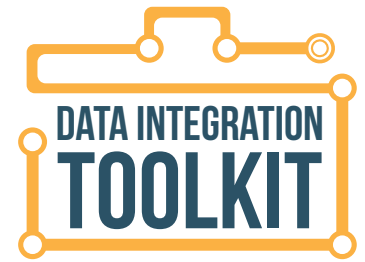| Activities in the Task	Participants	Duration
5.1.1 Verify destination location for each target element.	Database Administrator (DBA)	1 all day meeting
5.1.2 Review impact on existing applications.	DBA, Data Owners	1 all day meeting
5.1.3 Update/create the target database schema documentation.	DBA	1 all day meeting

What Are the Lessons Learned From Previous Efforts?

• Do not be afraid to raise questions. There may be errors or confusing target elements not identified during previous steps. Ask fellow project team members for clarification as the work from Steps 3 and 4 have not been fully implemented or reviewed.

Resources

• **ETL Checklist**

CENTER *for the*
INTEGRATION *of* **IDEA DATA**

DATA INTEGRATION
TOOLKIT

The CIID Data Integration Toolkit

STEP ⑤ Implement and Perform Extract, Transform, and Load (ETL) Procedures

Task 5.2: Identify all program business requirements needed for ETL(s).

Overall Description of the Task

This task focuses on gathering requirements for business needs used to inform the extraction, transformation, and load (ETL) process to integrate data from the source into the target database in the new environment.

Why Is This Task Important?

- This task ensures that those who rely on and use the data have a voice in identifying the requirements for developing ETL procedures and ensures all business needs related to ETL processing for data integration are captured and documented.
- Meetings with program staff to discuss requirements increase stakeholder engagement.
- Detailed ETL requirements become the basis for developers to write code and for quality assurance testing.

Activities

5.2.1 Gather existing requirements from source system documentation.

Based on the scope of the project established in Step 2, and the master data element and authoritative source determinations defined in Step 4, gather all existing documentation related to business requirements. This should include:

- equirements for producing current reports for internal use or public reporting;

- Processing rules and procedures for creating submission files for ED*Facts* and other required reports; and

- Commonly produced data files or reports in support of initiatives or programs.

5.2.2 Create preliminary ETL requirements document for target database.

After verifying the business needs in activity 5.2.1, draft preliminary ETL requirements based on the proposed data model for the integrated system. Typically, an ETL requirements document is created for each table in the target database.

5.2.3 Meet with the data manager or data stewards to discuss business needs.

Hold meetings between the data manager or data stewards and technical staff to discuss the business needs related to the use of the integrated data. During these meetings:

* Discuss what data are required to meet current business needs, when they are needed, how they should be formatted, and who needs access to them;

* Review documentation gathered in activity 5.2.1 to ensure that it is up to date and accurately reflects current requirements; and

* Discuss current or future business data requirements that are not currently being met. This may include new reports for internal use or public reporting, as well as data needed to support decision making related to current or future initiatives.

5.2.4 Verify ETL requirements with data stewards.

Provide a copy of the ETL requirements document to data stewards and additional staff who have a good understanding of the business logic. The document should include a place for comments to allow reviewers to provide feedback in the document itself. Review each data element in the target database and the proposed logic; and verify any assumptions made. Determine if there are exceptions or nuances that have not been adequately captured. Data stewards should identify if any additional concerns have not been addressed and verify that the proposed logic makes sense.

5.2.5 Revise requirements based on feedback.

Revise the ETL documentation based on feedback from the data stewards and validate as needed. Activity 5.2.4 may require several iterations, depending on the complexity of the requirements.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
5.2.1 Gather existing requirements from source system documentation.	Project team members, including data steward familiar with data sources and analysts	1-2 all day meetings*
5.2.2 Create preliminary ETL requirements document for target database.	Analysts	3-5 all day meetings*
5.2.3 Meet with the data manager or data stewards to discuss business needs.	Analysts, data stewards	3-5 all day meetings*
5.2.4 Verify ETL requirements with data stewards.	Analysts, data stewards	1-2 all day meetings*
5.2.5 Revise requirements based on feedback.	Analysts	1 all day meeting

Note: The duration of these activities will vary based on the project scope.

## What Are the Lessons Learned From Previous Efforts?

- When developing and reviewing the requirements, focus on one topic or table at a time (e.g., child count, discipline, personnel, etc.). Bring in subject matter experts from across the agency who use those data.

- In order to make best use of the meeting time between data stewards and analysts, it helps if the analyst is familiar with available requirements before the meetings. Consider creating a list of questions or observations that may be good starting points for the conversation or that may help keep conversations on track.

- Prioritize requirements. This can help with planning and determining resource needs. Prioritization can also help identify possible phases for implementation, if needed.

- Carefully document decisions made during the review and revision period. This record can be invaluable in the future when questions arise about the requirements.

Resources

- **ETL Checklist**

Contact Us: By email at CIIDTA@aemcorp.com.
Visit the CIID website for more information at CIIDTA.org.

![CIID - CENTER *for the* INTEGRATION *of* IDEA DATA]

![DATA INTEGRATION TOOLKIT]
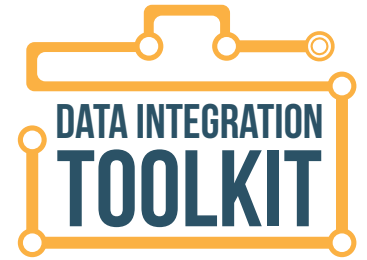
The CIID Data Integration Toolkit

STEP ⑤ Implement and Perform Extract, Transform, and Load (ETL) Procedures

Task 5.3: Create an ETL requirements document for the programmers.

Overall Description of the Task

As programmers prepare to write the code needed to extract, transform and load (ETL) the integrated data, they will need clear and complete requirements on where the data reside, the rules of transforming and loading, and rules for validating the process.

Why Is This Task Important?

- Requirements documents will provide the programmers the information needed to successfully code ETL.
- For a successful ETL, an accurate and timely process needs to be executed, with the appropriate validation steps at strategic points in the process. A requirements document can provide information needed for those validations.
- Data integration is only as good as the accuracy of the ETL.

Activities

5.3.1 Identify the specific data elements that will be integrated.

Based on the original element documentation produced in Steps 3 and 4 and the program requirements in Task 5.2, identify the individual data elements that will support the purpose and functionality required for data integration. For each data element, perform the following actions:

- Verify the source of the data element, including database, table names, source attribute or column;
- Confirm the data element type and length;
- Describe the data element's destination table and attribute or column;
- Determine the years/semester/terms, etc. that the data are available;
- Allow for appropriate access to the data;
- Define any edits that will take place on the extracted data elements;

- Define a process for responding if records are rejected because of accuracy or validity issues;

- Define any edit reports that will be used for extract validations; and

- Define the format of the extracted data elements that will be used for transformation.

5.3.2 Define and document transformation rules.

Transformation rules are used to foster the integration of data from multiple sources by resolving differences in data definitions and code sets. Create specific rules for transforming or cleaning the data in order to provide the programmer with necessary information for putting the code in place to make these transformations. Use the following steps for defining transformation rules:

- Determine the input format for the transformation;

- Determine the business rules for transforming data from the original source format to the transformed format;

- Define a process to capture the pre- and post-transformation data for validation purposes;

- Define any edit reports to be used for transformation validation;

- Determine how records with errors will be handled;

- Define the format of the transformed data elements that will be used for loading the integrated data; and

- Document other data elements that may be integrated during transformation.

5.3.3 Describe the load process.

Define processes that describe the database and tables where the data will be loaded. Make sure these processes include accompanying documentation or metadata that support any cross validation that may be needed. These processes should help confirm that all the expected data were loaded. Use the following steps for defining the load process:

- Document where the data are to be loaded;

- Identify any audit trail data that will need to be captured as the data are loaded; and

- Ensure that proper database backup and recovery procedures are documented and referenced for execution before the ETL takes place.

5.3.4 Compile the ETL requirements document.

Compile the information from activities 5.3.1-5.3.3 to create an updated ETL requirements document. Ensure the information is sufficient and includes the appropriate technical details so the programmers will be able to write and test the code to efficiently and accurately integrate and load the needed data. Document the sign off process confirming the accuracy of the data as loaded when compared to its original state in the source database.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
5.3.1 Identify the specific data elements that will be integrated.	Project lead, business analysts, programmers	5-7 all day meetings*
5.3.2 Define and document transformation rules.	Project lead, business analysts, programmers	3-5 all day meetings*
5.3.3 Describe the load process.	Project lead, business analysts, programmers	1-2 all day meetings*
5.3.4 Compile the ETL requirements document.	Project lead, business analysts, programmers, data stewards	3-5 all day meetings*

Note: The duration of these activities will vary based on the project scope.

What Are the Sustainability Considerations?

• Complete an ETL checklist for each data collection that is integrated before running the ETL. Make sure that both IT and program staff understand the checklists and know where they are stored.

• Assign maintenance of ETLs and ETL checklists to a position title and document that assignment.

What Are the Lessons Learned From Previous Efforts?

• Be careful to apply business rules accurately.

• Have the necessary documentation and ETL statistics in order to know the transformations and loading are done correctly. If the data are not accurately transformed correctly, integration and loading errors can occur.

• Provide sufficient detail so that the ETL programmer knows the specific data to extract, its location, precise specifications on the process of transforming the data, directions on where that data are to be loaded, and instructions for handling errors (both data errors and process errors).

• Execute proper database backup and recovery procedures before running the ETL.

Resources

• **ETL Checklist**

Contact Us: By email at CIIDTA@aemcorp.com.
Visit the CIID website for more information at CIIDTA.org.

CENTER *for the* **INTEGRATION** *of* **IDEA DATA**

53

The CIID Data Integration Toolkit

STEP (5) Implement and Perform Extract, Transform, and Load (ETL) Procedures

Task 5.4: Conduct pilot/test cases for ETL and validate data load with the program team.

Overall Description of the Task

A series of pilot/test cases will help confirm the ETL accurately processed the data and will help identify errors and needed corrections.

Why Is This Task Important?

- Test cases provide a mechanism for determining the completeness and accuracy of the ETL and allow stakeholders to help validate the accuracy of the ETL process.
- Test cases provide a documented process for managing errors.

Activities

5.4.1 Identify the specific data elements that will be integrated.

Stakeholders, program staff, and source data stewards should inform development of realistic test cases and review subsequent results.

- Identify stakeholders and business analysts to assist in identifying test cases to verify the accuracy of the ETL.
- Determine best means for communicating with those identified to review test cases and test results.
- Make sure expectations and time requirements are understood and they agree to be available.

5.4.2 Identify test cases.

Tests on the accuracy of each ETL should start once it is complete so the results can inform future work. To identify test cases for verifying the accuracy of the data load, include the following steps:

- Evaluate the various steps during ETL, specifically where data are transformed, cleaned, or loaded; and
- Identify the test cases based on the points where data transforms or moves.

5.4.3 Describe the load process.

Documentation should include what gets tested and whose responsibility it is to test it. This can be in the form of a list or "script," describing the test cases. The test case document can also include transformation rules. Once documentation is complete, ensure the project lead and other leaders as needed review and approve the test cases.

5.4.4 Establish documentation and a process to execute each test case.

Create documentation that clearly establishes the process for executing each test case. This document must include:

- The timing for executing the test;
- A test environment that accurately reflects the characteristics of the production environment;
    - » Test environment is typically 10% of production;
    - » Typically does not contain personally identifiable (PI) data;
- PI data have been modified with dummy data that still pass the edits;
    - » Has the same database structures and definitions as production;
    - » Test data have been validated and have been altered to "initiate" each test case at least once;
- The name of the person who needs to be available to monitor the test;
    - » This includes the business analyst, data system owners, and end users;
- The criteria for success;
    - » This is based on the requirements documentation; and
- A report for capturing the results for review by stakeholders.

5.4.5 Conduct initial testing, review results, and document needed corrections.

After executing the initial test, provide reviewers with the test results. These reviewers will determine if an error has occurred and will give instructions on how to reconcile the error. Establish feedback mechanisms for communicating any data accuracy issues with the source data system stewards and project team. These should be delivered immediately to the business analyst responsible for the ETL following the stakeholder review. The business analyst will compile the results and provide them to the ETL programmers for follow-up and correction. Once corrections are made, the ETL programmers will rerun the test case. Repeat this process until the expected test results are executed.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
5.4.1 Identify the specific data elements that will be integrated.	Project lead, database administrators	1 all day meeting
5.4.2 Identify test cases.	Project lead, analysts, database administrators, stakeholders, programmers	5-7 all day meetings*
5.4.3 Describe the load process.	Project lead, analysts, database administrators, other leaders as needed	3-5 all day meetings*
5.4.4 Establish a process to execute each test case.	Project lead, analysts, database administrators	5-7 all day meetings*
5.4.5 Conduct initial testing, review results, and document needed corrections.	Project lead, analysts, database administrators, stakeholders, programmers	3-5 all day meetings*

Note: The duration of these activities will vary based on the project scope.

What Are the Lessons Learned From Previous Efforts?

- Be careful not to overlook the importance of support staff responsibilities. Identify instances when any project member may need extra assistance or coordination is integral to efficiency.

- Ensure that every error condition is tested via one or more test cases.

- Reviewers of the test results should have familiarity with the data and purpose of the data.

- Define the feedback process between those reviewing the test results, the business analysist who will document test cases that did not perform as expected, and the ETL programmers who will need to make corrections and rerun the test.

Resources

- **ETL Checklist**

Contact Us: By email at CIIDTA@aemcorp.com.
Visit the CIID website for more information at CIIDTA.org.

CENTER *for the*
INTEGRATION *of* **IDEA DATA**

DATA INTEGRATION
TOOLKIT

The CIID Data Integration Toolkit

STEP ⑤ **Implement and Perform Extract, Transform, and Load (ETL) Procedures**

Task 5.5: Perform ETL.

Overall Description of the Task

Run the ETL processes to transition the data from a source system(s) to a target system (e.g., source system to an integrated data warehouse).

Why Is This Task Important?

- The ETL is the process that integrates the data.
- In order to protect the integrity and security of production data, all tasks performed during (and following) an ETL process should be done in the test environment first (see Task 5.4) to eliminate errors or bugs that could have a negative impact on "real" data.

Activities

5.5.1 Define ETL schedule.

As ETLs are executed, source data system performance and access can be affected. Reach agreement with source system data owners on the best time to run the ETL and establish a schedule. This decision will be influenced by source system utilization, regularly scheduled source data system maintenance, and amount of data to extract.

5.5.2 Perform ETL.

To perform the ETL, plan the execution according to the established schedule (determined during activity 5.5.1). Ensure source database(s) availability and notify the appropriate staff of the upcoming impact on the source system. Ensure target system availability as well, including any database maintenance in preparation for the new data being loaded. Once this is done, execute ETL.

5.5.3 Confirm ETL completion.

Once ETL is complete, confirm that the process worked as expected. Review ETL scripts and confirm the ETL's success. Evaluate any records rejected due to edit errors to determine their resolution and potential inclusion in the target system. Engage reviewers, including data stewards, to evaluate loaded data in the target system to confirm accuracy of data and record counts.

What Is the Timeline for This Task?

| Activities in the Task	Participants	Duration
5.5.1 Define ETL schedule.	Project lead, database administrators, programmers	1 all-day meeting
5.5.2 Perform ETL.	Database administrators, programmers	1-2 all-day meetings
5.5.3 Confirm ETL completion.	Database administrators, programmers, stakeholders	1-2 all-day meetings

What Are the Lessons Learned From Previous Efforts?

- Have a standard process for evaluating the load and that verifying record transformations and counts are correct in the target data system after each ETL is executed.

Resources

- **ETL Checklist**

CENTER *for the* INTEGRATION *of* IDEA DATA

DATA INTEGRATION TOOLKIT
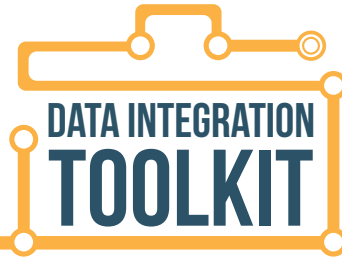
Step 6: Conduct Review and Validation of Data Integration

During Step 6, the project team checks the new system loaded in Step 5 for quality and ensures it meets the intended outcomes of the integrated system. By completing this Step, the project team validates the data integration's quality, gains final approval from state leaders, and finalizes documentation to sustain current and inform future integration efforts.

Task 6.1: Document and perform data quality checks.

Task 6.2: Secure final sign off on overall process.

Task 6.3: Review integration processes and plans to sustain integration.

CENTER *for the* **INTEGRATION** *of* **IDEA DATA**

The CIID Data Integration Toolkit

STEP 6 — Conduct Review and Validation of Data Integration

Task 6.1: Document and perform data quality checks.

Overall Description of the Task

This task ensures that integrated data adhere to established business rules and requirements. Documenting and performing these data quality checks will verify the completeness and accuracy of data after they have moved from one system to another.

Why Is This Task Important?

The extract, transform, and load (ETL) process in Step 5 applies code to move and transform data elements, option sets, and values from one data store to another. Thorough review of the data for accuracy and completeness is important considering the volume and complexity of the process. If left unchecked, the data reports could be incorrect.

Activities

6.1.1 Create and review documentation of data quality check process.

Create the documentation for the data quality check process in collaboration with the technical team and include the specific processes for review, individuals performing the validation and testing, and how to report validation and testing results. This documentation should cover:

- Process overview,
- Test strategy,
- Test plan,
- Test cases, and
- Issues log.

The documentation should be reviewed by all members of the project team. Project team members who have specific testing responsibilities should read the document thoroughly to be sure they understand the role for which they are responsible. Provide team members with a deadline for reviewing documentation and require sign off to keep the project on schedule. This documentation should be reviewed and updated annually to account for requirements updates or system changes.

6.1.2 Perform data quality checks.

Based on the approved documentation, perform data quality check tasks. Ensure all issues identified through each portion of the testing process are logged.

6.1.3 Review issues log and resolve issues.

Determine a schedule for addressing issues. Issues can be addressed and resolved as they occur or at a later date during a review with the project team. Identify and fully document each issue, its resolution, and the specific team member(s) involved. Issues may require modifying ETL steps within Step 5, edits to the master dataset in Step 4, or other edits. Once all issues have been addressed and data quality tests have been repeated without any outstanding issues, the sign off process may begin in Task 6.2.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
6.1.1 Create and review documentation of data quality check process.	Developer, requirements analyst, tester, data steward, project team members	2-5 all day meetings*
6.1.2 Perform data quality checks.	Developer, tester, data stewards	5-12 all day meetings*
6.1.3 Review issues log and resolve issues.	Project team, developer, tester, data stewards	2-10 all day meetings*

Note: The duration of these activities will vary based on the project scope.

What Are the Lessons Learned From Previous Efforts?

- Assign the right people to the right tasks. Always have someone familiar with a particular set of data do the testing, validation, or comparison as they are more likely to identify inconsistencies or errors.

- Include both technical and non-technical personnel in data quality tasks.

The CIID Data Integration Toolkit

STEP 6   Conduct Review and Validation of Data Integration

Task 6.2: Secure final sign off on overall process.

Overall Description of the Task

Conduct a final signoff of the overall data integration process when all parties agree that it is complete.

Why Is This Task Important?

It is important to formally conclude the data integration process with all key stakeholders involved; therefore, they should agree that the integration is completed before signing off on the final process. This sign off indicates that the SEA has approved the integration results and allows team members to begin developing communication documents.

Activities

6.2.1 Identify approvers.

Determine the key approvers who will need to sign off on the project. This will include the project team as well as other SEA leaders. Ensure all major user groups are represented.

6.2.2 Compile final sign off documentation.

Compile the necessary documentation to review the integrated data systems and updated data dictionary, established previously in Steps 4 and 5. Additional documentation should also demonstrate that all issue log items from Task 6.1 have been resolved and summarize the ongoing requirements to maintain and improve the data integration.

6.2.3 Obtain final sign off.

Obtain the final sign off documentation during a project team meeting with all identified approvers present. Any edits made during this review process should be communicated to the project lead. Once any edits are resolved, distribute the final version of the documentation from Task 6.2.2 to all project team members.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
6.2.1 Identify approvers.	Project lead	1 all-day meeting
6.2.2 Compile final sign off documentation.	Project lead, project team	2 all-day meetings
6.2.3 Obtain final sign off.	Project lead and approvers	1 all-day meeting

What Are the Lessons Learned From Previous Efforts?

Carefully select the approvers required during this step. Data users and SEA leaders should be represented. This sign off is focused on approving the completed integration process that was originally established in Steps 1 and 2. Future steps will help describe the process to external groups.

Contact Us: By email at CIIDTA@aemcorp.com.
Visit the CIID website for more information at CIIDTA.org.
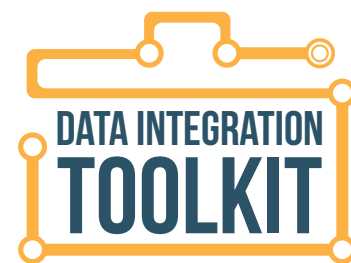
The CIID Data Integration Toolkit

STEP 6 — Conduct Review and Validation of Data Integration

Task 6.3: Review integration processes and plans to sustain integration.

Overall Description of the Task

Perform structured review of the processes involved in the entire integration process and implement follow-up activities and processes designed to sustain the integration effort and improve future integration projects.

Why Is This Task Important?

Data integration will likely happen again, and it is important to review the processes used during the entire data integration project to ensure future efforts are carried out as effectively and efficiently as possible. Tasks to sustain the integration are also important as they are designed to ensure the integrity of the integrated data and maintenance of the systems.

Activities

6.3.1 Conduct process review meeting annually.

Meet with the project team no later than two weeks after Task 6.2 has concluded to review all major processes, including those from Steps 3-6. Multiple meetings may be required to review all processes. During this meeting, the project team will review the documented processes for each step and will:

• Review each step, revising those requiring a change in sequence or tactic;

• Identify gaps in the documentation or analysis and how to address them;

• Update current documentation and create templates for future use;

• Draft a schedule for the ongoing maintenance of systems, e.g., annually or biannually;

• Identify issues or ideas that need additional discussion outside of the process review meeting; and

• Assign personnel to complete the tasks or duties that need to be carried out as identified.

6.3.2 Establish and implement processes for identifying issues and enhancements.

No data integration effort is perfect. Issues identified by end users or program staff should be reported to the helpdesk, or other central location, so they can be logged for resolution. Develop a process for identifying and addressing issues and potential enhancements. Users must have access to a system for recording issues or proposed enhancements.

Identify and document a secure process for submitting tickets and recommending enhancements, including the appropriate personnel to contact and how users will access and submit issues or suggestions. Identify and document how issues will be prioritized and addressed, including who is responsible for resolving the issue.

Review non-urgent issues and potential enhancements regularly. Identify a list of project team members and additional user group representatives who will meet annually or more frequently to review proposed

enhancements and identify additional enhancements through their work within the system. This process must align to the SEA data and IT governance processes.

6.3.3 Address the components of system sustainability.

Sustaining the integrated data system requires ongoing support. Identify a project team member(s) to develop a process for and manage implementation of the following:

- Review the use cases identified in Step 1 to ensure they still address stakeholder needs;
- Provide training(s) on how to access the integrated data and reports to ensure widespread system use;
- Ensure that proper funding is obtained to support the ongoing needs of the integrated system; and
- Develop evidence to demonstrate that the outcomes of the integrated system were worth the initial investments and justify sustaining the integrated system.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
6.3.1 Conduct process review meeting.	Project lead, project team	1-5 all day meetings
6.3.2 Establish and implement processes for identifying issues and enhancements.	Project lead, project team	2 all day meetings
6.3.3 Address the components of system sustainability.	Project lead, project team	Ongoing

What Are the Lessons Learned From Previous Efforts?

- Establish an issue and enhancement identification process shortly after integration is completed. All integration projects, no matter how well quality assurance efforts are executed, will contain residual issues that are not captured in testing. An already understood process for working out these remaining "bugs" after the official project is closed can speed up the correction, which will improve the overall user experience.
- Create templates for common tasks that occur during data integration to save effort in future integrations and assist staff members who may not have been involved in previous efforts.

Resources

- U.S. Department of Education: SLDS Sustainability Planning Guide

Contact Us: By email at CIIDTA@aemcorp.com.
Visit the CIID website for more information at CIIDTA.org.
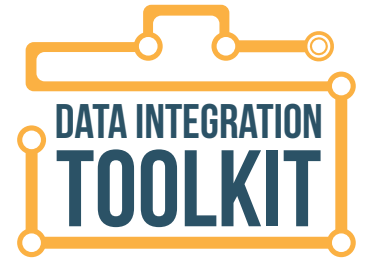
DATA INTEGRATION
TOOLKIT

Step 7: Conduct Post Integration Supports and Activities

During the final step, the team communicates the results of the data integration effort   in order to ensure long-term impact and continued success of the integration project. The team also performs a review of all documentation that was created during the active phases of the project, and creates a plan to continue any necessary tasks and for communicating with leadership. By completing this step, the project team informs all users of the new data system, how to access the integrated data, and how to sustain this effort moving forward using these tools and the documentation created during the integration process.

Task 7.1: Create and disseminate final calendar containing all processes and key dates for the newly integrated data system.

Task 7.2: Create documentation for users to understand how to access and maintain integrated data.

Task 7.3: Create and disseminate a final report and plan for ongoing communication with stakeholders.

CII
CENTER *for the*
INTEGRATION *of* **IDEA DATA**

Contact Us: By email at CIIDTA@aemcorp.com. Visit the CIID website for more information at CIIDTA.org.

CENTER *for the*
INTEGRATION *of* **IDEA DATA**

The CIID Data Integration Toolkit

DATA INTEGRATION
TOOLKIT

STEP ⑦ Conduct Post Integration Supports and Activities

Task 7.1: Create and disseminate a final calendar containing all processes and key dates for newly integrated data system.

Overall Description of the Task

Finalize and disseminate a unified calendar for the integrated system containing significant dates, processes, and timelines, to be distributed to all stakeholders, including internal and external users.

Why Is This Task Important?

- Establishing a common calendar of dates and processes at the outset will allow project members to set goals, organize the work, and manage the workflow and workload of the newly integrated system and existing data.
- Effective dissemination of the common calendar will allow the calendar to serve as a powerful communication tool for the project.
- The dissemination activity itself provides a platform for announcing the completion of the integration project, reinforcing earlier messages regarding benefits and sharing related resources.

Activities

7.1.1 Identify all new and changed processes and key dates affecting the calendar for the integrated system.

Review the processes documented in Task 6.3 to identify which changes need to be made in the calendar. Develop a document that lists new and/or changed deadlines and reporting periods and how they may affect data reporters and data users for the affected data elements. Determine the best way to incorporate the changes into the updated document. Document when data will be available to end users and assess any unexpected consequences.

7.1.2 Review and approve calendar changes.

Review the calendar changes with the project team. Discuss all major changes and their potential impact on stakeholder groups.

7.1.3 Create a new calendar or modify an existing one to document changed dates. Establish a process for maintaining the updated calendar.

Develop or update the calendar to allow all users to understand the new dates and processes. For dates and processes that require additional explanation, use a consistent format to further define or clarify the event (footnotes, comments that pop up with a "mouse- over," etc.). To help readers understand the changes, use the list approved in activity 7.1.2 to develop a separate document that details the changes made.

Once the calendar is updated, decide on a location for the calendar (internal intranet, internal network, or shared drive) as well as the permissions for viewing and editing the calendar. Multiple versions of the calendar may be required for different roles or organizations. Assign responsibility for calendar maintenance and determine an annual review process.

7.1.4 Align related resources.

Replace old versions of the calendar in locations where it is currently made available (websites, manuals, packets, internal networks) or referred to. Make necessary changes to introductory language or guidance that refers to the old calendar.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
7.1.1 Identify all new and changed processes and key dates affecting the calendar for the integrated system.	Project lead, project team	1-hour meeting
7.1.2 Review and approve calendar changes.	Project lead, project team	1.5-hour meeting
7.1.3 Create a new calendar or modify an existing one to document changed dates. Establish a process for maintaining the updated calendar	Project lead, project team	2-hour meeting
7.1.4 Align related resources.	Project lead, project team	1-hour meeting

What Are the Sustainability Considerations?

- Consider sending reminders throughout the year as collections come up to reinforce the new deadlines and processes. This will also help to keep team members engaged in the data collection and submission processes.

What Are the Lessons Learned From Previous Efforts?

- Recognize that there are dangers in waiting too long to create an updated calendar. It should be drafted in time to review the changes in dates and processes and the interaction with other state data system timelines. Until a calendar is clearly established, users may not embrace or support the integrated data system.

- Be thorough when compiling all necessary dates and processes prior to sharing it. The dangers of sending out an incomplete calendar include giving a false picture of the work, confusing staff, having a process that does not happen in a timely manner, and/or eroding confidence in the project. Include as many perspectives in the review process as possible.

Contact Us: By email at CIIDTA@aemcorp.com.
Visit the CIID website for more information at CIIDTA.org.

CENTER *for the*
INTEGRATION *of* **IDEA DATA**

DATA INTEGRATION
TOOLKIT

The CIID Data Integration Toolkit

STEP **7** **Conduct Post Integration Supports and Activities**

Task 7.2: Create documentation for users to understand how to access and maintain the integrated data system.

Overall Description of the Task

Create documentation of integrated datasets for current and future users to understand structure and flow of the data, how to access the data, and whom to contact for assistance.

Why Is This Task Important?

- This documentation will help users understand what data are available and how to access the newly integrated data.
- This document serves as a powerful tool to help users recognize the improved efficiencies, access, and capabilities within the newly integrated data system.
- This task also identifies an individual responsible for maintaining this document to ensure it is kept up to date and accurate.

Activities

7.2.1 Create system access documentation.

The same team that supported the calendar development within 7.1 should meet to create a draft or suggest modifications to existing documentation or diagrams. The team must include IT staff with a background in systems documentation as well as program staff. Collect and review examples of existing logic models and other states' data diagrams to inform the design. Determine the documentation's structure and perspective (organized by end user, data collection, system of origin). The documentation must include the following information:

- All incoming datasets;
- Specific access information/restrictions for each data system/set; and
- Data elements contained in the integrated system.

Document all incoming data sources, the data source steward and his/her contact information (office, email, phone), specific access information/restrictions, and the types of data contained within each module/collection. Determine other key pieces of information that end users will need to know, such as additional resources/user guides specific to the data.

7.2.2 Align related resources.

Replace the old documents (e.g., on websites, in manuals, in packets, on internal networks) as well as any references to those documents. Make necessary changes to introductory language or guidance where it is referenced.

7.2.3 Assign responsibility for ongoing system maintenance and documentation.

Using the original project plan, identify all roles that should be maintained and tasks that need completing. Consider all aspects of continuing the work including responsibility for monitoring requirement updates and ETL maintenance. Include assignments in the documentation as well as a schedule for task completion in the calendar.

The integration team should review documentation of the process including the project plan, ETL checklists, and calendars. Assign a position title as point of contact for all project documentation. This person will be responsible for knowing where information is stored and keeping all documentation up-to-date and accurate.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
7.2.1 Create system access documentation.	Project team, dataset owners	2 all-day meetings
7.2.2 Align related resources.	Project team	1 all-day meeting
7.2.3 Assign responsibility for ongoing system maintenance and documentation.	Project team	1 all-day meeting

## What Are the Sustainability Considerations?

- Clear documentation and record keeping is necessary to prevent situations where decisions need to be revisited or tasks re-done. Documentation also prevents loss of institutional knowledge as staff naturally turn over.

- Create a sustainability/project plan that outlines a schedule for post-integration activities and file submission (e.g., system upgrades, staff training, and leadership briefings).

What Are the Lessons Learned From Previous Efforts?

- Approach the task with a goal of providing simplicity and clarity about how to access the newly integrated data.

- Standardize a look for the integrated system by using common language, icons, and color schemes.

- Review documents, slides, and images prepared in earlier steps to ensure consistent use of language, icons, etc. in the documentation.

- Use an internal communications expert who specializes in infographics to help create the documentation so it depicts how to access the integrated data in an effective manner that is also visually appealing.

Contact Us: By email at CIIDTA@aemcorp.com.
Visit the CIID website for more information at CIIDTA.org.

CENTER *for the* INTEGRATION *of* IDEA DATA
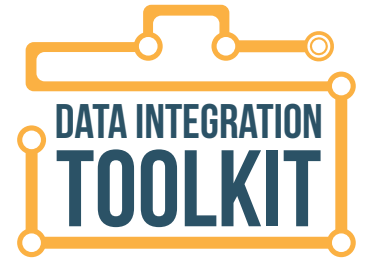
The CIID Data Integration Toolkit

Task 7.3: Create and disseminate a final report and plan for ongoing communication with stakeholders.

Overall Description of the Task

This task is designed to help communicate the impact of integration to all stakeholders and ensure they understand the changes made, efficiencies realized, and resources available to improve their understanding, such as an updated calendar or system documentation. Keep leadership informed about successes and ongoing needs of the integrated data system.

Why Is This Task Important?

- A key step in successful data integration is communicating the ways in which data integration affects related aspects of the work. This task conveys why integration occurred and how it affects multiple users, including those doing direct work on the project, funders, decision makers, and many others.
- People typically want to know how change will affect them and their work. Sharing a detailed summary that describes the data integration effort to external stakeholders will help convey the ways in which these stakeholders can expect to benefit from integration. Plans to keep stakeholders aware of ongoing success will maintain a high profile for integration and ensure continuity and growth.

Activities

7.3.1 List key aspects of integration efforts detailing the impact of the newly integrated system.

Gather a small group, including primary users and IT staff with knowledge of all data systems integrated, to create and identify all of the ways data integration will affect users and stakeholders. As a group, create a list of impacts and focus on efficiencies that have been realized, changes that have been made, things that have not changed, and new capabilities. This list should connect to the original use cases established in Step 1, describing how the needs articulated are now addressed within the integrated data system. Consider organizing the impacts around three categories of people, processes, and systems.

7.3.2 Draft communication materials that will form the basis for informing all users and stakeholders.

written content that can be used to communicate the impact and benefits of the integrated system, translating the content established in activity 7.2.1 into an accessible, easy-to-understand communication document(s), including a brief one- to two-page summary for quick reference and use.

Determine the format (PowerPoint presentations, written informational briefs, etc.) and method for organizing the materials (by user type, dataset, or established divisions of work). Create a draft for the report with the support of the project team, including both technical and program staff.

7.3.3 Identify intended audiences and develop dissemination plans.

uring a project team meeting, compile a list of users and other stakeholders to target for dissemination and the method of dissemination for each group. It is important to include known and potential data users in the state education agency, other state agencies, local educational agencies, and the public.

For each user or stakeholder group, identify a project team member to lead the plan development, material preparation, and outreach. Develop communication and dissemination plans and materials for each user or stakeholder group. Determine the ideal communication method, format, and focus for each group. Consider the group's use or interaction with the data and amount of information needed to ensure their understanding of the new integrated system. If needed, adjust the language and form of communication materials to meet the group's needs while ensuring the documentation is consistent with that created in activity 7.2.1.

Share adjusted materials with the project team to avoid duplication. Also determine the appropriate version of the updated calendar and data system documentation to include in the communication. Identify any timing factors specific to the group that would affect the communication schedule.

7.3.4 Review and implement dissemination plans.

During a project team meeting, review each dissemination plan developed in activity 7.3.3. Ensure each plan is appropriate for the audience and the version of the materials proposed. Once finalized, each designated individual will implement each dissemination plan, documenting the outreach and following up on questions or feedback provided.

7.3.5 Schedule regular discussions with stakeholders and agency leadership to confirm continued support for integrated data system.

These discussions, both informal and formal, serve to safeguard the progress that has been made by fostering high-level engagement and support. Share successes and milestones for the project. Bring evidence of increased efficiency, improved data quality or new ways to use data to keep the work at the forefront of discussions.

Include in the discussions financial and personnel requirements needed to maintain the work long-term.

What Is the Timeline for This Task?

Activities in the Task	Participants	Duration
7.3.1 List key aspects of integration efforts detailing the impact of the newly integrated system.	Project lead and project team	2-hour meeting
7.3.2 Draft communication materials that will form the basis for informing all users and stakeholders.	Project lead and project team	1 all-day meeting
7.3.3 Identify intended audiences and develop dissemination plans.	Project lead and designated project team member(s)	2 all-day meetings
7.3.4 Review and implement dissemination plans.	Project lead and designated project team member(s)	2 all-day meetings
7.3.5 Schedule regular discussions with stakeholders and agency leadership to confirm continued support for integration project.	Project lead	1-hour meeting

What Are the Sustainability Considerations?

State leadership have competing priorities. As such, the project manager should ensure the project remains visible to leadership and does not get lost in the shuffle. Having a plan to maintain the visibility and priority of the data integration project can help maintain needed funding and support to move forward.

What Are the Lessons Learned From Previous Efforts?

- Keep your materials consistent. It is very helpful to modify the materials to fit specific audiences; however, make sure the changes align with the original communication document developed in activity 7.3.2.

Contact Us: By email at CIIDTA@aemcorp.com.
Visit the CIID website for more information at CIIDTA.org.

CENTER *for the* INTEGRATION *of* IDEA DATA

IDEAs that Work
U.S. Office of Special
Education Programs

aem

TA&D
NETWORK